

УКРАЇНА

(19) UA (11) 116140 (13) C2

(51) МПК (2017.01)

G10L 19/038 (2013.01)

G10L 19/008 (2013.01)

H04R 5/00

**МІНІСТЕРСТВО
ЕКОНОМІЧНОГО
РОЗВИТКУ І ТОРГІВЛІ
УКРАЇНИ**

(12) ОПИС ДО ПАТЕНТУ НА ВІНАХІД

[illegible]

UA 116140 C2

(54) СТИСНЕННЯ РОЗКЛАДЕНИХ ПРЕДСТАВЛЕНЬ ЗВУКОВОГО ПОЛЯ

(57) Реферат:

Загалом, описані методи стиснення розкладеного представлення звукового поля. Пристрої, що містять один або більше процесорів, можуть бути виконані з можливістю здійснення методів. Один або більше процесорів можуть бути виконані з можливістю одержання бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[1] Дана заявка вимагає пріоритет по попередній патентній заявці США № 61/828,445, поданій 29 травня 2013 р., попередній патентній заявці США № 61/829,791, поданій 31 травня 2013 р., попередній патентній заявці США № 61/899,034 поданій 1 листопада 2013 р., попередній патентній заявці США № 61/899,041, поданій 1 листопада 2013 р., попередній патентній заявці США № 61/829,182, поданій 30 травня 2013 р., попередній патентній заявці США № 61/829,174, поданій 30 травня 2013 р., попередній патентній заявці США № 61/829,155, поданій 30 травня 2013 р., попередній патентній заявці США № 61/933,706 поданій 30 січня 2014 р., попередній патентній заявці США № 61/829,846, поданій 31 травня 2013 р., попередній патентній заявці США № 61/886,605, поданій 3 жовтня 2013 р., попередній патентній заявці США № 61/886,617, поданій 3 жовтня 2013 р., попередній патентній заявці США № 61/925,158, поданій 8 січня 2014 р., попередній патентній заявці США № 61/933,721, поданій 30 січня 2014 р., попередній патентній заявці США № 61/925,074, поданій 8 січня 2014 р., попередній патентній заявці США № 61/925,112, поданій 8 січня 2014 р., попередній патентній заявці США № 61/925,126, поданій 8 січня 2014 р., попередній патентній заявці США № 62/003,515, поданій 27 травня 2014 р., і попередній патентній заявці США № 61/828,615, поданій 29 травня 2013 р., зміст яких у повному обсязі включений в даний опис у порядку посилання.

Галузь техніки, до якої належить винахід

[2] Даний винахід стосується аудіоданих, зокрема стиснення аудіоданих.

Рівень техніки

[3] Сигнал амбіофонії більш високого порядку (HOA) (часто представлений множиною коефіцієнтів сферичних гармонік (SHC) або іншими ієрархічними елементами) є тривимірним представленням звукового поля. Це представлення HOA або SHC може представляти це звукове поле незалежно від локальної геометрії гучномовців, використовуваних для відтворення багатоканального аудіосигналу, рендеризованого з цього сигналу SHC. Цей сигнал SHC також може полегшувати зворотну сумісність, оскільки цей сигнал SHC можна рендеризувати у загальновідомі і широко розповсюджені багатоканальні формати, наприклад формат аудіоканалів 5.1 або формат аудіоканалів 7.1. Таким чином, представлення SHC може забезпечувати краще представлення звукового поля, яке також має зворотну сумісність.

Сутність винаходу

[4] У цілому, описані методи стиснення і зняття стиснення аудіоданих амбіофонії більш високого порядку.

[5] В одному аспекті, спосіб включає одержання одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів сферичних гармонік.

[6] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю визначення одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів сферичних гармонік.

[7] В іншому аспекті, пристрій містить засіб для одержання одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів сферичних гармонік, і засіб для збереження одного або більше перших векторів.

[8] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам одержувати один або більше перших векторів, що описують характерні компоненти звукового поля, і один або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів сферичних гармонік.

[9] В іншому аспекті, спосіб включає вибір однієї з множини схем зняття стиснення на основі указання, чи генерується стиснена версія коефіцієнтів сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, і зняття стиснення стисненої версії коефіцієнтів сферичних гармонік з використанням вибраної однієї з множини схем зняття стиснення.

[10] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю вибору однієї з множини схем зняття стиснення на основі указання, чи генерується стиснена

версія коефіцієнтів сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, і зняття стиснення стисненої версії коефіцієнтів сферичних гармонік з використанням вибраної однієї з множини схем зняття стиснення.

5 [11] В іншому аспекті, пристрій містить засіб для вибору однієї з множини схем зняття стиснення на основі указання, чи генерується стиснена версія коефіцієнтів сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, і засіб для зняття стиснення стисненої версії коефіцієнтів сферичних гармонік з використанням вибраної однієї з множини схем зняття стиснення.

10 [12] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам вбудованого пристрою декодування вибирати одну з множини схем зняття стиснення на основі указання, чи генерується стиснена версія коефіцієнтів сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, і знімати стиснення стисненої версії коефіцієнтів сферичних гармонік з використанням вибраної однієї з множини схем зняття стиснення.

15 [13] В іншому аспекті, спосіб включає одержання указання, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

[14] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю одержання указання, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

20 [15] В іншому аспекті, пристрій містить засіб для збереження коефіцієнтів сферичних гармонік, що представляють звукове поле, і засіб для одержання указання, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта.

25 [16] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам одержувати указання, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

[17] В іншому аспекті, спосіб включає квантування одного або більше перших векторів, що представляють один або більше компонентів звукового поля, і компенсацію помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля.

30 [18] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю квантування одного або більше перших векторів, що представляють один або більше компонентів звукового поля, і компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля.

35 [19] В іншому аспекті, пристрій містить засіб для квантування одного або більше перших векторів, що представляють один або більше компонентів звукового поля, і засіб для компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля.

40 [20] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам квантувати один або більше перших векторів, що представляють один або більше компонентів звукового поля, і компенсувати помилку, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля.

45 [21] В іншому аспекті, спосіб включає здійснення, на основі цільової бітової швидкості, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань, причому множина коефіцієнтів сферичних гармонік представляє звукове поле.

50 [22] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю здійснення, на основі цільової бітової швидкості, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань, причому множина коефіцієнтів сферичних гармонік представляє звукове поле.

55 [23] В іншому аспекті, пристрій містить засіб для збереження множини коефіцієнтів сферичних гармонік або їх розкладань і засіб для здійснення, на основі цільової бітової швидкості, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань, причому множина коефіцієнтів сферичних гармонік представляє звукове поле.

[24] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати, на основі цільової бітової швидкості, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань, причому множина коефіцієнтів сферичних гармонік представляє звукове поле.

[25] В іншому аспекті, спосіб включає одержання першого ненульового набору коефіцієнтів вектора, які представляють характерний компонент звукового поля, причому вектор розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[26] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю одержання першого ненульового набору коефіцієнтів вектора, що представляють характерний компонент звукового поля, причому вектор розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[27] В іншому аспекті, пристрій містить засіб для одержання першого ненульового набору коефіцієнтів вектора, що представляють характерний компонент звукового поля, причому вектор розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле, і засіб для збереження першого ненульового набору коефіцієнтів.

[28] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам визначати перший ненульовий набір коефіцієнтів вектора, що представляють характерний компонент звукового поля, причому вектор розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[29] В іншому аспекті, спосіб включає одержання, з бітового потоку, щонайменше одного з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, які рекомбінували з коефіцієнтами сферичних гармонік фону, причому коефіцієнти сферичних гармонік описують звукове поле, і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[30] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю визначення, з бітового потоку, щонайменше одного з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, які рекомбінували з коефіцієнтами сферичних гармонік фону, причому коефіцієнти сферичних гармонік описують звукове поле, і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[31] В іншому аспекті, пристрій містить засіб для одержання, з бітового потоку, щонайменше одного з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, які рекомбінували з коефіцієнтами сферичних гармонік фону, причому коефіцієнти сферичних гармонік описують звукове поле, і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[32] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам одержувати, з бітового потоку, щонайменше один з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, які рекомбінували з коефіцієнтами сферичних гармонік фону, причому коефіцієнти сферичних гармонік описують звукове поле, і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[33] В іншому аспекті, спосіб включає ідентифікацію одного або більше характерних аудіооб'єктів з одного або більше коефіцієнтів сферичних гармонік (SHC), пов'язаних з аудіооб'єктами, на основі направленості, визначеної для одного або більше з аудіооб'єктів.

[34] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю ідентифікації одного або більше характерних аудіооб'єктів з одного або більше коефіцієнтів сферичних гармонік (SHC), пов'язаних з аудіооб'єктами, на основі направленості, визначеної для одного або більше з аудіооб'єктів.

[35] В іншому аспекті, пристрій містить засіб для збереження одного або більше коефіцієнтів сферичних гармонік (SHC) і засіб для ідентифікації одного або більше характерних аудіооб'єктів з одного або більше коефіцієнтів сферичних гармонік (SHC), пов'язаних з аудіооб'єктами, на основі направленості, визначеної для одного або більше з аудіооб'єктів.

[36] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам ідентифікувати один або більше характерних аудіооб'єктів з одного або більше коефіцієнтів

сферичних гармонік (SHC), пов'язаних з аудіооб'єктами, на основі направленості, визначеної для одного або більше з аудіооб'єктів.

[37] В іншому аспекті, спосіб включає здійснення синтезу на векторній основі відносно множини коефіцієнтів сферичних гармонік для генерації розкладених представлень з множини коефіцієнтів сферичних гармонік, що представляють один або більше аудіооб'єктів і відповідну інформацію направленості, причому коефіцієнти сферичних гармонік пов'язані з порядком і описують звукове поле, визначення окремої і фонової інформації направленості з інформації направленості, зниження порядку інформації направленості, пов'язаної з фоновими аудіооб'єктами, для генерації перетвореної інформації направленості фону, застосування компенсації для збільшення значень перетвореної інформації направленості для збереження повної енергії звукового поля.

[38] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю здійснення синтезу на векторній основі відносно множини коефіцієнтів сферичних гармонік для генерації розкладених представлень з множини коефіцієнтів сферичних гармонік, що представляють один або більше аудіооб'єктів і відповідну інформацію направленості, причому коефіцієнти сферичних гармонік пов'язані з порядком і описують звукове поле, визначення окремої і фонової інформації направленості з інформації направленості, зниження порядку інформації направленості, пов'язаної з фоновими аудіооб'єктами, для генерації перетвореної інформації направленості фону, застосування компенсації для збільшення значень перетвореної інформації направленості для збереження повної енергії звукового поля.

[39] В іншому аспекті, пристрій містить засіб для здійснення синтезу на векторній основі відносно множини коефіцієнтів сферичних гармонік для генерації розкладених представлень з множини коефіцієнтів сферичних гармонік, що представляють один або більше аудіооб'єктів і відповідну інформацію направленості, причому коефіцієнти сферичних гармонік пов'язані з порядком і описують звукове поле, засіб для визначення окремої і фонової інформації направленості з інформації направленості, засіб для зниження порядку інформації направленості, пов'язаної з фоновими аудіооб'єктами, для генерації перетвореної інформації направленості фону, і засіб для застосування компенсації для збільшення значень перетвореної інформації направленості для збереження повної енергії звукового поля.

[40] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати синтез на векторній основі відносно множини коефіцієнтів сферичних гармонік для генерації розкладених представлень з множини коефіцієнтів сферичних гармонік, що представляють один або більше аудіооб'єктів і відповідну інформацію направленості, причому коефіцієнти сферичних гармонік пов'язані з порядком і описують звукове поле, визначати окрему і фонову інформацію направленості з інформації направленості, знижувати порядок інформації направленості, пов'язаної з фоновими аудіооб'єктами, для генерації перетвореної інформації направленості фону, і застосовувати компенсацію для збільшення значень перетвореної інформації направленості для збереження повної енергії звукового поля.

[41] В іншому аспекті, спосіб включає одержання розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, щонайменше частково, шляхом здійснення інтерполяції відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік.

[42] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю одержання розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, щонайменше частково, шляхом здійснення інтерполяції відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік.

[43] В іншому аспекті, пристрій містить засіб для збереження першої множини коефіцієнтів сферичних гармонік і другої множини коефіцієнтів сферичних гармонік і засіб для одержання розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, щонайменше частково, шляхом здійснення інтерполяції відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік.

[44] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам одержувати розкладені інтерпольовані коефіцієнти сферичних гармонік для відрізка часу, щонайменше частково, шляхом здійснення інтерполяції відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік.

основі порядку просторового компонента відносно просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

5 [58] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю ідентифікації кодової книги Хаффмана для використання при стисненні просторового компонента з множини просторових компонентів на основі порядку просторового компонента відносно просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

10 [59] В іншому аспекті, пристрій містить засіб для збереження кодової книги Хаффмана і засіб для ідентифікації кодової книги Хаффмана для використання при стисненні просторового компонента з множини просторових компонентів на основі порядку просторового компонента відносно просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

15 [60] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам ідентифікувати кодову книгу Хаффмана для використання при стисненні просторового компонента з множини просторових компонентів на основі порядку просторового компонента відносно просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

20 [61] В іншому аспекті, спосіб включає визначення розміру кроку квантування, що підлягає використанню при стисненні просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

25 [62] В іншому аспекті, пристрій містить один або більше процесорів, виконаних з можливістю визначення розміру кроку квантування, що підлягає використанню при стисненні просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

30 [63] В іншому аспекті, пристрій містить засіб для визначення розміру кроку квантування, що підлягає використанню при стисненні просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік, і засіб для збереження розміру кроку квантування.

35 [64] В іншому аспекті, на довгостроковому зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам визначати розміру кроку квантування, що підлягає використанню при стисненні просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

40 [65] Деталі одного або більше аспектів методів викладені в прикладених кресленнях і нижченаведеному описі. Інші ознаки, задачі і переваги цих методів стають зрозумілі з опису і креслень, а також з формули винаходу.

Короткий опис креслень

45 [66] Фіг. 1 і 2 - схеми, що демонструють сферичні гармонічні базисні функції різних порядків і підпорядків.

[67] Фіг. 3 - схема, що демонструє систему, яка може здійснювати різні аспекти методів, описаних у цьому винаході.

50 [68] Фіг. 4 - блок-схема, що демонструє, більш детально, один приклад пристрою аудіокодування, показаного в прикладі, наведеному на фіг. 3, який може здійснювати різні аспекти методів, описаних у цьому винаході.

[69] Фіг. 5 - блок-схема, що більш детально демонструє пристрій аудіодекодування, показаний на фіг. 3.

[70] Фіг. 6 - блок-схема операцій, що демонструє роботу блока аналізу контенту пристрою аудіокодування при здійсненні різних аспектів методів, описаних у цьому винаході.

55 [71] Фіг. 7 - блок-схема операцій, що демонструє роботу пристрою аудіокодування при здійсненні різних аспектів методів синтезу на векторній основі, описаних у цьому винаході.

[72] Фіг. 8 - блок-схема операцій, що демонструє роботу пристрою аудіодекодування при здійсненні різних аспектів методів, описаних у цьому винаході.

60 [73] Фіг. 9A-9L - блок-схеми, що більш детально демонструють різні аспекти пристрою аудіокодування прикладу, наведеного на фіг. 4.

[74] Фіг. 10A-10O(ii) - схеми, що демонструють частину бітового потоку або інформації побічного каналу, яка може більш детально вказувати стиснені просторові компоненти.

[75] Фіг. 11A-11G - блок-схеми, що демонструють, більш детально, різні блоки пристрою аудіодекодування, показаного в прикладі, наведеному на фіг. 5.

5 [76] Фіг. 12 - схема, що демонструє приклад аудіо-екосистеми, яка може здійснювати різні аспекти методів, описаних у цьому винаході.

[77] Фіг. 13 - схема, що більш детально демонструє один приклад аудіо-екосистеми, показаної на фіг. 12.

10 [78] Фіг. 14 - схема, що більш детально демонструє один приклад аудіо-екосистеми, показаної на фіг. 12.

[79] Фіг. 15A і 15B - схеми, що більш детально демонструють інші приклади аудіо-екосистеми, показаної на фіг. 12.

[80] Фіг. 16 - схема, що демонструє приклад пристрою аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході.

15 [81] Фіг. 17 - схема, що більш детально демонструє один приклад пристрою аудіокодування, показаного на фіг. 16.

[82] Фіг. 18 - схема, що демонструє приклад пристрою аудіодекодування, який може здійснювати різні аспекти методів, описаних у цьому винаході.

20 [83] Фіг. 19 - схема, що більш детально демонструє один приклад пристрою аудіодекодування, показаного на фіг. 18.

[84] Фіг. 20A-20G - схеми, що демонструють приклад пристроїв захоплення аудіосигналу, які можуть здійснювати різні аспекти методів, описаних у цьому винаході.

[85] Фіг. 21A-21E - схеми, що демонструють приклад пристроїв відтворення аудіосигналу, які можуть здійснювати різні аспекти методів, описаних у цьому винаході.

25 [86] Фіг. 22A-22H - схеми, що демонструють приклад середовищ відтворення аудіосигналу відповідно до одного або більше методів, описаних в цьому винаході.

[87] Фіг. 23 - схема, що демонструє приклад випадку використання, коли користувач може випробувати 3D звукове поле спортивної гри при носінні головних телефонів відповідно до одного або більше методів, описаних в цьому винаході.

30 [88] Фіг. 24 - схема, що демонструє спортивний стадіон, де 3D звукове поле може записуватися відповідно до одного або більше методів, описаних в цьому винаході.

[89] Фіг. 25 - блок-схема операцій, що демонструє метод рендеризації 3D звукового поля на основі локального аудіоландшафту відповідно до одного або більше методів, описаних в цьому винаході.

35 [90] Фіг. 26 - схема, що демонструє приклад ігрової студії відповідно до одного або більше методів, описаних в цьому винаході.

[91] Фіг. 27 - схема, що демонструє множину ігрових систем, які включають в себе машини рендеризації, відповідно до одного або більше методів, описаних в цьому винаході.

40 [92] Фіг. 28 - схема, що демонструє конфігурацію гучномовців, яку можна імітувати головними телефонами, відповідно до одного або більше методів, описаних в цьому винаході.

[93] Фіг. 29 - схема, що демонструє множину мобільних пристроїв, які можна використовувати для захоплення і/або редагування 3D звукового поля відповідно до одного або більше методів, описаних в цьому винаході.

45 [94] Фіг. 30 - схема, що демонструє кадр відео, пов'язаний з 3D звуковим полем, який може оброблятися відповідно до одного або більше методів, описаних в цьому винаході.

[95] Фіг. 31A-31M - схеми, що демонструють графіки, де показані різні результати імітації здійснення синтетичної або записаної категоризації звукового поля відповідно до різних аспектів методів, описаних у цьому винаході.

50 [96] Фіг. 32 - схема, що демонструє графік сингулярних значень з матриці S, розкладеної з коефіцієнтів амбіофонії більш високого порядку, відповідно до методів, описаних в цьому винаході.

[97] Фіг. 33A і 33B - схеми, що демонструють відповідні графіки, де показаний потенційний вплив переупорядкування при кодуванні векторів, що описують компоненти переднього плану звукового поля відповідно до методів, описаних в цьому винаході.

55 [98] Фіг. 34 і 35 - принципові схеми, що демонструють відмінності між ідентифікацією характерних аудіооб'єктів тільки на основі енергії і на основі направленості, відповідно до цього розкриття.

60 [99] Фіг. 36A-36G - схеми, що демонструють проекції щонайменше частини розкладеної версії коефіцієнтів сферичних гармонік у просторову область для здійснення інтерполяції відповідно до різних аспектів методів, описаних у цьому винаході.

[100] Фіг. 37 демонструє представлення описаних тут методів одержання просторово-часової інтерполяції.

[101] Фіг. 38 - блок-схема, що демонструє штучні матриці US , US_1 і US_2 для послідовних блоків SVD для багатовимірного сигналу відповідно до описаних тут методів.

5 [102] Фіг. 39 - блок-схема, що демонструє розкладання наступних кадрів сигналу амбіофонії більш високого порядку (HOA) з використанням розкладання на сингулярні значення і згладжування просторово-часових компонентів відповідно до методів, описаних у цьому винаході.

10 [103] Фіг. 40A-40J показані блок-схеми, що демонструють приклади пристроїв аудіокодування, які можуть здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля.

[104] Фіг. 41A-41D - блок-схеми, що демонструють ілюстративний пристрій аудіодекодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для декодування коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля.

15 [105] Фіг. 42A-42C - блок-схеми, що демонструють блок зниження порядку, показаний в прикладах, наведених на фіг. 40B-40J більш детально.

[106] Фіг. 43 - схема, що більш детально демонструє блок стиснення V , показаний на фіг. 40I.

20 [107] Фіг. 44 - схема, що демонструє ілюстративні операції, здійснювані пристроєм аудіокодування для компенсації помилки квантування відповідно до різних аспектів методів, описаних у цьому винаході.

[108] Фіг. 45A і 45B - схеми, що демонструють інтерполяцію підкадрів з ділянок двох кадрів відповідно до різних аспектів методів, описаних у цьому винаході.

25 [109] Фіг. 46A-46E - схеми, що демонструють поперечний переріз проекції одного або більше векторів розкладеної версії множини коефіцієнтів сферичних гармонік, інтерпольованих відповідно до методів, описаних в цьому винаході.

[110] Фіг. 47 - блок-схема, що демонструє, більш детально, блок витягання пристроїв аудіодекодування, показаних у прикладах, наведених на фіг. 41A-41D.

30 [111] Фіг. 48 - блок-схема, що більш детально демонструє блок аудіорендеризації пристрою аудіодекодування, показаного в прикладах, наведених на фіг. 41A-41D.

[112] Фіг. 49A-49E(ii) - схеми, що демонструють відповідні системи аудіокодування, які можуть реалізувати різні аспекти методів, описаних у цьому винаході.

35 [113] Фіг. 50A і 50B - блок-схеми, що демонструють один із двох різних підходів для потенційного зниження порядку фонового контенту відповідно до методів, описаних в цьому винаході.

[114] Фіг. 51 - блок-схема, що демонструє приклади тракту стиснення характерних компонентів пристрою аудіокодування, який може реалізувати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік.

40 [115] Фіг. 52 - блок-схема, що демонструє інший приклад пристрою аудіодекодування, який може реалізувати різні аспекти методів, описаних у цьому винаході, для реконструкції або зразкової реконструкції коефіцієнтів сферичних гармонік (SHC).

[116] Фіг. 53 - блок-схема, що демонструє інший приклад пристрою аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході.

45 [117] Фіг. 54 - блок-схема, що демонструє, більш детально, ілюстративну реалізацію пристрою аудіокодування, показаного в прикладі, наведеному на фіг. 53.

[118] Фіг. 55A і 55B - схеми, що демонструють приклад здійснення різних аспектів методів, описаних у цьому винаході, для повороту звукового поля.

50 [119] Фіг. 56 - схема, що демонструє приклад звукового поля, захопленого відносно першої системи координат, яке потім повертається відповідно до методів, описаних в цьому винаході, для вираження звукового поля відносно другої системи координат.

[120] Фіг. 57A-57E - схеми, що демонструють бітові потоки, сформовані відповідно до методів, описаних в цьому винаході.

55 [121] Фіг. 58 - блок-схема операцій, що демонструє приклад роботи пристрою аудіокодування, показаного в прикладі, наведеному на фіг. 53, при реалізації аспектів повороту методів, описаних у цьому винаході.

[122] Фіг. 59 - блок-схема операцій, що демонструє приклад роботи пристрою аудіокодування, показаного в прикладі, наведеному на фіг. 53, при здійсненні аспектів перетворення методів, описаних у цьому винаході.

Докладний опис

[123] Розвиток об'ємного звуку зробив доступними багато які вихідні формати для розваги в наш час. Приклади таких споживацьких форматів об'ємного звуку, більшою частиною, є основаними на 'каналах', оскільки вони неявно вказують підведення до гучномовців у визначених геометричних координатах. Вони включають в себе популярний формат 5.1 (який

5 включає в себе наступні шість каналів: передній лівий (FL), передній правий (FR), центральний або передній центральний, задній лівий або лівий оточення, задній правий або правий оточення і низькочастотні ефекти (LFE)), перспективний формат 7.1, різні формати, що включають в себе висотні гучномовці, наприклад формат 7.1.4 і формат 22.2 (наприклад, для використання зі стандартом телебачення надвисокої чіткості). Неспоживацькі формати можуть займати будь-яку

10 кількість гучномовців (у симетричних і несиметричних геометріях), що часто іменуються 'решітками оточення'. Один приклад такої решітки включає в себе 32 гучномовці, розташовані в координатах кутів зрізаного ікосаедра.

[124] Вхід у майбутній кодер MPEG, у необов'язковому порядку, є одним із трьох можливих форматів: (i) традиційний аудіосигнал на каналній основі (як розглянуто вище), що означає відтворення через гучномовці в попередньо зазначених позиціях; (ii) аудіосигнал на основі об'єктів, що передбачає дискретну імпульсно-кодovu модуляцію (IKM) даних для одиничних аудіооб'єктів з відповідними метаданими, що містять координати їх положення (крім іншої інформації); і (iii) аудіосигнал на основі сцени, що передбачає представлення звукового поля з використанням коефіцієнтів сферичних гармонічних базисних функцій (також іменованих

20 "коефіцієнтами сферичних гармонік" або SHC, "амбіофонією більш високого порядку" або HOA і "коефіцієнтами HOA"). Цей майбутній кодер MPEG описаний більш детально в документі під назвою "Call for Proposals for 3D Audio", виданому Міжнародною організацією по стандартизації/Міжнародною електротехнічною комісією (ISO)/(IEC) JTC1/SC29/WG11/N13411, січень 2013 р. Женева, Швейцарія, і доступному за адресою

25 <http://mpeg.chiariglione.org/sites/default/files/files/standards/parts/docs/w13411.zip>.

[125] На ринку присутні різні формати на каналній основі 'оточуючого звуку'. Вони простягаються, наприклад, від 5.1 для системи домашнього кінотеатру (який був найбільш успішним відносно поширення в житлових приміщеннях після стерео) до системи 22.2, розробленої NHK (Nippon Hoso Kyokai або Japan Broadcasting Corporation). Творці контенту (наприклад, голлівудські студії), імовірно, будуть створювати звукову доріжку для кінофільму один раз, замість того, щоб витрачати зусилля для її реміксу для кожної конфігурації гучномовців. Нещодавно, організації по розробці стандартів розглянули можливості забезпечення кодування в стандартизований бітовий потік і наступне декодування, що адаптується і не залежить від геометрії (і кількості) гучномовців і акустичних умов у положенні

35 відтворення (з використанням рендеризатора).

[126] Для забезпечення такої гнучкості для творців контенту, можна використовувати ієрархічний набір елементів для представлення звукового поля. Ієрархічний набір елементів може означати набір елементів, у якому елементи упорядковані таким чином, що базовий набір елементів більш низького порядку забезпечує повне представлення модельованого звукового поля. Коли набір розширюється для включення елементів більш високого порядку, представлення стає більш деталізованим і розрізнення збільшується.

40

[127] Одним прикладом ієрархічного набору елементів є набір коефіцієнтів сферичних гармонік (SHC). Наступне вираз демонструє опис або представлення звукового поля з використанням SHC:

45

$$p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t},$$

[128] цей вираз показує, що тиск p_i у будь-якій точці $\{r_r, \theta_r, \varphi_r\}$ звукового поля, у момент часу t , можна однозначно представити за допомогою SHC, $A_n^m(k)$. Тут, $k = \frac{\omega}{c}$, c - швидкість

50 звуку (~ 343 м/с), $\{r_r, \theta_r, \varphi_r\}$ - точка відліку (або точка спостереження), $j_n(\cdot)$ - сферична

бесселева функція порядку n , і $Y_n^m(\theta_r, \varphi_r)$ - сферичні гармонічні базисні функції порядку n і підпорядку m . Ясно, що член у квадратних дужках є представленням у частотній області сигналу (тобто $S(\omega, r_r, \theta_r, \varphi_r)$), який можна апроксимувати різними частотно-часовими

перетвореннями, наприклад дискретним перетворенням Фур'є (DFT), дискретним косинусним перетворенням (DCT) або вейвлетним перетворенням. Інші приклади ієрархічних наборів включають в себе набори коефіцієнтів вейвлетного перетворення й інші набори коефіцієнтів базисних функцій з різними розрізненнями.

5 [129] На фіг. 1 показана схема, що демонструє сферичні гармонічні базисні функції від нульового порядку ($n=0$) до четвертого порядку ($n=4$). Як можна бачити, для кожного порядку, існує розширення підпорядків m , що показані, але явно не зазначені в прикладі, наведеному на фіг. 1, для спрощення ілюстрації.

10 [130] На фіг. 2 показана інша схема, що демонструє сферичні гармонічні базисні функції від нульового порядку ($n=0$) до четвертого порядку ($n=4$). На фіг. 2, сферичні гармонічні базисні функції показані в тривимірному координатному просторі, де показані порядок і підпорядок.

[131] SHC $A_n^m(k)$ можуть або фізично захоплюватися (наприклад, записуватися) різними конфігураціями мікрофонної решітки, або, альтернативно, можуть виводитися з описів звукового поля на каналній основі або на основі об'єктів. SHC представляють аудіосигнал на основі сцени, де SHC може вводитися на аудіокодер для одержання кодованого SHC, що може сприяти більш ефективній передачі або збереженню. Наприклад, можна використовувати представлення четвертого порядку, що передбачає $(1+4)^2$ (25, і, таким чином, четвертого порядку) коефіцієнтів.

20 [132] Як згадано вище, SHC можна виводити з мікрофонного запису з використанням мікрофона. Різні приклади, як SHC можна виводити з мікрофонних решіток, описані в Poletti M., "Three-Dimensional Surround Sound Systems Based on Spherical Harmonics", J. Audio Eng. Soc., Vol. 53, № 11, 2005 November, pp. 1004-1025.

[133] Для ілюстрації, як ці SHC можна виводити з опису на основі об'єктів, розглянемо наступне рівняння. Коефіцієнти $A_n^m(k)$ для звукового поля, що відповідає окремому аудіооб'єкту, можна виразити у вигляді:

$$A_n^m(k) = g(\omega)(-4\pi i k) h_n^{(2)}(k r_s) Y_n^{m*}(\theta_s, \varphi_s),$$

де i позначає $\sqrt{-1}$, $h_n^{(2)}(\cdot)$ - сферична функція Ханкеля (другого роду) порядку n , і

30 $\{r_s, \theta_s, \varphi_s\}$ - положення об'єкта. Знання енергії $g(\omega)$ джерела об'єкта як функції частоти (наприклад, з використанням методів частотно-часового аналізу, наприклад, здійснюючи швидке перетворення Фур'є на потоці IKM) дозволяє перетворювати кожен об'єкт IKM і його положення в SHC $A_n^m(k)$. Крім того, можна показати (оскільки вищенаведене є лінійним і ортогональним розкладанням), що коефіцієнти $A_n^m(k)$ для кожного об'єкта адитивні. Таким чином, велику кількість об'єктів IKM можна представити коефіцієнтами $A_n^m(k)$ (наприклад, як суму векторів коефіцієнтів для окремих об'єктів). Власне кажучи, ці коефіцієнти містять інформацію про звукове поле (тиск як функцію 3D координат), і вищенаведене представляє перетворення з окремих об'єктів у представлення загального звукового поля, поблизу точки спостереження $\{r_r, \theta_r, \varphi_r\}$. Фігури, що залишилися, описані нижче в контексті аудіокодування на основі об'єктів і на основі SHC.

40 [134] На фіг. 3 показана схема, що демонструє систему 10, що може здійснювати різні аспекти методів, описаних у цьому винаході. Як показано в прикладі, наведеному на фіг. 3, система 10 включає в себе творця 12 контенту і споживача 14 контенту. Хоча вони описані в контексті творця 12 контенту і споживача 14 контенту, методи можна реалізувати в будь-якому контексті, у якому SHC (які також можуть іменуватися коефіцієнтами HOA) або будь-яке інше ієрархічне представлення звукового поля кодується для формування бітового потоку, що представляє аудіодані. Крім того, творець 12 контенту може представляти будь-яку форму обчислювального пристрою, здатного реалізувати методи, описані в цьому винаході, у тому числі телефонну трубку (або стільниковий телефон), планшетний комп'ютер, смартфон або настільний комп'ютер для забезпечення декількох прикладів. Аналогічно, споживач 14 контенту може представляти будь-яку форму обчислювального пристрою, здатного реалізувати методи, описані в цьому винаході, у тому числі телефонну трубку (або стільниковий телефон),

планшетний комп'ютер, смартфон, телевізійну приставку або настільний комп'ютер для забезпечення декількох прикладів.

[135] Творець 12 контенту може представляти кіностудію або інший суб'єкт, що може генерувати багатоканальний аудіоконтент для споживання споживачами контенту, наприклад споживачем 14 контенту. У деяких прикладах, творець 12 контенту може представляти окремого користувача, що бажає стиснути коефіцієнти 11 НОА. Часто, цей творець контенту генерує аудіоконтент спільно з відеоконтентом. Споживач 14 контенту представляє фізичну особу, яка володіє або має доступ до системи відтворення аудіосигналу, що може означати будь-яку форму системи відтворення аудіосигналу, здатної рендеризувати SHC для відтворення як багатоканального аудіоконтенту. У прикладі, наведеному на фіг. 3, споживач 14 контенту включає в себе систему 16 відтворення аудіосигналу.

[136] Творець 12 контенту включає в себе систему 18 аудіоредагування. Творець 12 контенту може одержувати живі записи 7 у різних форматах (у тому числі безпосередньо як коефіцієнти НОА) і аудіооб'єкти 9, які творець 12 контенту може редагувати з використанням системи 18 аудіоредагування. Творець 12 контенту може, при виконанні процесу редагування, рендеризувати коефіцієнти 11 НОА з аудіооб'єктів 9, прослуховуючи рендеризовані підведення гучномовця в спробі ідентифікації різних аспектів звукового поля, що вимагають додаткового редагування. Потім творець 12 контенту може редагувати коефіцієнти 11 НОА (потенційно опосередковано за допомогою маніпуляції різними аудіооб'єктами 9, з яких можна виводити вихідні коефіцієнти НОА вищеописаним чином). Творець 12 контенту може використовувати систему 18 аудіоредагування для генерації коефіцієнтів 11 НОА. Система 18 аудіоредагування представляє будь-яку систему, здатну редагувати аудіодані і виводити ці аудіодані як один або більше вихідних коефіцієнтів сферичних гармонік.

[137] Після завершення процесу редагування, творець 12 контенту може генерувати бітовий потік 21 на основі коефіцієнтів 11 НОА. Таким чином, творець 12 контенту включає в себе пристрій 20 аудіокодування, що представляє пристрій, виконаний з можливістю кодування або іншого стиснення коефіцієнтів 11 НОА відповідно до різних аспектів методів, описаних у цьому винаході, для генерації бітового потоку 21. Пристрій 20 аудіокодування може генерувати бітовий потік 21 для передачі, у порядку одного прикладу, по каналу передачі, який може бути дротовим або бездротовим каналом, пристроєм збереження даних і т. п. Бітовий потік 21 може представляти кодовану версію коефіцієнтів 11 НОА і може включати в себе первинний бітовий потік і інший побічний бітовий потік, які можуть іменуватися інформацією побічного каналу.

[138] Хоча більш докладно описаний нижче, пристрій 20 аудіокодування може бути виконаний з можливістю кодування коефіцієнтів 11 НОА на основі синтезу на векторній основі або синтезу на основі напрямку. Для визначення, чи здійснювати методологію синтезу на векторній основі або методологію синтезу на основі напрямку, пристрій 20 аудіокодування може визначати, щонайменше частково, на основі коефіцієнтів 11 НОА, чи згенеровані коефіцієнти 11 НОА за допомогою природного запису звукового поля (наприклад, живого запису 7) або створені штучно (тобто синтетично), у порядку одного прикладу, з аудіооб'єктів 9, наприклад об'єкта ІКМ. Коли коефіцієнти 11 НОА згенеровані з аудіооб'єктів 9, пристрій 20 аудіокодування може кодувати коефіцієнти 11 НОА з використанням методології синтезу на основі напрямку. Коли коефіцієнти 11 НОА захоплені наживо з використанням, наприклад, власного мікрофона, пристрій 20 аудіокодування може кодувати коефіцієнти 11 НОА на основі методології синтезу на векторній основі. Вищеописана відмінність представляє один приклад, де може бути розгорнута методологія синтезу на векторній основі або на основі напрямку. Можуть існувати інші випадки, коли одна або обидві можуть бути корисні для природних записів, штучно генерованого контенту або їх суміші (гібридного контенту). Крім того, можна також використовувати обидві методології одночасно для кодування одиничного часового кадру коефіцієнтів НОА.

[139] Передбачаючи з метою ілюстрації, що пристрій 20 аудіокодування визначає, що коефіцієнти 11 НОА захоплені наживо або інакше представляють живі записи, наприклад живий запис 7, пристрій 20 аудіокодування може бути виконаний з можливістю кодування коефіцієнтів 11 НОА з використанням методології синтезу на векторній основі, що передбачає застосування лінійного оборотного перетворення (LIT). Один приклад лінійного оборотного перетворення іменується "розкладанням на сингулярні значення" (або "SVD"). У цьому прикладі, пристрій 20 аудіокодування може застосовувати SVD до коефіцієнтів 11 НОА для визначення розкладеної версії коефіцієнтів 11 НОА. Потім пристрій 20 аудіокодування може аналізувати розкладену версію коефіцієнтів 11 НОА для ідентифікації різних параметрів, що може полегшувати переупорядкування розкладеної версії коефіцієнтів 11 НОА. Потім пристрій 20 аудіокодування може переупорядковувати розкладену версію коефіцієнтів 11 НОА на основі ідентифікованих параметрів, де таке переупорядкування, як описано більш докладно нижче,

може поліпшувати ефективність кодування, за умови, що перетворення може переупорядковувати коефіцієнти HOA по кадрах коефіцієнтів HOA (де кадр звичайно включає в себе M вибірок коефіцієнтів 11 HOA і M, у деяких прикладах, задане дорівнюючим 1024). Після переупорядкування розкладеної версії коефіцієнтів 11 HOA, пристрій 20 аудіокодування може

5 вибирати ті з розкладеної версії коефіцієнтів 11 HOA, що представляють передньопланові (або, іншими словами, характерні, переважні або помітні) компоненти звукового поля. Пристрій 20 аудіокодування може вказувати розкладену версію коефіцієнтів 11 HOA, що представляє компоненти переднього плану, як аудіооб'єкт і відповідну інформацію направленості.

10 [140] Пристрій 20 аудіокодування також може здійснювати аналіз звукового поля відносно коефіцієнтів 11 HOA по порядку, щонайменше частково, для ідентифікації з коефіцієнтів 11 HOA, що представляють один або більше компонентів фону (або, іншими словами, оточення) звукового поля. Пристрій 20 аудіокодування може здійснювати компенсацію енергії відносно фонових компонентів, за умови, що, у деяких прикладах, фонові компоненти можуть включати в себе тільки піднабір будь-якої даної вибірки коефіцієнтів 11 HOA (наприклад, що відповідають

15 сферичним базисним функціям нульового і першого порядку, а не відповідним сферичним базисним функціям другого або більш високого порядку). При здійсненні зниження порядку, іншими словами, пристрій 20 аудіокодування може доповнювати (наприклад, додавати/віднімати енергію) до/з коефіцієнтів HOA, що залишилися, фону коефіцієнтів 11 HOA для компенсації зміни повної енергії, яка обумовлена здійсненням зниження порядку.

20 [141] Потім пристрій 20 аудіокодування може здійснювати форму психоакустичного кодування (наприклад, MPEG surround, MPEG-AAC, MPEG-USAC або інші відомі форми психоакустичного кодування) відносно кожного з коефіцієнтів 11 HOA, що представляють фонові компоненти, і кожного з аудіооб'єктів переднього плану. Пристрій 20 аудіокодування може здійснювати форму інтерполяції відносно інформації направленості переднього плану і

25 потім здійснювати зниження порядку відносно інтерпольованої інформації направленості переднього плану для генерації інформації направленості переднього плану зниженого порядку. Пристрій 20 аудіокодування може додатково здійснювати, у деяких прикладах, квантування відносно інформації направленості переднього плану зниженого порядку, виводячи кодовану інформацію направленості переднього плану. У ряді випадків, це квантування може містити скалярне/ентропійне квантування. Потім пристрій 20 аудіокодування може формувати бітовий

30 потік 21 для включення кодованих фонових компонентів, кодованих аудіооб'єктів переднього плану і квантованої інформації направленості. Потім пристрій 20 аудіокодування може передавати або інакше виводити бітовий потік 21 споживачу 14 контенту.

35 [142] Хоча на фіг. 3 показана безпосередня передача споживачу 14 контенту, творець 12 контенту може виводити бітовий потік 21 на проміжний пристрій, розташований між творцем 12 контенту і споживачем 14 контенту. Цей проміжний пристрій може зберігати бітовий потік 21 для доставки у подальшому споживачу 14 контенту, таким чином, можуть запитувати цей бітовий потік. Проміжний пристрій може містити файловий сервер, веб-сервер, настільний комп'ютер, портативний комп'ютер, планшетний комп'ютер, мобільний телефон, смартфон або будь-який

40 інший пристрій, здатний зберігати бітовий потік 21 для витягання у подальшому аудіодекодером. Цей проміжний пристрій може розташовуватися в мережі доставки контенту, здатний до потокової передачі бітового потоку 21 (і, можливо, спільно з передачею відповідного бітового потоку відеоданих) абонентам, наприклад споживачу 14 контенту, який запитує бітовий потік 21.

45 [143] Альтернативно, творець 12 контенту може зберігати бітовий потік 21 на носії даних, наприклад компакт-диск, цифровий відеодиск, диск відео високої чіткості або інші носії даних, більшість з яких можуть зчитуватися комп'ютером і, таким чином, можуть іменуватися зчитувані комп'ютером носії даних або довгострокові зчитувані комп'ютером носії даних. У цьому контексті, канал передачі може означати канали, по яких передається контент, що зберігається

50 на цих носіях (і може включати в себе магазини роздрібної торгівлі й інший механізм доставки зі складу). У будь-якому випадку, методи цього винаходу не підлягають обмеженню в цьому відношенні прикладом, наведеним на фіг. 3.

[144] Як показано далі в прикладі, наведеному на фіг. 3, споживач 14 контенту включає в себе систему 16 відтворення аудіосигналу. Система 16 відтворення аудіосигналу може

55 представляти будь-яку систему відтворення аудіосигналу, здатну відтворювати багатоканальні аудіодані. Система 16 відтворення аудіосигналу може включати в себе декілька різних рендеризаторів 22. Кожний з рендеризаторів 22 може забезпечувати ту чи іншу форму рендеризації, де різні форми рендеризації можуть включати в себе один або більше з різних способів здійснення амплітудного панорамування на векторній основі (VBAP) і/або один або

більше з різних способів здійснення синтезу звукового поля. Використовуване тут "A і/або B" означає "A або B" або обидва "A і B".

[145] Система 16 відтворення аудіосигналу може додатково включати в себе пристрій 24 аудіодекодування. Пристрій 24 аудіодекодування може представляти пристрій, виконаний з можливістю декодування коефіцієнтів 11' HOA з бітового потоку 21, де коефіцієнти 11' HOA можуть бути аналогічні коефіцієнтам 11 HOA, але відрізнятися внаслідок втратних операцій (наприклад, квантування) і/або передачі по каналу передачі. Таким чином, пристрій 24 аудіодекодування може деквантувати інформацію направленості переднього плану, зазначену в бітовому потоці 21, здійснюючи також психоакустичне декодування відносно аудіооб'єктів переднього плану, зазначених у бітовому потоці 21, і кодованих коефіцієнтів HOA, що представляють фонові компоненти. Пристрій 24 аудіодекодування може додатково здійснювати інтерполяцію відносно декодованої інформації направленості переднього плану і потім визначати коефіцієнти HOA, що представляють компоненти переднього плану, на основі декодованих аудіооб'єктів переднього плану й інтерпольованої інформації направленості переднього плану. Пристрій 24 аудіодекодування може потім визначати коефіцієнти 11' HOA на основі визначених коефіцієнтів HOA, що представляють компоненти переднього плану, і декодованих коефіцієнтів HOA, що представляють фонові компоненти.

[146] Система 16 відтворення аудіосигналу може, після декодування бітового потоку 21, одержувати коефіцієнти 11' HOA і рендеризувати коефіцієнти 11' HOA для виведення підведень 25 гучномовця. Підведення 25 гучномовця можуть збуджувати один або більше гучномовців (які не показані в прикладі, наведеному на фіг. 3, для спрощення ілюстрації).

[147] Для вибору належного рендеризатора або, у ряді випадків, генерації належного рендеризатора, система 16 відтворення аудіосигналу може одержувати інформацію 13 гучномовців, що вказує кількість гучномовців і/або просторову геометрію гучномовців. У ряді випадків, система 16 відтворення аудіосигналу може одержувати інформацію 13 гучномовців з використанням опорного мікрофона і збуджувати гучномовці таким чином, щоб динамічно визначати інформацію 13 гучномовців. В інших випадках або разом з динамічним визначенням інформації 13 гучномовців, система 16 відтворення аудіосигналу може пропонувати користувачу взаємодіяти з системою 16 відтворення аудіосигналу і вводити інформацію 13 гучномовців.

[148] Потім система 16 відтворення аудіосигналу може вибирати один з аудіорендеризаторів 22 на основі інформації 13 гучномовців. У ряді випадків, система 16 відтворення аудіосигналу може, коли жоден з аудіорендеризаторів 22 не знаходиться в межах деякої порогової міри подібності (що визначається геометрією гучномовців) з зазначеними в інформації 13 гучномовців, система 16 відтворення аудіосигналу може генерувати один з аудіорендеризаторів 22 на основі інформації 13 гучномовців. Система 16 відтворення аудіосигналу, у ряді випадків, може генерувати один з аудіорендеризаторів 22 на основі інформації 13 гучномовців, не намагаючись спочатку вибрати один з існуючих аудіорендеризаторів 22.

[149] На фіг. 4 показана блок-схема, що демонструє, більш детально, один приклад пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 3, який може здійснювати різні аспекти методів, описаних у цьому винаході. Пристрій 20 аудіокодування включає в себе блок 26 аналізу контенту, блок 27 методології синтезу на векторній основі і блок 28 методології синтезу на основі напрямку.

[150] Блок 26 аналізу контенту представляє блок, виконаний з можливістю аналізу контенту коефіцієнтів 11 HOA для ідентифікації, чи представляють коефіцієнти 11 HOA контент, генерований з живого запису або аудіооб'єкта. Блок 26 аналізу контенту може визначати, чи згенеровані коефіцієнти 11 HOA із запису фактичного звукового поля або зі штучного аудіооб'єкта. Блок 26 аналізу контенту може здійснювати це визначення по-різному. Наприклад, блок 26 аналізу контенту може кодувати $(N+1)^2-1$ каналів і прогнозувати останній канал, що залишився (який може бути представлений як вектор). Блок 26 аналізу контенту може застосовувати скаляри до щонайменше деяких з $(N+1)^2-1$ каналів і підсумовувати результуючі значення для визначення останнього каналу, що залишився. Крім того, у цьому прикладі, блок 26 аналізу контенту може визначати точність прогнозованого каналу. У цьому прикладі, якщо точність прогнозованого каналу відносно висока (наприклад, точність перевищує конкретний поріг), коефіцієнти 11 HOA, імовірно, генеруються із синтетичного аудіооб'єкта. Навпаки, якщо точність прогнозованого каналу відносно низька (наприклад, точність нижче конкретного порога), коефіцієнти 11 HOA з більшою мірою імовірності представляють записане звукове поле. Наприклад, у цьому прикладі, якщо відношення сигнал/шум (SNR) прогнозованого каналу більше 100 децибелів (дБ), коефіцієнти 11 HOA з більшою мірою імовірності представляють звукове поле, генероване із синтетичного аудіооб'єкта. Навпаки, SNR звукового поля, записаного з використанням власного мікрофона, може складати від 5 до 20 дБ. Таким чином,

може існувати очевидна відмінність відносно SNR між звуковим полем, представленим коефіцієнтами 11 HOA, генерованими з фактичного прямого запису і з синтетичного аудіооб'єкта.

[151] Зокрема, блок 26 аналізу контенту може, при визначенні, чи генеруються коефіцієнти 11 HOA, що представляють звукове поле, із синтетичного аудіооб'єкта, одержувати кадровані коефіцієнти HOA, що можуть мати розмір 25 на 1024 для представлення четвертого порядку (тобто $N=4$). Після одержання кадрованих коефіцієнтів HOA (які також можуть бути позначені тут як матриця 11 кадрованих SHC, і подальші матриці кадрованих SHC можуть бути позначені як матриці 27B, 27C і т. д. кадрованих SHC). Блок 26 аналізу контенту може виключити перший вектор кадрованих коефіцієнтів 11 HOA для генерації приведених кадрованих коефіцієнтів HOA. У деяких прикладах, цей перший вектор, виключений з кадрованих коефіцієнтів 11 HOA, може відповідати коефіцієнтам 11 HOA, пов'язаним з базисною функцією у вигляді сферичної гармоніки нульового порядку, нульового підпорядку.

[152] Потім блок 26 аналізу контенту може прогнозувати перший ненульовий вектор приведених кадрованих коефіцієнтів HOA з векторів, що залишилися, приведених кадрованих коефіцієнтів HOA. Перший ненульовий вектор може означати перший вектор, від першого порядку (і з урахуванням кожного з підпорядків, що залежать від порядку) до четвертого порядку (і з урахуванням кожного з підпорядків, що залежать від порядку), що має значення, відмінні від нуля. У деяких прикладах, перший ненульовий вектор приведених кадрованих коефіцієнтів HOA стосується коефіцієнтів 11 HOA, пов'язаних з базисною функцією у вигляді сферичної гармоніки першого порядку, нульового підпорядку. Хоча вони описані відносно першого ненульового вектора, методи можуть прогнозувати інші вектори приведених кадрованих коефіцієнтів HOA з векторів, що залишилися, приведених кадрованих коефіцієнтів HOA. Наприклад, блок 26 аналізу контенту може прогнозувати ті з приведених кадрованих коефіцієнтів HOA, які пов'язані з базисною функцією у вигляді сферичної гармоніки першого порядку, першого підпорядку або базисною функцією у вигляді сферичної гармоніки першого порядку, мінус першого підпорядку. В інших прикладах, блок 26 аналізу контенту може прогнозувати ті з приведених кадрованих коефіцієнтів HOA, які пов'язані з базисною функцією у вигляді сферичної гармоніки другого порядку, нульового підпорядку.

[153] Для прогнозування першого ненульового вектора, блок 26 аналізу контенту може діяти згідно з наступним рівнянням:

$$\sum_i (\alpha_i v_i)$$

де i приймає значення від 1 до $(N+1)^2-2$, що дорівнює 23 для представлення четвертого

порядку, α_i позначає деяку сталу для i -го вектора, і v_i означає i -й вектор. Після прогнозування першого ненульового вектора, блок 26 аналізу контенту може одержувати помилку на основі прогнозованого першого ненульового вектора і фактичного ненульового вектора. У деяких прикладах, блок 26 аналізу контенту віднімає прогнозований перший ненульовий вектор з фактичного першого ненульового вектора для виведення помилки. Блок 26 аналізу контенту може обчислювати помилку як суму абсолютних значень різниць між кожним елементом у прогнозованому першому ненульовому векторі і фактичному першому ненульовому векторі.

[154] Одержавши помилку, блок 26 аналізу контенту може обчислювати відношення на основі енергії фактичного першого ненульового вектора і помилки. Блок 26 аналізу контенту може визначати цю енергію шляхом піднесення до квадрата кожного елемента першого ненульового вектора і підсумовування квадратів елементів. Потім блок 26 аналізу контенту може порівнювати це відношення з порогом. Коли відношення не перевищує поріг, блок 26 аналізу контенту може визначити, що кадровані коефіцієнти 11 HOA генеруються з запису, і указати в бітовому потоці, що відповідне кодоване представлення коефіцієнтів 11 HOA згенероване із запису. Коли відношення перевищує поріг, блок 26 аналізу контенту може визначити, що кадровані коефіцієнти 11 HOA генеруються із синтетичного аудіооб'єкта, і вказати в бітовому потоці, що відповідне кодоване представлення кадрованих коефіцієнтів 11 HOA згенероване із синтетичного аудіооб'єкта.

[155] Указання, чи згенеровані кадровані коефіцієнти 11 HOA з запису або синтетичного аудіооб'єкта, може містити одиничний біт для кожного кадру. Одиничний біт може вказувати, що різні операції кодування використовувалися для кожного кадру, ефективно переключаючись між різними способами кодування відповідного кадру. У ряді випадків, коли кадровані коефіцієнти 11 HOA згенеровані з запису, блок 26 аналізу контенту передає коефіцієнти 11 HOA на блок 27 синтезу на векторній основі. У ряді випадків, коли кадровані коефіцієнти 11 HOA згенеровані із

синтетичного аудіооб'єкта, блок 26 аналізу контенту передає коефіцієнти 11 НОА на блок 28 синтезу на основі напрямку. Блок 28 синтезу на основі напрямку може представляти блок, виконаний з можливістю здійснення синтезу на основі напрямку коефіцієнтів 11 НОА для генерації бітового потоку 21 на основі напрямку.

5 [156] Іншими словами, методи базуються на кодуванні коефіцієнтів НОА з використанням зовнішнього класифікатора. Класифікатор може діяти наступним чином:

почати з матриці кадрованих SHC (наприклад 4-го порядку, розмір кадру 1024, що також можуть іменуватися кадрованими коефіцієнтами НОА або коефіцієнтами НОА), де матриця одержується розміром 25×1024 ;

10 виключити 1-й вектор (SH 0-го порядку), виходить матриця розміром 24×1024 ;

спрогнозувати перший ненульовий вектор у матриці (вектор розміром 1×1024) з векторів, що залишилися, у матриці (23 векторів розміром 1×1024).

Прогнозування здійснюється наступним чином: прогнозований вектор=сума по i [альфа- $i \times$ вектор- i] (де підсумовування по i здійснюється по 23 індексах, $i=1 \dots 23$).

15 Потім перевірити помилку: фактичний вектор - прогнозований вектор=помилка.

Якщо відношення енергії вектора/помилка велике (тобто помилка мала), то звукове поле, що лежить в основі (у цьому кадрі), є розрідженим/синтетичним. Інакше, звукове поле, що лежить в основі, є записаним (з використанням, наприклад, мікрофонної решітки) звуковим полем.

20 Залежно від рішення записане/синтетичне, по-різному здійснювати кодування/декодування (яке може означати стиснення смуги). Рішення є 1-бітовим рішенням, що передається в бітовому потоці для кожного кадру.

[157] Як показано в прикладі, наведеному на фіг. 4, блок 27 синтезу на векторній основі може включати в себе блок 30 лінійного оборотного перетворення (LIT), блок 32 обчислення параметрів, блок 34 переупорядковування, блок 36 вибору переднього плану, блок 38 компенсації енергії, блок 40 психоакустичного аудіокодера, блок 42 генерації бітового потоку, блок 44 аналізу звукового поля, блок 46 приведення коефіцієнтів, блок 48 вибору фону (BG), блок 50 просторово-часової інтерполяції і блок 52 квантування.

30 [158] Блок 30 лінійного оборотного перетворення (LIT) приймає коефіцієнти 11 НОА у формі каналів НОА, причому кожен канал представляє блок або кадр коефіцієнта, пов'язаного з даним порядком, підпорядком сферичних базисних функцій (який може бути позначений як $\text{НОА}[k]$, де k може позначати поточний кадр або блок вибірок). Матриця коефіцієнтів 11 НОА може мати розміри $D: M \times (N+1)^2$.

35 [159] Таким чином, блок 30 LIT може представляти блок, виконаний з можливістю здійснення форми аналізу, іменованої розкладання на сингулярні значення. Хоча вони описані відносно SVD, методи, описані в цьому винаході, можна здійснювати відносно будь-якого аналогічного перетворення або розкладання, що забезпечує набори лінійно розкорельованих, ущільнених по енергії виходів. Також посилення на "набори" у цьому винаході, у загальному випадку, має означати ненульові набори, якщо прямо не зазначено зворотне, і не має означати класичне математичне визначення множин, що включає в себе так названу "порожню множину".

40 [160] Альтернативне перетворення може включати аналіз головних компонентів, часто іменований "PCA". PCA означає математичну процедуру, яка використовує ортогональне перетворення для перетворення набору спостережень можливо корелюючих змінних у набір лінійно розкорельованих змінних, іменованих головними компонентами. Лінійно розкорельовані змінні представляють змінні, котрі не мають лінійного статистичного співвідношення (або залежності) одна від одної. Ці головні компоненти можна описати як такі, що мають малу міру статистичної кореляції один з одним. У будь-якому випадку, кількість так званих головних компонентів менше або дорівнює кількості вихідних змінних. У деяких прикладах, перетворення визначається таким чином, що перший головний компонент має найбільшу можливу дисперсію (або, іншими словами, враховує максимально можливу мінливість даних), і кожен наступний компонент по черзі має найвищу дисперсію, можливу при обмеженні, що цей наступний компонент ортогональний попереднім компонентам (іншими словами, розкорельований з ними). PCA може здійснювати форму зниження порядку, що відносно коефіцієнтів 11 НОА може приводити до стиснення коефіцієнтів 11 НОА. Залежно від контексту, PCA може іменуватися по-різному, наприклад дискретне перетворення Карунена-Лева, перетворення Хотеллінга, правильне ортогональне розкладання (POD) і розкладання по власних значеннях (EVD), і т. д. Властивостями таких операцій, які приводять до мети стиснення аудіоданих, що лежить в основі, є 'ущільнення по енергії' і 'декореляція' багатоканальних аудіоданих.

60 [161] У будь-якому випадку, блок 30 LIT здійснює розкладання на сингулярні значення (яке, знову ж, може іменуватися "SVD") для перетворення коефіцієнтів 11 НОА у два або більше

наборів перетворених коефіцієнтів HOA. Ці "набори" перетворених коефіцієнтів HOA можуть включати в себе вектори перетворених коефіцієнтів HOA. У прикладі, наведеному на фіг. 4, блок 30 LIT може здійснювати SVD відносно коефіцієнтів 11 HOA для генерації так названої матриці V , матриці S і матриці U . SVD, у лінійній алгебрі, може представляти факторизацію дійсної або комплексної матриці X у на z (де X може представляти багатоканальні аудіодані, наприклад коефіцієнти 11 HOA) у наступній формі:

$$X=USV^*,$$

де U може представляти дійсну або комплексну унітарну матрицю у на u , де u стовпців U звичайно відомі як ліві сингулярні вектори багатоканальних аудіоданих. S може представляти прямокутну діагональну матрицю у на z з ненегативними дійсними числами на діагоналі, де діагональні значення S звичайно відомі як сингулярні значення багатоканальних аудіоданих. V^* (яка може позначати транспоновану і комплексно спряжену матрицю V) може представляти дійсну або комплексну унітарну матрицю z на z , де z стовпців V^* звичайно відомі як праві сингулярні вектори багатоканальних аудіоданих.

[162] Хоча в цьому винаході вони описані як застосовувані до багатоканальних аудіоданих, що містять коефіцієнти 11 HOA, методи можуть застосовуватися до будь-якої форми багатоканальних аудіоданих. Таким чином, пристрій 20 аудіокодування може здійснювати розкладання на сингулярні значення відносно багатоканальних аудіоданих, що представляють щонайменше частину звукового поля, для генерації матриці U , яка представляє ліві сингулярні вектори багатоканальних аудіоданих, матриці S , яка представляє сингулярні значення багатоканальних аудіоданих, і матриці V , яка представляє праві сингулярні вектори багатоканальних аудіоданих, і представлення багатоканальних аудіоданих як функції щонайменше частини однієї або більше з матриці U , матриці S і матриці V .

[163] У деяких прикладах, матриця V^* у згаданому вище математичному виразі SVD позначається як спряжена транспонована матриця V для відображення того, що SVD може застосовуватися до матриць, які містять комплексні числа. У випадку застосування до матриць, які містять тільки дійсні числа, комплексно-спряжену матрицю V (або, іншими словами, матрицю V^*) можна розглядати як транспоновану матрицю V . Нижче передбачається, для спрощення ілюстрації, що коефіцієнти 11 HOA містять дійсні числа, у результаті чого матриця V виводиться через SVD замість матриці V^* . Крім того, хоча в цьому винаході вона позначена як матриця V , посилення на матрицю V варто розуміти як транспоновану матрицю V , коли це доцільно. Хоча передбачається, що це матриця V , методи можуть застосовуватися аналогічним чином до коефіцієнтів 11 HOA, що мають комплексні коефіцієнти, де виходом SVD є матриця V^* . Відповідно, методи не підлягають обмеженню в цьому відношенні тільки забезпеченням застосування SVD для генерації матриці V , але можуть включати в себе застосування SVD до коефіцієнтів 11 HOA, що мають комплексні компоненти, для генерації матриці V^* .

[164] У будь-якому випадку, блок 30 LIT може здійснювати поблокову форму SVD відносно кожного блока (який може означати кадр) аудіоданих амбіофонії більш високого порядку (HOA) (де ці аудіодані амбіофонії включають в себе блоки або вибірки коефіцієнтів 11 HOA або будь-яку іншу форму багатоканальних аудіоданих). Як згадано вище, змінну M можна використовувати для позначення довжини кадру аудіо у вибірках. Наприклад, коли кадр аудіо включає в себе 1024 вибірки аудіосигналу, M дорівнює 1024. Хоча вони описані відносно цього типового значення для M , методи цього винаходу не підлягають обмеженню цим типовим значенням для M . Таким чином, блок 30 LIT може здійснювати поблокове SVD відносно блока коефіцієнтів 11 HOA, що має M на $(N+1)^2$ коефіцієнтів HOA, де N , знову ж, позначає порядок аудіоданих HOA. Блок 30 LIT може генерувати, за допомогою здійснення цього SVD, матрицю V , матрицю S і матрицю U , де кожна з матриць може представляти відповідні матриці V , S і U , описані вище. Таким чином, блок 30 лінійного оборотного перетворення може здійснювати SVD відносно коефіцієнтів 11 HOA для виведення векторів $33 US[k]$ (які можуть представляти комбіновану версію векторів S і векторів U), що мають розміри $D: M \times (N+1)^2$, і векторів $35 V[k]$, що мають розміри $D: (N+1)^2 \times (N+1)^2$. Окремі елементи вектора в матриці $US[k]$ також можуть позначатися $X_{PS}(k)$, тоді як характерні вектори матриці $V[k]$ також можуть позначатися $v(k)$.

[165] Аналіз матриць U , S і V може показати, що ці матриці несуть або представляють просторові і часові характеристики звукового поля, що лежить в основі, представленого вище як X . Кожний з N векторів у U (довжиною M вибірок) може представляти нормалізовані розділені аудіосигнали як функцію часу (протягом періоду часу, представленого M вибірками), які ортогональні один одному і які відділені від будь-яких просторових характеристик (які також можуть іменуватися інформацією направленості). Просторові характеристики, що представляють просторову форму і позицію (r , тета, ϕ) ширини, можуть бути альтернативно

представлені окремими i -ими векторами, $v^{(i)}(k)$, у матриці V (кожен довжиною $(N+1)^2$). Вектори в матриці U і в матриці V нормалізовані таким чином, що їх середньоквадратичні енергії дорівнюють одиниці. Енергія аудіосигналів у U , таким чином, представляється діагональними елементами в S . Множення U і S для формування $US[k]$ (з окремими елементами вектора $X_{PS}(k)$), таким чином, представляє аудіосигнал істинними енергіями. Здатність розкладання SVD відділяти часові аудіосигнали (у U), їх енергії (у S) і їх просторові характеристики (у V) може підтримувати різні аспекти методів, описаних у цьому винаході. Крім того, ця модель синтезу коефіцієнтів HOA[k], що лежать в основі, X , за допомогою векторного множення $US[k]$ і $V[k]$ породжує термін "методологія векторного синтезу", який використовується протягом цього документа.

[166] Хоча він описаний як здійснюваний безпосередньо відносно коефіцієнтів 11 HOA, блок 30 LIT може застосовувати лінійне оборотне перетворення до похідних коефіцієнтів 11 HOA. Наприклад, блок 30 LIT може застосовувати SVD відносно матриці спектральної щільності потужності, виведеної з коефіцієнтів 11 HOA. Матриця спектральної щільності потужності може бути позначена як PSD і одержана шляхом матричного множення транспонованої $hoaFrame$ на $hoaFrame$, що викладено в нижченаведеному псевдокоді. Позначення $hoaFrame$ означає кадр коефіцієнтів 11 HOA.

[167] Блок 30 LIT може, після застосування SVD (svd) до PSD, одержувати матрицю $S[k]^2$ ($S_squared$) і матрицю $V[k]$. Матриця $S[k]^2$ може позначати квадрат матриці $S[k]$, у зв'язку з чим блок 30 LIT може застосовувати операцію витягання квадратного кореня з матриці $S[k]^2$ для одержання матриці $S[k]$. Блок 30 LIT може, у ряді випадків, здійснювати квантування відносно матриці $V[k]$ для одержання квантованої матриці $V[k]$ (яка може бути позначена як матриця $V[k]'$). Блок 30 LIT може одержувати матрицю $U[k]$, спочатку множачи матрицю $S[k]$ на квантовану матрицю $V[k]'$ для одержання матриці $SV[k]'$. Потім блок 30 LIT може одержувати псевдообернену ($pinv$) матрицю $SV[k]'$ і потім множити коефіцієнти 11 HOA на псевдообернену матрицю $SV[k]'$ для одержання матриці $U[k]$. Вищенаведений опис може бути представлений наступним псевдокодом:

```

PSD=hoaFrame*hoaFrame;
[V, S_squared]=svd(PSD, 'econ');
S=sqrt(S_squared);
U=hoaFrame * pinv(S*V');

```

[168] Шляхом здійснення SVD відносно спектральної щільності потужності (PSD) коефіцієнтів HOA замість самих коефіцієнтів, блок 30 LIT може потенційно знижувати обчислювальну складність здійснення SVD відносно одного або більше з циклів процесора і простору зберігання, досягаючи при цьому такої ж ефективності аудіокодування джерела, як якби SVD застосовувалося безпосередньо до коефіцієнтів HOA. Таким чином, вищеописане SVD типу PSD може потенційно вимагати менше обчислювальної потужності, оскільки SVD здійснюється на матриці F^*F (де F - кількість коефіцієнтів HOA). У порівнянні з матрицею M^*F , де M - довжина кадру, тобто 1024 або більше вибірок. Тепер складність SVD, завдяки застосуванню до PSD замість коефіцієнтів 11 HOA, може бути близько $O(L^3)$ у порівнянні з $O(M^*L^2)$ у випадку застосування до коефіцієнтів 11 HOA (де $O(*)$ позначає символ великого O обчислювальної складності, звичайно використовуваний у комп'ютерній техніці).

[169] Блок 32 обчислення параметрів представляє блок, виконаний з можливістю обчислення різних параметрів, наприклад параметр кореляції (R), параметри властивостей направленості (θ , ϕ , r) і властивість енергії (e). Кожний з цих параметрів для поточного кадру може бути позначений як $R[k]$, $\theta[k]$, $\phi[k]$, $r[k]$ і $e[k]$. Блок 32 обчислення параметрів може здійснювати аналіз енергії і/або кореляцію (або так називану крос-кореляцію) відносно векторів 33 $US[k]$ для ідентифікації цих параметрів. Блок 32 обчислення параметрів також може визначати ці параметри для попереднього кадру, де параметри попереднього кадру можуть бути позначені $R[k-1]$, $\theta[k-1]$, $\phi[k-1]$, $r[k-1]$ і $e[k-1]$, на основі попереднього кадру вектора $US[k-1]$ і векторів $V[k-1]$. Блок 32 обчислення параметрів може виводити поточні параметри 37 і попередні параметри 39 на блок 34 переупорядкування.

[170] Таким чином, блок 32 обчислення параметрів може здійснювати аналіз енергії відносно кожного з L перших векторів 33 $US[k]$, що відповідають першому часу, і кожного з других векторів 33 $US[k-1]$, що відповідають другому часу, обчислювати середньоквадратичну енергію для щонайменше частини (але часто усього) першого кадру аудіо і частини (але часто усього) другого кадру аудіо і, таким чином, генерувати 2L енергії, по одному для кожного з L перших векторів 33 $US[k]$ першого кадру аудіо і по одному для кожного з других векторів 33 $US[k-1]$ другого кадру аудіо.

[171] В інших прикладах, блок 32 обчислення параметрів може здійснювати крос-кореляцію між деякою частиною (якщо не усього) набору вибірок для кожного з перших векторів 33 $US[k]$ і кожного з других векторів 33 $US[k-1]$. Крос-кореляція може означати крос-кореляцію в тому розумінні, у якому її розуміють фахівці з обробки сигналів. Іншими словами, крос-кореляція може означати міру подібності між двома формами хвилі (які в цьому випадку визначаються як дискретний набір M вибірок) як функцію запізнювання за часом, застосовуваного до однієї з них. У деяких прикладах, для здійснення крос-кореляції, блок 32 обчислення параметрів порівнює останні L вибірок кожного з перших векторів 27 $US[k]$, по черзі, з першими L вибірками кожного з других векторів 33 $US[k-1]$, що залишилися, для визначення параметра кореляції. Використовувана тут "почергова" операція означає поелементну операцію, виконувану відносно першого набору елементів і другого набору елементів, де операція витягає один елемент із кожного з першого і другого наборів елементів "по черзі" відповідно до упорядкування наборів.

[172] Блок 32 обчислення параметрів також може аналізувати вектори 35 $V[k]$ і/або $V[k-1]$ для визначення параметрів властивостей направленості. Ці параметри властивостей направленості можуть забезпечувати указання переміщення і положення аудіооб'єкта, представлених відповідними векторами 33 $US[k]$ і/або $US[k-1]$. Блок 32 обчислення параметрів може забезпечувати будь-яку комбінацію вищенаведених поточних параметрів 37 (визначених відносно векторів 33 $US[k]$ і/або векторів 35 $V[k]$) і будь-яку комбінацію попередніх параметрів 39 (визначених відносно векторів 33 $US[k-1]$ і/або векторів 35 $V[k-1]$) на блок 34 переупорядкування.

[173] Розкладання SVD не гарантує, що аудіосигнал/об'єкт представлений r -им вектором у векторах 33 $US[k-1]$, який може бути позначений як вектор $US[k-1][r]$ (або, альтернативно, як $X_{PS}^{(p)}(k-1)$), буде тим же аудіосигналом/об'єктом (що розвився у часі), представленим r -им вектором у векторах 33 $US[k]$, які також можуть бути позначені як вектори 33 $US[k][r]$ (або, альтернативно, як $X_{PS}^{(p)}(k)$). Параметри, обчислені блоком 32 обчислення параметрів, можуть використовуватися блоком 34 переупорядкування для переупорядкування аудіооб'єктів для представлення їх природного оцінювання або неперервності протягом часу.

[174] Таким чином, блок 34 переупорядкування потім може порівнювати кожний з параметрів 37 з перших векторів 33 $US[k]$ по черзі з кожним з параметрів 39 для других векторів 33 $US[k-1]$. Блок 34 переупорядкування може переупорядковувати (з використанням, у порядку одного прикладу, угорського алгоритму) різні вектори в матриці 33 $US[k]$ і матриці 35 $V[k]$ на основі поточних параметрів 37 і попередніх параметрів 39 для виведення переупорядкованої матриці 33' $US[k]$ (яка може бути математично позначена як $\overline{US[k]}$) і

переупорядкованої матриці 35' $V[k]$ (яка може бути математично позначена як $\overline{V[k]}$) на блок 36 вибору звуку переднього плану (або переважного звуку - PS) ("блок 36 вибору переднього плану") і блок 38 компенсації енергії.

[175] Іншими словами, блок 34 переупорядкування може представляти блок, виконаний з можливістю переупорядкування векторів у матриці 33 $US[k]$ для генерації переупорядкованої матриці 33' $US[k]$. Блок 34 переупорядкування може переупорядковувати матрицю 33 $US[k]$, оскільки порядок векторів 33 $US[k]$ (де, знову ж, кожен вектор з векторів 33 $US[k]$, який, знову ж, можна альтернативно позначати як $X_{PS}^{(p)}(k)$), може представляти один або більше характерних (або, іншими словами, переважних) монофонічних аудіооб'єктів, присутніх у звуковому полі) може відрізнятися від ділянок аудіоданих. Таким чином, за умови, що пристрій 12 аудіокодування, у деяких прикладах, діє на цих ділянках аудіоданих, у загальному випадку, іменованих кадрами аудіо, позиція векторів, що відповідають цим окремим монофонічним аудіооб'єктам, представлених у матриці 33 $US[k]$ як виведені, може відрізнятися від кадру аудіо до кадру аудіо внаслідок застосування SVD до кадрів і зміни помітності кожного аудіооб'єкта від кадру до кадру.

[176] Передача векторів у матриці 33 $US[k]$ безпосередньо на блок 40 психоакустичного аудіокодера без переупорядкування векторів у матриці 33 $US[k]$ від кадру аудіо до кадру аудіо може знижувати міру стиснення, досягнувши для деяких схем стиснення, наприклад успадкованих схем стиснення, що діють краще, коли монофонічні аудіооб'єкти є неперервними (поканально, що визначається в цьому прикладі позиційним порядком векторів у матриці 33 $US[k]$ один відносно одного) по кадрах аудіо. Крім того, за відсутності переупорядкування, кодування векторів у матриці 33 $US[k]$ може знижувати якість аудіоданих при декодуванні. Наприклад, кодера AAC, що можуть бути представлені в прикладі, наведеному на фіг. 3, блоком 40

психоакустичного аудіокодера, можуть більш ефективно стискати переупорядковані один або більше векторів у матриці 33' US[k] від кадру до кадру в порівнянні зі стисненням, що досягається при безпосередньому кодуванні векторів у матриці 33 US[k] від кадру до кадру. Хоча вище приведений опис відносно кодерів AAC, методи можна здійснювати відносно будь-якого кодера, який забезпечує краще стиснення, коли монофонічні аудіооб'єкти задаються по кадрах у конкретному порядку або позиції (поканально).

[177] Різні аспекти методів, таким чином, можуть передбачати, що пристрій 12 аудіокодування преупорядковує один або більше векторів (наприклад, вектори в матриці 33 US[k] для генерації переупорядкованих одного або більше векторів у переупорядкованій матриці 33' US[k] і, таким чином, полегшує стиснення векторів у матриці 33 US[k] успадкованим аудіокодером, наприклад блоком 40 психоакустичного аудіокодера.

[178] Наприклад, блок 34 переупорядковування може переупорядковувати один або більше векторів у матриці 33 US[k] з першого кадру аудіо, що іде за часом за другим кадром, якому відповідають один або більше других векторів у матриці 33 US[k-1] на основі поточних параметрів 37 і попередніх параметрів 39. Хоча описано, що перший кадр аудіо іде за часом за другим кадром аудіо, перший кадр аудіо може передувати за часом другому кадру аудіо. Відповідно, методи не підлягають обмеженню прикладом, описаним в цьому винаході.

[179] Для ілюстрації розглянемо нижченаведену таблицю 1, де кожний з p векторів у матриці 33 US[k] позначається як US[k][p], де k позначає, чи одержаний відповідний вектор з k -го кадру або попереднього ($k-1$)-го кадру, і p позначає рядок вектора відносно векторів того ж кадру аудіо (де матриця US[k] має $(N+1)^2$ таких векторів). Як згадано вище, передбачаючи N заданим рівним одиниці, p може позначати вектори з першого (1) по четвертий (4).

Таблиця 1

Енергія, що розглядається	У порівнянні з
US[k-1][1]	US[k][1], US[k][2], US[k][3], US[k][4]
US[k-1][2]	US[k][1], US[k][2], US[k][3], US[k][4]
US[k-1][3]	US[k][1], US[k][2], US[k][3], US[k][4]
US[k-1][4]	US[k][1], US[k][2], US[k][3], US[k][4]

[180] У вищенаведеній таблиці 1, блок 34 переупорядковування порівнює енергію, обчислену для US[k-1][1], з енергією, обчисленою для кожного з US[k][1], US[k][2], US[k][3], US[k][4], енергію, обчислену для US[k-1][2], з енергією, обчисленою для кожного з US[k][1], US[k][2], US[k][3], US[k][4], і т. д. Блок 34 переупорядковування потім може відкидати один або більше із других векторів 33 US[k-1] другого попереднього кадру аудіо (за часом). Для ілюстрації, розглянемо нижченаведену таблицю 2, що демонструє другі вектори, що залишилися, 33 US[k-1].

Таблиця 2

Вектор, що розглядається	Інші, що розглядаються
US[k-1][1]	US[k][1], US[k][2]
US[k-1][2]	US[k][1], US[k][2]
US[k-1][3]	US[k][3], US[k][4]
US[k-1][4]	US[k][3], US[k][4]

[181] У вищенаведеній таблиці 2, блок 34 переупорядковування може визначати, на основі порівняння енергії, що енергія, обчислена для US[k-1][1], аналогічна енергії, обчисленій для кожного з US[k][1] і US[k][2], енергія, обчислена для US[k-1][2], аналогічна енергії, обчисленій для кожного з US[k][1] і US[k][2], енергія, обчислена для US[k-1][3], аналогічна енергії, обчисленій для кожного з US[k][3] і US[k][4], і енергія, обчислена для US[k-1][4], аналогічна енергії, обчисленій для кожного з US[k][3] і US[k][4]. У деяких прикладах, блок 34 переупорядковування може здійснювати додатково аналіз енергії для ідентифікації подібності між кожним з перших векторів матриці 33 US[k] і кожним із других векторів матриці 33 US[k-1].

[182] В інших прикладах, блок 32 переупорядковування може переупорядковувати вектори на основі поточних параметрів 37 і попередніх параметрів 39, що стосуються крос-кореляції. У цих прикладах, повертаючись до вищенаведеної таблиці 2, блок 34 переупорядковування може визначати наступну ілюстративну кореляцію, виражену в таблиці 3, на основі цих параметрів крос-кореляції.

Таблиця 3

Вектор, що розглядається	Корелює з
$US[k-1][1]$	$US[k][2]$
$US[k-1][2]$	$US[k][1]$
$US[k-1][3]$	$US[k][3]$
$US[k-1][4]$	$US[k][4]$

[183] З вищенаведеної таблиці 3, блок 34 переупорядкування визначає, у порядку одного прикладу, що вектор $US[k-1][1]$ корелює з інакше розташованим вектором $US[k][2]$, вектор $US[k-1][2]$ корелює з інакше розташованим вектором $US[k][1]$, вектор $US[k-1][3]$ корелює з аналогічно розташованим вектором $US[k][3]$ і вектор $US[k-1][4]$ корелює з аналогічно розташованим вектором $US[k][4]$. Іншими словами, блок 34 переупорядкування визначає те, що може іменуватися інформацією переупорядкування, яка описує, як преупорядковувати перші вектори матриці 33 $US[k]$ таким чином, що вектор $US[k][2]$ змінює положення в першому рядку перших векторів матриці 33 $US[k]$ і вектор $US[k][1]$ змінює положення в другому рядку перших векторів 33 $US[k]$. Потім блок 34 переупорядкування може переупорядковувати перші вектори матриці 33 $US[k]$ на основі цієї інформації переупорядкування для генерації переупорядкованої матриці 33' $US[k]$.

[184] Додатково, блок 34 переупорядкування може, хоча це не показано в прикладі, наведеному на фіг. 4, забезпечувати цю інформацію переупорядкування на пристрій 42 генерації бітового потоку, який може генерувати бітовий потік 21 для включення цієї інформації переупорядкування, завдяки чому пристрій аудіодекодування, наприклад пристрій 24 аудіодекодування, показаний в прикладі, наведеному на фіг. 3 і 5, може визначати, як преупорядковувати переупорядковані вектори матриці 33' $US[k]$ для відновлення векторів матриці 33 $US[k]$.

[185] Хоча він описаний вище як здійснюючий двоетапний процес, що передбачає аналіз на основі спочатку залежних від енергії параметрів і потім параметрів крос-кореляції, блок 32 переупорядкування може здійснювати цей аналіз тільки відносно параметрів енергії для визначення інформації переупорядкування, здійснювати цей аналіз тільки відносно параметрів крос-кореляції для визначення інформації переупорядкування або здійснювати аналіз відносно як параметрів енергії, так і параметрів крос-кореляції вищеописаним чином. Додатково, методи можуть використовувати інші типи процесів для визначення кореляції, які не передбачають здійснення одного або обох з порівняння енергії і/або крос-кореляції. Відповідно, методи не підлягають обмеженню в цьому відношенні вищевикладеними прикладами. Крім того, інші параметри, одержані від блока 32 обчислення параметрів (наприклад, параметри просторового положення, виведені з векторів V або кореляції векторів у $V[k]$ і $V[k-1]$), також можна використовувати (або одночасно/спільно, або послідовно) з параметрами енергії і крос-кореляції, одержаними з $US[k]$ і $US[k-1]$, для визначення правильного упорядкування векторів у US .

[186] У порядку одного прикладу використання кореляції векторів у матриці V , блок 34 обчислення параметрів може визначити, що вектори матриці 35 $V[k]$ корелюють, як зазначено в нижченаведеній таблиці 4..

Таблиця 4

Вектор, що розглядається	Корелює з
$V[k-1][1]$	$V[k][2]$
$V[k-1][2]$	$V[k][1]$
$V[k-1][3]$	$V[k][3]$
$V[k-1][4]$	$V[k][4]$

З вищенаведеної таблиці 4, блок 34 переупорядкування визначає, у порядку одного прикладу, що вектор $V[k-1][1]$ корелює з інакше розташованим вектором $V[k][2]$, вектор $V[k-1][2]$ корелює з інакше розташованим вектором $V[k][1]$, вектор $V[k-1][3]$ корелює з аналогічно розташованим вектором $V[k][3]$ і вектор $V[k-1][4]$ корелює з аналогічно розташованим вектором $V[k][4]$. Блок 34 переупорядкування може виводити переупорядковану версію векторів матриці 35 $V[k]$ як переупорядковану матрицю 35' $V[k]$.

[187] У деяких прикладах, те ж переупорядкування, яке застосовується до векторів у матриці US , також застосовується до векторів у матриці V . Іншими словами, будь-який аналіз, використовуваний при переупорядкуванні векторів V , можна використовувати спільно з будь-яким аналізом, використовуваним для переупорядкування векторів US . Для ілюстрації прикладу, у якому інформація переупорядкування визначається не тільки відносно параметрів енергії і/або параметрів крос-кореляції відносно векторів $35\ US[k]$, блок 34 переупорядкування також може здійснювати цей аналіз відносно векторів $35\ V[k]$ на основі параметрів крос-кореляції і параметрів енергії аналогічно описаному вище відносно векторів $35\ V[k]$. Крім того, хоча вектори $33\ US[k]$ не мають ніяких властивостей направленості, вектори $35\ V[k]$ можуть забезпечувати інформацію, що стосується направленості відповідних векторів $33\ US[k]$. У цьому розумінні, блок 34 переупорядкування може ідентифікувати кореляції між векторами $35\ V[k]$ і векторами $35\ V[k-1]$ на основі аналізу відповідних параметрів властивостей направленості. Таким чином, у деяких прикладах, аудіооб'єкт рухається в звуковому полі неперервно, коли переміщується або знаходиться у стані спокою у відносно стійкому положенні. Таким чином, блок 34 переупорядкування може ідентифікувати ті вектори матриці $35\ V[k]$ і матриці $35\ V[k-1]$, які демонструють деякий відомий фізично реалістичний рух або стоять нерухомо в звуковому полі, як корелюючі, переупорядковуючи вектори $33\ US[k]$ і вектори $35\ V[k]$ на основі кореляції цих властивостей направленості. У будь-якому випадку, блок 34 переупорядкування може виводити переупорядковані вектори $33'\ US[k]$ і переупорядковані вектори $35'\ V[k]$ на блок 36 вибору переднього плану.

[188] Додатково, методи можуть використовувати інші типи процесів для визначення правильного порядку, які не передбачають здійснення одного або обох з порівняння енергії і/або крос-кореляції. Відповідно, методи не підлягають обмеженню в цьому відношенні вищевикладеними прикладами.

[189] Хоча воно описане вище як переупорядкування векторів матриці V для відображення переупорядкування векторів матриці US , у ряді випадків, вектори V можна переупорядковувати інакше, ніж вектори US , де окремі елементи синтаксису можуть генеруватися для указання переупорядкування векторів US і переупорядкування векторів V . У ряді випадків, вектори V можуть не переупорядковуватися, і тільки вектори US можуть переупорядковуватися за умови, що вектори V можуть не піддаватися психоакустичному кодуванню.

[190] Варіант здійснення, де переупорядкування векторів матриці V і векторів матриці US відрізняються, відповідають наміру обмінюватися аудіооб'єктами в просторі, тобто переміщення їх з вихідної записаної позиції (коли звукове поле, що лежить в основі, є природним записом) або майстерно призначеної позиції (коли звукове поле, що лежить в основі, є штучною сумішшю об'єктів). У порядку прикладу, припустимо, що існує два аудіоджерела A і B , A може бути звуком котячого нявкання, що виходить з "лівої" частини звукового поля, і B може бути звуком собачого гавкоту, що виходить з "правої" частини звукового поля. Коли переупорядкування V і US відрізняються, позиції двох джерел звуку міняються місцями. Після обміну A (нявкання) виходить із правої частини звукового поля і B (гавкіт) виходить з лівої частини звукового поля.

[191] Блок 44 аналізу звукового поля може представляти блок, виконаний з можливістю здійснення аналізу звукового поля відносно коефіцієнтів $11\ HOA$, щоб потенційно досягти цільової бітової швидкості 41 . Блок 44 аналізу звукового поля може, на основі цього аналізу і/або прийнятої цільової бітової швидкості 41 , визначати сумарну кількість екземплярів психоакустичного кодера (яка може бути функцією сумарної кількості каналів оточення або фону (BG_{TOT}) і кількості каналів переднього плану або, іншими словами, переважних каналів). Сумарну кількість екземплярів психоакустичного кодера можна позначити як $n_{umHOATransportChannels}$. Блок 44 аналізу звукового поля також може визначати, знову ж для потенційного досягнення цільової бітової швидкості 41 , сумарну кількість каналів переднього плану (n_{FG}) 45 , мінімальний порядок звукового поля фону (або, іншими словами, оточення) (N_{BG} або, альтернативно, $MinAmbHoaOrder$), відповідну кількість фактичних каналів, що представляють мінімальний порядок фонового звукового поля ($n_{BGa} = (MinAmbHoaOrder + 1)^2$), і індекси (i) додаткових каналів $HOA\ BG$ для відправлення (які можуть спільно позначатися як інформація 43 фонового каналу в прикладі, наведеному на фіг. 4). Інформація 42 фонового каналу також може іменуватися інформацією 43 каналу оточення. Кожний з каналів, що залишилися з $n_{umHOATransportChannels} - n_{BGa}$, може бути або "додатковим каналом фону/оточення", "активним переважним каналом на векторній основі", "активним переважним сигналом на основі направленості", або "повністю неактивним". В одному варіанті здійснення, ці типи каналів можуть указуватися (як "ChannelType") елемент синтаксису двома бітами (наприклад 00 : додатковий фоновий канал; 01 : переважний сигнал на векторній основі; 10 :

неактивний сигнал; 11: сигнал на основі направленості). Сумарну кількість сигналів фону або оточення, $nBGa$, можна задавати у вигляді $(MinAmbHoaOrder+1)^2$ + число разів, коли індекс 00 (у вищенаведеному прикладі) виглядає як тип каналу в бітовому потоці для цього кадру.

[192] У будь-якому випадку, блок 44 аналізу звукового поля може вибирати кількість каналів фону (або, іншими словами, оточення) і кількість каналів переднього плану (або, іншими словами, переважних) на основі цільової бітової швидкості 41, вибирати більше каналів фону і/або переднього плану, при відносно більш високій цільовій бітовій швидкості 41 (наприклад, коли цільова бітова швидкість 41 більше або дорівнює 512 кбіт/с). В одному варіанті здійснення, $numHOATransportChannels$ можна задати рівним 8, тоді як $MinAmbHoaOrder$ можна задати рівним 1 у секції заголовка бітового потоку (яка описана більш детально з посиланням на фіг. 10-10O(ii)). У цьому сценарії, у кожному кадрі, чотири канали можуть бути призначені для представлення фонові або оточуючої частини звукового поля, тоді як інші 4 канали можуть, на основі кадрів, залежати від типу каналу, наприклад, використовуваного або як додатковий канал фону/оточення, або як передньоплановий/переважний канал. Сигнали переднього плану/переважні сигнали можуть бути або сигналами на векторній основі, або сигналами на основі направленості, як описано вище.

[193] У ряді випадків, сумарну кількість переважних сигналів на векторній основі для кадру можна задавати у вигляді числа разів, коли індекс $ChannelType$ дорівнює 01, у бітовому потоці цього кадру, у вищенаведеному прикладі. У вищенаведеному варіанті здійснення, для кожного додаткового каналу фону/оточення (наприклад, відповідного $ChannelType$, що дорівнює 00), відповідна інформація кожного з можливих коефіцієнтів НОА (крім перших чотирьох) може бути представлена в цьому каналі. Ця інформація, для контенту НОА четвертого порядку, може бути індексом для вказання між 5-25 (перші чотири 1-4 можуть відправлятися увесь час, коли $minAmbHoaOrder$ заданий рівним 1, тому необхідно тільки вказати між 5-25). Ця інформація, таким чином, може відправлятися з використанням 5-бітового елемента синтаксису (для контенту 4-го порядку), який може бути позначений як "CodedAmbCoeffIdx".

[194] У другому варіанті здійснення, усі сигнали переднього плану/переважні сигнали є сигналами на векторній основі. У цьому другому варіанті здійснення, сумарну кількість сигналів переднього плану/переважних сигналів можна задавати у вигляді $nFG = numHOATransportChannels - [(MinAmbHoaOrder+1)^2 + \text{число разів, коли індекс дорівнює 00}]$.

[195] Блок 44 аналізу звукового поля виводить інформацію 43 фоновому каналу і коефіцієнти 11 НОА на блок 46 вибору фону (BG), інформацію 43 фоновому каналу на блок 46 приведення коефіцієнтів і блок 42 генерації бітового потоку, і 45 nFG на блок 36 вибору переднього плану.

[196] У деяких прикладах, блок 44 аналізу звукового поля може вибирати, на основі аналізу векторів матриці 33 $US[k]$ і цільової бітової швидкості 41, змінну nFG , кількість цих компонентів, що мають найбільше значення. Іншими словами, блок 44 аналізу звукового поля може визначати значення змінної A (яке може бути аналогічне або по суті аналогічне N_{BG}), яка розділяє два підпростори, шляхом аналізу нахилу кривої, створеної зменшуваними діагональними значеннями векторів матриці $S[k]$ 33, де великі сингулярні значення представляють звуки переднього плану або характерні звуки і низькі сингулярні значення представляють фонові компоненти звукового поля. Таким чином, змінна A може сегментувати загальне звукове поле на підпростір переднього плану і підпростір фону.

[197] У деяких прикладах, блок 44 аналізу звукового поля може використовувати першу і другу похідні кривої сингулярних значень. Блок 44 аналізу звукового поля може також обмежувати значення для змінної A величиною між одиницею і п'ятьма. У порядку іншого прикладу, блок 44 аналізу звукового поля може обмежувати значення змінної A величиною між одиницею і $(N+1)^2$. Альтернативно, блок 44 аналізу звукового поля може попередньо задавати значення для змінної A , наприклад значення чотири. У будь-якому випадку, на основі значення A , блок 44 аналізу звукового поля визначає сумарну кількість каналів 45 переднього плану (nFG), порядок фоновому звукового поля (N_{BG}) і кількість ($nBGa$) і індекси (i) додаткових каналів НОА BG для відправлення.

[198] Крім того, блок 44 аналізу звукового поля може визначати енергію векторів у матриці 35 $V[k]$ для кожного вектора. Блок 44 аналізу звукового поля може визначати енергію для кожного з векторів у матриці 35 $V[k]$ і ідентифікувати ті, що мають високу енергію, як компоненти переднього плану.

[199] Крім того, блок 44 аналізу звукового поля може здійснювати різні інші аналізи відносно коефіцієнтів 11 НОА, що включають у себе просторовий аналіз енергії, аналіз просторового маскування, аналіз дифузії або інші форми звукових аналізів. Блок 44 аналізу звукового поля може здійснювати просторовий аналіз енергії шляхом перетворення коефіцієнтів 11 НОА у просторову область і ідентифікації областей високої енергії, що представляють направлені

компоненти звукового поля, які потрібно зберегти. Блок 44 аналізу звукового поля може здійснювати перцептивний аналіз просторового маскування аналогічно просторовому аналізу енергії, за винятком того, що блок 44 аналізу звукового поля може ідентифікувати просторові області, які маскуються просторово близькими звуками більш високої енергії. Потім блок 44 аналізу звукового поля може, на основі перцепційно маскованих областей, ідентифікувати менше компонентів переднього плану в ряді випадків. Блок 44 аналізу звукового поля може додатково здійснювати аналіз дифузії відносно коефіцієнтів 11 HOA для ідентифікації областей дифузної енергії, які можуть представляти фонові компоненти звукового поля.

[200] Блок 44 аналізу звукового поля може також представляти блок, виконаний з можливістю визначення помітності, характерності або переваги аудіоданих, що представляють звукове поле, з використанням інформації на основі направленості, пов'язаної з аудіоданими. Хоча визначення на основі енергії можуть поліпшувати рендеризацію звукового поля, розкладеного за допомогою SVD для ідентифікації характерних аудіокомпонентів звукового поля, визначення на основі енергії також можуть приводити до того, що пристрій помилково ідентифікує фонові аудіокомпоненти як характерні аудіокомпоненти, у випадках, коли фонові аудіокомпоненти демонструють високий рівень енергії. Таким чином, розділення характерних і фонових аудіокомпонентів тільки на основі енергії може бути ненадійним, оскільки енергетичні (наприклад, більш голосні) фонові аудіокомпоненти можуть неправильно ідентифікуватися як характерні аудіокомпоненти. Щоб більш надійно розрізняти між характерними і фоновими аудіокомпонентами звукового поля, різні аспекти методів, описаних у цьому винаході, можуть передбачати, що блок 44 аналізу звукового поля здійснює аналіз на основі направленості коефіцієнтів 11 HOA для відділення аудіокомпонентів переднього плану й оточення від розкладених версій коефіцієнтів 11 HOA.

[201] У цьому відношенні, блок 44 аналізу звукового поля може представляти блок, сконфігурований або інакше виконаний з можливістю ідентифікації характерних (або передньопланових) елементів з фонових елементів, включених в один або більше з векторів у матриці 33 US[k] і векторів у матриці 35 V[k]. Відповідно до деяких методів на основі SVD, найбільш енергетичні компоненти (наприклад, перші декілька векторів однієї або більше з матриці 33 US[k] і матриці 35 V[k] або виведених з них векторів) можна розглядати як характерні компоненти. Однак, найбільш енергетичні компоненти (які представлені векторами) з одного або більше з векторів у матриці 33 US[k] і векторів у матриці 35 V[k] можуть не у всіх сценаріях представляти найбільш направлені компоненти/сигнали.

[202] Блок 44 аналізу звукового поля може реалізувати один або більше аспектів методів, описаних тут, для ідентифікації передній план/прямий/переважний елементи на основі направленості векторів з одного або більше з векторів у матриці 33 US[k] і векторів у матриці 35 V[k] або виведених з них векторів. У деяких прикладах, блок 44 аналізу звукового поля може ідентифікувати або вибрати як характерні аудіокомпоненти (де компоненти також можуть іменуватися "об'єктами") один або більше векторів на основі енергії і направленості векторів. Наприклад, блок 44 аналізу звукового поля може ідентифікувати ті вектори з одного або більше з векторів у матриці 33 US[k] і векторів у матриці 35 V[k] (або виведених з них векторів), які демонструють високу енергію і високу направленість (наприклад, представлену як показник направленості) як характерні аудіокомпоненти. У результаті, якщо блок 44 аналізу звукового поля визначає, що конкретний вектор є відносно менш направленим у порівнянні з іншими векторами з одного або більше з векторів у матриці 33 US[k] і векторів у матриці 35 V[k] (або виведених з них векторів), то, незалежно від рівня енергії, пов'язаного з конкретним вектором, блок 44 аналізу звукового поля може визначити, що конкретний вектор представляє аудіокомпоненти фону (або оточення) звукового поля, представленого коефіцієнтами 11 HOA.

[203] У деяких прикладах, блок 44 аналізу звукового поля може ідентифікувати характерні аудіооб'єкти (які, як згадано вище, також можуть іменуватися "компонентами") на основі направленості, шляхом здійснення наступних операцій. Блок 44 аналізу звукового поля може множити (наприклад, з використанням одного або більше процесів матричного множення) вектори в матриці S[k] (які можна виводити з векторів 33 US[k] або, хоча це не показано в прикладі, наведеному на фіг. 4, окремо виводити блоком 30 LIT) на вектори в матриці 35 V[k]. Шляхом множення матриці 35 V[k] і векторів S[k], блок 44 аналізу звукового поля може одержувати матрицю VS[k]. Додатково, блок 44 аналізу звукового поля може підносити до квадрата (тобто, підносити до степеня два) щонайменше деякі з елементів кожного з векторів у матриці VS[k]. У ряді випадків, блок 44 аналізу звукового поля може підсумовувати квадрати елементів кожного вектора, що пов'язані з порядком, який більше 1.

[204] У порядку одного прикладу, якщо кожен вектор матриці VS[k] включає в себе 25 елементів, блок 44 аналізу звукового поля може, відносно кожного вектора, підносити до

квадрата елементи кожного вектора, починаючи з п'ятого елемента і закінчуючи двадцять п'ятим елементом, підсумовуючи квадрати елементів для визначення показника направленості (або індикатора направленості). Кожна операція підсумовування може давати показник направленості для відповідного вектора. У цьому прикладі, блок 44 аналізу звукового поля може визначити, що ті елементи кожного рядка, які пов'язані з порядком, що менше або дорівнює 1, а саме елементи з першого по четвертий, більшою мірою належать до величини енергії й меншою мірою до направленості цих елементів. Таким чином, амбіофонія більш низького порядку, пов'язана з порядком нуль або один, відповідають сферичним базисним функціям, які, як показано на фіг. 1 і фіг. 2, не забезпечують багато відносно напрямку хвилі тиску, але, проте, забезпечують деяку голосність (яка представляє енергію).

[205] Операції, описані у вищенаведеному прикладі, також можуть бути виражені відповідно до наступного псевдокоду. Нижченаведений псевдокод включає в себе анотації, у формі операторів коментарів, які поміщені між послідовними екземплярами рядків символів "/"* і "*/" (без лапок).

```
[U, S, V]=svd(audioframe, 'ecom');
VS=V*S;
/* Наступний рядок стосується незалежного аналізу кожного рядка і підсумовування значень
у першому (у порядку одного прикладу) рядку від п'ятого елемента до двадцять п'ятого
елемента для визначення показника направленості або метрики направленості для відповідного
вектора. Піднести до квадрата елементи до підсумовування. Елементи в кожному рядку, що
пов'язані з порядком, який більше 1, пов'язані з амбіофонією більш високого порядку, і, таким
чином, з більшою імовірністю є направленими. */
sumVS=sum(VS(5:end, :).^2,1);
/* Наступний рядок стосується сортування суми квадратів для генерованої матриці VS і
вибору набору найбільших значень (наприклад, трьох або чотирьох найбільших значень) */
[~,idxVS]=sort(sumVS, 'descend');
U=U(:,idxVS);
V=V(:,idxVS);
S=S(idxVS, idxVS);
```

[206] Іншими словами, відповідно до вищенаведеного псевдокоду, блок 44 аналізу звукового поля може вибирати елементи кожного вектора матриці VS[k], розкладеної з коефіцієнтів 11 НОА, відповідних сферичній базисній функції, що має порядок, який більше одиниці. Потім блок 44 аналізу звукового поля може підносити до квадрата ці елементи для кожного вектора матриці VS[k], підсумовувати квадрати елементів для ідентифікації, обчислювати або інакше визначати метрику або показник направленості для кожного вектора матриці VS[k]. Потім, блок 44 аналізу звукового поля може сортувати вектори матриці VS[k] на основі відповідних метрик направленості кожного з векторів. Блок 44 аналізу звукового поля може сортувати ці вектори в порядку убывання метрик направленості таким чином, що вектори з найвищою відповідною направленістю є першими і вектори з найнижчою відповідною направленістю є останніми. Потім блок 44 аналізу звукового поля може вибрати ненульовий піднабір векторів, що мають найвищу відносну метрику направленості.

[207] Блок 44 аналізу звукового поля може здійснювати будь-яку комбінацію вищенаведених аналізів для визначення сумарної кількості екземплярів психоакустичного кодера (яка може бути функцією сумарної кількості каналів оточення або фону (BG_{TOT}) і кількості каналів переднього плану. Блок 44 аналізу звукового поля може, на основі будь-якої комбінації вищенаведених аналізів, визначати сумарну кількість каналів переднього плану (nFG), порядок фонового звукового поля (N_{BG}) і кількість ($nBGa$) і індекси (i) додаткових каналів НОА BG для відправлення (які можуть спільно позначатися як інформація 43 фонового каналу в прикладі, наведеному на фіг. 4).

[208] У деяких прикладах, блок 44 аналізу звукового поля може здійснювати цей аналіз кожні М вибірок, що може відновлюватися на основі кадрів. У цьому відношенні, значення А може змінюватися від кадру до кадру. Приклад бітового потоку, де рішення виробляється кожні М вибірок, показаний на фіг. 10-10O(ii). В інших прикладах, блок 44 аналізу звукового поля може здійснювати цей аналіз більше одного разу за кадр, аналізуючи дві або більше ділянок кадру. Відповідно, методи не підлягають обмеженню в цьому відношенні прикладами, наведеними в цьому винаході.

[209] Блок 48 вибору фону може представляти блок, виконаний з можливістю визначення коефіцієнтів 47 НОА фону або оточення на основі інформації фонового каналу (наприклад, фонового звукового поля (N_{BG}) і кількості ($nBGa$) і індексів (i) додаткових каналів НОА BG для відправлення). Наприклад, коли N_{BG} дорівнює одиниці, блок 48 вибору фону може вибирати

коефіцієнти 11 НОА для кожної вибірки кадру аудіо, що має порядок, який менше або дорівнює одиниці. Потім блок 48 вибору фону може, у цьому прикладі, вибрати коефіцієнти 11 НОА, що мають індекс, ідентифікований одним з індексів (i) як додаткові коефіцієнти НОА BG, де nBGa надходить на блок 42 генерації бітового потоку для указання в бітовому потоці 21, щоб пристрій аудіодекодування, наприклад пристрій 24 аудіодекодування, показаний в прикладі, наведеному на фіг. 3, виділяв коефіцієнти 47 НОА BG з бітового потоку 21. Потім блок 48 вибору фону може виводити коефіцієнти 47 НОА оточення на блок 38 компенсації енергії. Коефіцієнти 47 НОА оточення можуть мати розміри D: $M \times [(N_{BG}+1)^2 + nBGa]$.

[210] Блок 36 вибору переднього плану може представляти блок, виконаний з можливістю вибору тих з переупорядкованої матриці 33' US[k] і переупорядкованої матриці 35' V[k], які представляють компоненти переднього плану або характерні компоненти звукового поля на основі 45 nFG (який може представляти один або більше індексів, що ідентифікують ці вектори переднього плану). Блок 36 вибору переднього плану може виводити сигнали 49 nFG (які можуть бути позначені як переупорядковані $US[k]_{1,..., 49}$ nFG, $FG_{1,..., nFG}[k]$ 49, або

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases} \quad 49)$$

на блок 40 психоакустичного аудіокодера, де сигнали 49 nFG можуть мати розміри D: $M \times nFG$, і кожний представляє монофонічні аудіооб'єкти. Блок 36 вибору переднього плану також може виводити переупорядковану матрицю 35' V[k] (або

$$residual = |v_q| - 2^{cid-1} \quad 35'),$$

яка відповідає компонентам переднього плану звукового поля, на блок 50 просторово-часової інтерполяції, де ті з переупорядкованої матриці 35' V[k], що відповідають компонентам переднього плану, можуть бути позначені як матриця V[k] 51_k переднього плану (яка може математично позначатися як $\bar{V}_{1,...,nFG}[k]$), що має розміри D: $(N+1)^2 \times nFG$.

[211] Блок 38 компенсації енергії може представляти блок, виконаний з можливістю здійснення компенсації енергії відносно коефіцієнтів 47 НОА оточення для компенсації втрати енергії внаслідок видалення різних каналів НОА блоком 48 вибору фону. Блок 38 компенсації енергії може здійснювати аналіз енергії відносно однієї або більше з переупорядкованої матриці 33' US[k], переупорядкованої матриці 35' V[k], сигналів 49 nFG, векторів 51_k переднього плану V[k] і коефіцієнтів 47 НОА оточення і потім здійснювати компенсацію енергії на основі цього аналізу енергії для генерації коефіцієнтів 47' НОА оточення з компенсацією енергії. Блок 38 компенсації енергії може виводити коефіцієнти 47' НОА оточення з компенсацією енергії на блок 40 психоакустичного аудіокодера.

[212] По суті, блок 38 компенсації енергії можна використовувати для компенсації можливого зниження повної енергії фонових звукових компонентів звукового поля, обумовленого зниженням порядку компонентів звукового поля оточення, описаних коефіцієнтами 11 НОА, для генерації коефіцієнтів 47 НОА оточення зниженого порядку (які, у деяких прикладах, мають порядок, менший ніж N, відносно тільки включених коефіцієнтів, що відповідають сферичним базисним функціям, які мають наступні порядки/підпорядки: $[(N_{BG}+1)^2 + nBGa]$). У деяких прикладах, блок 38 компенсації енергії компенсує ці втрати енергії шляхом визначення коефіцієнта підсилення компенсації у формі значень посилення для застосування до кожного з $[(N_{BG}+1)^2 + nBGa]$ стовпців коефіцієнтів 47 НОА оточення для збільшення середньоквадратичної (RMS) енергії коефіцієнтів 47 НОА оточення, щоб вони дорівнювали або щонайменше були більш наближені до RMS коефіцієнтів 11 НОА (які визначаються за допомогою аналізу сукупної енергії однієї або більше з переупорядкованої матриці 33' US[k], переупорядкованої матриці 35' V[k], сигналів 49 nFG, векторів 51_k переднього плану V[k] і коефіцієнтів 47 НОА оточення зниженого порядку), до виведення коефіцієнтів 47 НОА оточення на блок 40 психоакустичного аудіокодера.

[213] У ряді випадків, блок 38 компенсації енергії може ідентифікувати RMS для кожного рядка і/або стовпця однієї або більше з переупорядкованої матриці 33' US[k] і переупорядкованої матриці 35' V[k]. Блок 38 компенсації енергії також може ідентифікувати RMS для кожного рядка і/або стовпця одного або більше з вибраних каналів переднього плану, який може включати в себе сигнали 49 nFG і вектори 51_k переднього плану V[k], і коефіцієнти 47 НОА оточення зниженого порядку. RMS для кожного рядка і/або стовпця однієї або більше з переупорядкованої матриці 33' US[k] і переупорядкованої матриці 35' V[k] може зберігатися у вектор, позначений RMS_{FULL} , тоді як RMS для кожного рядка і/або стовпця одного або більше із

сигналів 49 nFG, векторів 51_k переднього плану $V[k]$ і коефіцієнтів 47 HOA оточення зниженого порядку можуть зберігатися у вектор, позначений $RMS_{REDUCED}$. Блок 38 компенсації енергії потім може обчислювати вектор Z значень посилення, згідно з наступним рівнянням: $Z = RMS_{FULL}/RMS_{REDUCED}$. Потім блок 38 компенсації енергії може застосовувати цей

- 5 вектор Z значень посилення або різні його ділянки до одного або більше із сигналів 49 nFG, векторів 51_k переднього плану $V[k]$ і коефіцієнтів 47 HOA оточення зниженого порядку. У ряді випадків, вектор Z значень посилення застосовується тільки до коефіцієнтів 47 HOA оточення зниженого порядку відповідно до наступного рівняння $HOA_{BG-RED}' = HOA_{BG-RED} Z^T$, де HOA_{BG-RED} позначає коефіцієнти 47 HOA оточення зниженого порядку, HOA_{BG-RED}' позначає
10 приведені коефіцієнти 47' HOA оточення з компенсацією енергії і Z^T позначає транспонування вектора Z .

- [214] У деяких прикладах, для визначення кожного RMS відповідних рядків і/або стовпців однієї або більше з переупорядкованої матриці 33' $US[k]$, переупорядкованої матриці 35' $V[k]$, сигналів 49 nFG, векторів 51_k переднього плану $V[k]$ і коефіцієнтів 47 HOA оточення зниженого
15 порядку, блок 38 компенсації енергії може спочатку застосовувати рендеризатор опорних коефіцієнтів сферичних гармонік (SHC) до стовпців. Застосування рендеризатора опорних SHC блоком 38 компенсації енергії дозволяє визначати RMS в області SHC для визначення енергії загального звукового поля, описаного кожним рядком і/або стовпцем кадру, представленого рядками і/або стовпцями однієї або більше з переупорядкованої матриці 33' $US[k]$,
20 переупорядкованої матриці 35' $V[k]$, сигналів 49 nFG, векторів 51_k переднього плану $V[k]$ і коефіцієнтів 47 HOA оточення зниженого порядку, як більш докладно описано нижче.

- [215] Блок 50 просторово-часової інтерполяції може представляти блок, виконаний з можливістю прийому векторів 51_k переднього плану $V[k]$ для k -го кадру і векторів 51_{k-1} переднього плану $V[k-1]$ для попереднього кадру (тому позначених $k-1$) і здійснення просторово-
25 часової інтерполяції для генерації інтерпольованих векторів переднього плану $V[k]$. Блок 50 просторово-часової інтерполяції може рекомбінувати сигнали 49 nFG з векторами 51_k переднього плану $V[k]$ для відновлення переупорядкованих коефіцієнтів HOA переднього плану. Потім блок 50 просторово-часової інтерполяції може ділити переупорядковані коефіцієнти HOA переднього плану на інтерпольовані вектори $V[k]$ для генерації інтерпольованих сигналів 49' nFG. Блок 50 просторово-часової інтерполяції також може виводити ті з векторів 51_k переднього
30 плану $V[k]$, які використовувалися для генерації інтерпольованих векторів переднього плану $V[k]$, завдяки чому пристрій аудіодекодування, наприклад пристрій 24 аудіодекодування, може генерувати інтерпольовані вектори переднього плану $V[k]$ і, таким чином, відновлювати вектори 51_k переднього плану $V[k]$. Вектори 51_k переднього плану $V[k]$, використовувані для генерації інтерпольованих векторів переднього плану $V[k]$, позначені як вектори 53, що залишилися, переднього плану $V[k]$. Щоб гарантувати, що одні і ті ж $V[k]$ і $V[k-1]$ використовуються на кодері і декодері (для створення інтерпольованих векторів $V[k]$), на кодері і декодері можна використовувати їх квантованні/деквантовані версії.

- [216] У цьому відношенні, блок 50 просторово-часової інтерполяції може представляти блок,
40 що інтерполує першу частину першого кадру аудіо з яких-небудь інших ділянок першого кадру аудіо і другого в часі наступного або попереднього кадру аудіо. У деяких прикладах, ділянки можуть бути позначені як підкадри, де інтерполяція, здійснювана відносно підкадрів, більш докладно описана нижче з посиланням на фіг. 45-46Е. В інших прикладах, блок 50 просторово-часової інтерполяції може діяти відносно деякої кількості останніх вибірок попереднього кадру і
45 деякої кількості перших вибірок наступного кадру, як описано більш детально з посиланням на фіг. 37-39. Блок 50 просторово-часової інтерполяції може, при здійсненні цієї інтерполяції, зменшувати кількість вибірок векторів 51_k переднього плану $V[k]$, які необхідно вказувати в бітовому потоці 21, оскільки тільки ті з векторів 51_k переднього плану $V[k]$, які використовуються для генерації інтерпольованих векторів $V[k]$, представляють піднабір векторів 51_k переднього
50 плану $V[k]$. Таким чином, щоб потенційно зробити стиснення коефіцієнтів 11 HOA більш ефективним (шляхом зменшення кількості векторів 51_k переднього плану $V[k]$, що задаються в бітовому потоці 21), різні аспекти методів, описаних у цьому винаході, можуть передбачати інтерполяцію однієї або більше ділянок першого кадру аудіо, де кожна з ділянок може представляти розкладені версії коефіцієнтів 11 HOA.

- [217] Просторово-часова інтерполяція може забезпечувати ряд переваг. По-перше, сигнали 49 nFG можуть не бути неперервними від кадру до кадру внаслідок поблокового характеру здійснення SVD або іншого LIT. Іншими словами, за умови, що блок 30 LIT застосовує SVD на основі кадрів, певні порушення неперервності можуть існувати в результируючих перетворених коефіцієнтах HOA як свідчення, наприклад, неупорядкованого характеру матриці 33 $US[k]$ і

матриці 35 $V[k]$. Завдяки здійсненню цієї інтерполяції, порушення неперервності може знижуватися за умови, що інтерполяція може мати згладжуючу дію, яка потенційно знижує будь-які артефакти, що вносяться внаслідок границь кадру (або, іншими словами, сегментування коефіцієнтів 11 НОА на кадри). Завдяки використанню векторів 51_k переднього плану $V[k]$ для здійснення цієї інтерполяції і потім генерації інтерпольованих сигналів 49' nFG на основі інтерпольованих векторів 51_k переднього плану $V[k]$ з відновлених переупорядкованих коефіцієнтів НОА, можна згладжувати щонайменше деякі ефекти, обумовлені покадровою операцією, а також переупорядкуванням сигналів 49 nFG .

[218] У ході експлуатації, блок 50 просторово-часової інтерполяції може інтерпольовувати один або більше підкадрів першого кадру аудіо з першого розкладання, наприклад векторів 51_k переднього плану $V[k]$, частини першої множини коефіцієнтів 11 НОА, включених у перший кадр, і другого розкладання, наприклад векторів переднього плану $V[k]$ 51_{k-1} , частини другої множини коефіцієнтів 11 НОА, включених у другий кадр, для генерації розкладених інтерпольованих коефіцієнтів сферичних гармонік для одного або більше підкадрів.

[219] У деяких прикладах, перше розкладання містить перші вектори 51_k переднього плану $V[k]$, що представляють праві сингулярні вектори частини коефіцієнтів 11 НОА. Аналогічно, у деяких прикладах, друге розкладання містить другі вектори 51_k переднього плану $V[k]$, що представляють праві сингулярні вектори частини коефіцієнтів 11 НОА.

[220] Іншими словами, 3D аудіо на основі сферичних гармонік може бути параметричним представленням 3D поля тиску відносно ортогональних базисних функцій на сфері. Чим вище порядок N представлення, тим потенційно вище просторове розрізнення і, часто, тим більше кількість коефіцієнтів сферичних гармонік (SH) (для усього $(N+1)^2$ коефіцієнтів). Для багатьох застосувань, може знадобитися, щоб стиснення смуги коефіцієнтів дозволяло ефективно передавати і зберігати коефіцієнти. Ці методи, передбачені в цьому винаході, можуть забезпечувати процес зниження розмірності на кадровій основі з використанням розкладання на сингулярні значення (SVD). Аналіз SVD може розкладати кожен кадр коефіцієнтів на три матриці U , S і V . У деяких прикладах, методи можуть обробляти деякі з векторів у матриці $US[k]$ як компоненти переднього плану звукового поля, що лежить в основі. Однак, при такій обробці, ці вектори (у матриці U $S[k]$) є розривними від кадру до кадру, хоча вони представляють один і той же характерний аудіокомпонент. Ці порушення неперервності можуть приводити до значних артефактів, коли компоненти надходять через перетворюючі аудіокодекери.

[221] Методи, описані в цьому винаході, дозволяють справитися з цим порушенням неперервності. Таким чином, методи можуть ґрунтуватися на тому спостереженні, що матрицю V можна інтерпретувати як ортогональні просторові осі в області сферичних гармонік. Матриця $U[k]$ може представляти проєкцію даних сферичних гармонік (НОА) відносно цих базисних функцій, де порушення неперервності може приписуватися ортогональній просторовій осі ($V[k]$), які змінюють кожен кадр і, таким чином, самі є розривними. У цьому полягає відмінність від аналогічного розкладання, наприклад перетворення Фур'є, де базисні функції, у деяких прикладах, постійні від кадру до кадру. У цьому відношенні, SVD можна розглядати як алгоритм переслідування збігу. Методи, описані в цьому винаході, можуть передбачати, що блок 50 просторово-часової інтерполяції підтримує неперервність між базисними функціями ($V[k]$) від кадру до кадру шляхом інтерполяції між ними.

[222] Як згадано вище, інтерполяцію можна здійснювати відносно вибірок. Цей випадок узагальнений у вищенаведеному описі, коли підкадри містять одиничний набір вибірок. В обох випадках інтерполяції по вибірках і по підкадрах операція інтерполяції може приймати форму наступного рівняння:

$$A = V_{BG}^{RMS} / V_{BG}'^{RMS},$$

У цьому рівнянні, інтерполяцію можна здійснювати відносно одиничного V -вектора

$$A = [\alpha_1 \dots \alpha_{(N+1)^2 - D}],$$

з одиничного V -вектора $v(k-1)$, який в одному варіанті здійснення може представляти V -вектори із сусідніх кадрів k і $k-1$. У вищенаведеному рівнянні, l представляє розрізнення, з яким здійснюється інтерполяція, де l може вказувати цілочислову вибірку і $l=1, \dots, T$ (де T - довжина вибірок, по яких здійснюється інтерполяція і по яких вимагаються вихідні інтерпольовані вектори, $v(l)$, і також вказує, що вихід цього процесу створює l з цих векторів). Альтернативно, l може вказувати підкадри, що складаються з множинних вибірок. Коли, наприклад, кадр

поділяється на чотири підкадри, l може містити значення 1, 2, 3 і 4, для кожного з підкадрів. Значення l може сигналізуватися як поле за назвою "CodedSpatialInterpolationTime" за допомогою бітового потоку, завдяки чому операція інтерполяції може повторюватися на декодері.

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

5

може містити значення вагових коефіцієнтів інтерполяції. Коли інтерполяція є лінійною,

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

може змінюватися лінійно і монотонно між 0 і 1, як функція l . В інших випадках,

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

10

може змінюватися між 0 і 1 нелінійно, але монотонно (наприклад, чвертьперіод піднятого косинуса) як функція l . Функція,

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

може індексуватися між декількома різними можливостями функцій і сигналізуватися в бітовому потоці як поле за назвою "SpatialInterpolationMethod" таким чином, що ідентична операція інтерполяції може повторюватися декодером. Коли

15

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

має значення, близьке до 0, вихід, $\overline{v(l)}$, може мати високу вагу або визначатися $residual = |v_q| - 2^{cid-1}$. Тоді як, коли

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

20

має значення, близьке до 1, це гарантує, що вихід, $\overline{v(l)}$, має високу вагу або визначається $residual = |v_q| - 2^{cid-1}$.

[223] Блок 46 приведення коефіцієнтів може представляти блок, виконаний з можливістю здійснення приведення коефіцієнтів відносно векторів 53, що залишилися, переднього плану $V[k]$ на основі інформації 43 фоновому каналу для виведення приведених векторів 55 переднього плану $V[k]$ на блок 52 квантування. Приведені вектори 55 переднього плану $V[k]$ можуть мати розміри $D: [(N+1)^2 - (N_{BG}+1)^2 - n_{BG}] \times n_{FG}$.

25

[224] Блок 46 приведення коефіцієнтів може, у цьому відношенні, представляти блок, виконаний з можливістю зниження кількості коефіцієнтів векторів 53, що залишилися, переднього плану $V[k]$. Іншими словами, блок 46 приведення коефіцієнтів може представляти блок, виконаний з можливістю виключення коефіцієнтів векторів переднього плану $V[k]$ (які утворюють вектори 53, що залишилися, переднього плану $V[k]$), які мають мало або не мають інформації направленості. Як описано вище, у деяких прикладах, коефіцієнти характерних або,

30

іншими словами, векторів переднього плану $V[k]$, що відповідають базисним функціям першого і нульового порядку (який може бути позначений як N_{BG}), забезпечують мало інформації направленості і, таким чином, можуть бути видалені з векторів переднього плану V (за допомогою процесу, що може іменуватися "приведення коефіцієнтів"). У цьому прикладі, підвищена гнучкість може забезпечуватися не тільки для ідентифікації цих коефіцієнтів, що відповідають N_{BG} , але і для ідентифікації додаткових каналів НОА (які можуть позначатися змінною $TotalOfAddAmbHOAChan$) з набору $[(N_{BG}+1)^2+1, (N+1)^2]$. Блок 44 аналізу звукового поля може аналізувати коефіцієнти 11 НОА для визначення BG_{TOT} , який може ідентифікувати не тільки $(N_{BG}+1)^2$, але і $TotalOfAddAmbHOAChan$, що можуть спільно іменуватися інформацією 43 фонового каналу. Блок 46 приведення коефіцієнтів потім може видаляти коефіцієнти, що відповідають $(N_{BG}+1)^2$, і $TotalOfAddAmbHOAChan$ з векторів 53, що залишилися, переднього плану $V[k]$ для генерації матриці 55 $V[k]$ меншого розміру розміром $(N+1)^2-(BG_{TOT}) \times nFG$, що також можуть іменуватися приведеними векторами 55 переднього плану $V[k]$.

[225] Блок 52 квантування може представляти блок, виконаний з можливістю здійснення будь-якої форми квантування для стиснення приведених векторів 55 переднього плану $V[k]$ для генерації кодованих векторів 57 переднього плану $V[k]$, виведення цих кодованих векторів 57 переднього плану $V[k]$ на блок 42 генерації бітового потоку. У ході експлуатації, блок 52 квантування може представляти блок, виконаний з можливістю стиснення просторового компонента звукового поля, тобто одного або більше із приведених векторів 55 переднього плану $V[k]$ у цьому прикладі. Для прикладу, передбачається, що приведені вектори 55 переднього плану $V[k]$ включають в себе два вектори-рядки, що мають, у результаті приведення коефіцієнтів, менше 25 елементів кожний (що передбачає представлення НОА звукового поля четвертого порядку). Хоча опис наведений відносно двох векторів-рядків, будь-яка кількість векторів може бути включена в приведені вектори 55 переднього плану $V[k]$ аж до $(n+1)^2$, де n позначає порядок представлення НОА звукового поля. Крім того, хоча він описаний нижче як здійснюючий скалярне і/або ентропійне квантування, блок 52 квантування може здійснювати будь-яку форму квантування, яка приводить до стиснення приведених векторів 55 переднього плану $V[k]$.

[226] Блок 52 квантування може приймати приведені вектори 55 переднього плану $V[k]$ і здійснювати схему стиснення для генерації кодованих векторів 57 переднього плану $V[k]$. Ця схема стиснення може передбачати, у загальному випадку, будь-яку припустиму схему стиснення для стиснення елементів вектора або даних, і не підлягає обмеженню прикладом, більш детально описаним нижче. Блок 52 квантування може здійснювати, у порядку прикладу, схему стиснення, яка включає в себе одне або більше з перетворення представлень з плаваючою комою кожного елемента приведених векторів 55 переднього плану $V[k]$ у цілочислові представлення кожного елемента приведених векторів 55 переднього плану $V[k]$, однорідного квантування цілочислових представлень приведених векторів 55 переднього плану $V[k]$ і категоризації і кодування квантованих цілочислових представлень векторів 55, що залишилися, переднього плану $V[k]$.

[227] У деяких прикладах, різні один або більше процесів цієї схеми стиснення можуть динамічно керуватися параметрами для досягнення або приблизного досягнення, у порядку одного прикладу, цільової бітової швидкості для результуючого бітового потоку 21. За умови, що кожний із приведених векторів 55 переднього плану $V[k]$ ортонормальні один одному, кожний із приведених векторів 55 переднього плану $V[k]$ можна кодувати незалежно. У деяких прикладах, як більш докладно описано нижче, кожен елемент кожного з приведених векторів 55 переднього плану $V[k]$ можна кодувати з використанням одного і того ж режиму кодування (заданого різними підрежимами).

[228] У будь-якому випадку, як згадано вище, ця схема кодування може спочатку використовувати перетворення представлень з плаваючою комою кожного елемента (яке, у деяких прикладах, є 32-бітовим числом з плаваючою комою) кожного з приведених векторів 55 переднього плану $V[k]$ у 16-бітове цілочислове представлення. Блок 52 квантування може здійснювати це перетворення плаваючої коми у ціле число шляхом множення кожного елемента даного одного з приведених векторів 55 переднього плану $V[k]$ на 2^{15} , яке, у деяких прикладах, здійснюється за допомогою зсуву вправо на 15.

[229] Блок 52 квантування потім може здійснювати однорідне квантування відносно всіх елементів даного одного з приведених векторів 55 переднього плану $V[k]$. Блок 52 квантування може ідентифікувати розмір кроку квантування на основі значення, яке може бути позначене як параметр $nbits$. Блок 52 квантування може динамічно визначати цей параметр $nbits$ на основі цільової бітової швидкості 41. Блок 52 квантування може визначати розмір кроку квантування як функцію цього параметра $nbits$. У порядку одного прикладу, блок 52 квантування може

визначати розмір кроку квантування (позначений як "дельта" або " Δ " у цьому винаході) рівним $2^{16-nbits}$. У цьому прикладі, якщо $nbits$ дорівнює шести, дельта дорівнює 2^{10} , і існує 2^6 рівнів квантування. У цьому відношенні, для елемента v вектора, квантований елемент v_q вектора дорівнює $[v/\Delta]$ і $-2^{nbits-1} < v_q < 2^{nbits-1}$.

- 5 [230] Блок 52 квантування потім може здійснювати категоризацію і залишкове кодування квантованих елементів вектора. У порядку одного прикладу, блок 52 квантування може, для даного квантованого елемента v_q вектора, ідентифікувати категорію (визначаючи ідентифікатор категорії cid), якій відповідає цей елемент, з використанням наступного рівняння:

$$cid = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}$$

- 10 Блок 52 квантування потім може кодувати по Хаффману цей індекс категорії cid , ідентифікуючи також знаковий біт, який вказує, чи має v_q позитивне значення або негативне значення. Блок 52 квантування потім може ідентифікувати залишок у цій категорії. У порядку одного прикладу, блок 52 квантування може визначати цей залишок згідно з наступним рівнянням:

15 $residual = |v_q| - 2^{cid-1}$

Блок 52 квантування потім може блоково кодувати цей залишок $cid-1$ бітами.

[231] Наступний приклад демонструє спрощений приклад цього процесу категоризації і залишкового кодування. Спочатку припустимо, що $nbits$ дорівнює шести, так що $v_q \in [-31, 31]$.

Потім припустимо наступне:

20

cid	v_q	Код Хаффмана для cid
0	0	'1'
1	-1, 1	'01'
2	-3,-2, 2,3	'000'
3	-7,-6,-5,-4, 4,5,6,7	'0010'
4	-15,-14,...,-8, 8,...,14,15	'00110'
5	-31,-30,...,-16, 16,...,30,31	'00111'

Також припустимо наступне:

cid	Блоковий код для залишку
0	N/A
1	0, 1
2	01,00, 10,11
3	011,010,001,000, 100,101,110,111
4	0111, 0110...,0000, 1000,...,1110,1111
5	01111,...,00000, 10000,...,11111

Таким чином, для $v_q=[6, -17, 0, 0, 3]$, можна визначити наступне:

- 25 » $cid=3,5,0,0,2$
 » $знак=1,0,x, x,1$
 » $залишок=2,1,x, x,1$
 » біти для 6='0010'+1'+10'
 » біти для -17='00111'+0'+0001'
 » біти для 0='0'
 30 » біти для 0='0'
 » біти для 3='000'+1'+1'
 » усього бітів=7+10+1+1+5=24
 » середня кількість бітів=24/5=4.8

- 35 [232] Хоча це не показано у вищевизначеному спрощеному прикладі, блок 52 квантування може вибирати різні кодові книги Хаффмана для різних значень $nbits$ при кодуванні cid . У деяких прикладах, блок 52 квантування може забезпечувати різні таблиці хаффманівського кодування для значень $nbits$ 6,..., 15. Крім того, блок 52 квантування може включати в себе п'ять різних кодових книг Хаффмана для кожного з різних значень $nbits$ у межах 6,..., 15 для всього 50 кодових книг Хаффмана. У цьому відношенні, блок 52 квантування може включати в себе 40 множину різних кодових книг Хаффмана для розміщення кодування cid у декількох різних статистичних контекстах.

[233] Для ілюстрації, блок 52 квантування може, для кожного зі значень $nbits$, включати в себе першу кодову книгу Хаффмана для кодування елементів вектора з першого по четвертий, другу кодову книгу Хаффмана для кодування елементів вектора з п'ятого по дев'ятий, третю кодову книгу Хаффмана для кодування елементів вектора з дев'ятого і вище. Ці перші три кодові книги Хаффмана можна використовувати, коли один із приведених векторів 55 переднього плану $V[k]$, що підлягає стисненню, не прогнозується з наступного у часі відповідного одного з приведених векторів 55 переднього плану $V[k]$ і не представляє просторову інформацію синтетичного аудіооб'єкта (заданого, наприклад, початково аудіооб'єктом, підданим імпульсно-кодовій модуляції (ІКМ)). Блок 52 квантування може додатково включати в себе, для кожного зі значень $nbits$, четверту кодову книгу Хаффмана для кодування одного з приведених векторів 55 переднього плану $V[k]$, коли цей один із приведених векторів 55 переднього плану $V[k]$ прогнозується з наступного у часі відповідного одного з приведених векторів 55 переднього плану $V[k]$. Блок 52 квантування також може включати в себе, для кожного зі значень $nbits$, п'яту кодову книгу Хаффмана для кодування одного з приведених векторів 55 переднього плану $V[k]$, коли цей один із приведених векторів 55 переднього плану $V[k]$ представляє синтетичний аудіооб'єкт. Різні кодові книги Хаффмана можна розробляти для кожного з цих різних статистичних контекстів, тобто непрогнозованого і несинтетичного контексту, прогнозованого контексту і синтетичного контексту в цьому прикладі.

[234] Нижченаведена таблиця демонструє вибір таблиці Хаффмана і біти, що підлягають указанню в бітовому потоці, щоб блок зняття стиснення міг вибирати належну таблицю Хаффмана.

Режим прогнозування	Інформація НТ	Таблиця НТ
0	0	НТ5
0	1	НТ{1,2,3}
1	0	НТ4
1	1	НТ5

У вищенаведеній таблиці, режим прогнозування ("режим прогнозування") вказує, чи здійснювалося прогнозування для поточного вектора, тоді як таблиця Хаффмана ("інформація НТ") вказує додаткову інформацію кодової книги (або таблиці) Хаффмана, використовувану для вибору однієї з таблиць Хаффмана з першої по п'яту.

[235] Нижченаведена таблиця додатково демонструє цей процес вибору таблиці Хаффмана в різних статистичних контекстах або сценаріях.

	Запис	Синтетичний
без прогнозування	НТ{1,2,3}	НТ5
з прогнозуванням	НТ4	НТ5

У вищенаведеній таблиці, стовпець "запис" вказує контекст кодування, коли вектор представляє аудіооб'єкт, який був записаний, тоді як стовпець "синтетичний" вказує контекст кодування, коли вектор представляє синтетичний аудіооб'єкт. Рядок "без прогнозування" вказує контекст кодування, коли прогнозування не здійснюється відносно елементів вектора, тоді як рядок "з прогнозуванням" вказує контекст кодування, коли прогнозування здійснюється відносно елементів вектора. Як показано в цій таблиці, блок 52 квантування вибирає НТ{1,2,3}, коли вектор представляє записаний аудіооб'єкт, і прогнозування не здійснюється відносно елементів вектора. Блок 52 квантування вибирає НТ5, коли аудіооб'єкт представляє синтетичний аудіооб'єкт, і прогнозування не здійснюється відносно елементів вектора. Блок 52 квантування вибирає НТ4, коли вектор представляє записаний аудіооб'єкт, і прогнозування здійснюється відносно елементів вектора. Блок 52 квантування вибирає НТ5, коли аудіооб'єкт представляє синтетичний аудіооб'єкт, і прогнозування здійснюється відносно елементів вектора.

[236] У цьому відношенні, блок 52 квантування може здійснювати вищезгадане скалярне квантування і/або хаффманівське кодування для стиснення приведених векторів 55 переднього плану $V[k]$, виводячи кодовані вектори 57 переднього плану $V[k]$, які можуть іменуватися інформацією 57 побічного каналу. Ця інформація 57 побічного каналу може включати в себе елементи синтаксису, використовувані для кодування векторів 55, що залишилися, переднього плану $V[k]$. Блок 52 квантування може виводити інформацію 57 побічного каналу аналогічно прикладу, показаному на одній з фіг. 10В і 10С.

[237] Як згадано вище, блок 52 квантування може генерувати елементи синтаксису для інформації 57 побічного каналу. Наприклад, блок 52 квантування може вказувати елемент

синтаксису в заголовку одиниці доступу (який може включати в себе один або більше кадрів), що позначає, який з множини режимів конфігурації вибраний. Хоча він описаний як зазначений для кожної одиниці доступу, блок 52 квантування може вказувати цей елемент синтаксису на покaдровій основі або будь-якій іншій періодичній основі або неперіодичній основі (наприклад, один раз для всього бітового потоку). У будь-якому випадку, цей елемент синтаксису може містити два біти, що вказують, який з чотирьох режимів конфігурації був вибраний для указання ненульового набору коефіцієнтів приведених векторів 55 переднього плану $V[k]$ для представлення аспектів направленості цього характерного компонента. Елемент синтаксису може бути позначений як "codedVVecLength". Таким чином, блок 52 квантування може сигналізувати або інакше вказувати в бітовому потоці, який з чотирьох режимів конфігурації використовувався для указання кодованих векторів 57 переднього плану $V[k]$ у бітовому потоці. Хоча вони описані відносно чотирьох режимів конфігурації, методи не підлягають обмеженню чотирма режимами конфігурації, але передбачають будь-яку кількість режимів конфігурації, яка включає в себе одиничний режим конфігурації або множину режимів конфігурації. Блок 53 скалярного/ентропійного квантування також може вказувати прапор 63 як інший елемент синтаксису в інформації 57 побічного каналу.

[238] Блок 40 психоакустичного аудіокодера, включений у пристрій 20 аудіокодування, може представляти множинні екземпляри психоакустичного аудіокодера, кожний з яких використовується для кодування окремого аудіооб'єкта або каналу НОА кожного з коефіцієнтів 47' НОА оточення з компенсацією енергії й інтерпольованих сигналів 49' nFG для генерації кодованих коефіцієнтів 59 НОА оточення і кодованих сигналів 61 nFG. Блок 40 психоакустичного аудіокодера може виводити кодовані коефіцієнти 59 НОА оточення і кодовані сигнали 61 nFG на блок 42 генерації бітового потоку.

[239] У ряді випадків, цей блок 40 психоакустичного аудіокодера може представляти один або більше екземплярів блока кодування удосконаленого аудіокодування (AAC). Блок 40 психоакустичного аудіокодера може кодувати кожен стовпець або кожен рядок коефіцієнтів 47' НОА оточення з компенсацією енергії й інтерпольованих сигналів 49' nFG. Часто, блок 40 психоакустичного аудіокодера може викликати екземпляр блока кодування AAC для кожної з комбінацій порядків/підпорядків, що залишилися в коефіцієнтах 47' НОА оточення з компенсацією енергії й інтерпольованих сигналів 49' nFG. Додаткову інформацію про те, як можна кодувати коефіцієнти 31 сферичних гармонік фону з використанням блока кодування AAC, можна знайти в матеріалах з'їзду Eric Hellerud et al., за назвою "Encoding Higher Order Ambisonics with AAC", представлених на 124-ому з'їзді, 17-20 травня 2008 р., і доступних за адресою: <http://ro.uow.edu.au/cgi/viewcontent.cgi?article=8025&context=engpapers>. У ряді випадків, блок 14 аудіокодування може аудіокодувати коефіцієнти 47' НОА оточення з компенсацією енергії з використанням більш низької цільової бітової швидкості, ніж використовували для кодування інтерпольованих сигналів 49' nFG, таким чином, потенційно стискаючи коефіцієнти 47' НОА оточення з компенсацією енергії більше, ніж інтерпольовані сигнали 49' nFG.

[240] Блок 42 генерації бітового потоку, включений у пристрій 20 аудіокодування, представляє блок, який форматує дані для узгодження з відомим форматом (який може означати формат, відомий пристрою декодування), таким чином генеруючи бітовий потік 21 на векторній основі. Блок 42 генерації бітового потоку в деяких прикладах може представляти мультіплексор, який може приймати кодовані вектори 57 переднього плану $V[k]$, кодовані коефіцієнти 59 НОА оточення, кодовані сигнали 61 nFG і інформацію 43 фонового каналу. Потім блок 42 генерації бітового потоку може генерувати бітовий потік 21 на основі кодованих векторів 57 переднього плану $V[k]$, кодованих коефіцієнтів 59 НОА оточення, кодованих сигналів 61 nFG і інформації 43 фонового каналу. Бітовий потік 21 може включати в себе первинний або головний бітовий потік і один або більше бітових потоків побічних каналів.

[241] Хоча це не показано в прикладі, наведеному на фіг. 4, пристрій 20 аудіокодування також може включати в себе блок виведення бітового потоку, який переключає бітовий потік, що виводиться з пристрою 20 аудіокодування (наприклад, між бітовим потоком 21 на основі напрямку і бітовим потоком 21 на векторній основі), на основі того, чи підлягає поточний кадр кодуванню з використанням синтезу на основі напрямку або синтезу на векторній основі. Цей блок виведення бітового потоку може здійснювати це переключення на основі елемента синтаксису, що виводиться блоком 26 аналізу контенту, який вказує, чи здійснювався синтез на основі напрямку (у результаті виявлення, що коефіцієнти 11 НОА згенеровані із синтетичного аудіооб'єкта) або чи здійснювався синтез на векторній основі (у результаті виявлення, що коефіцієнти НОА були записані). Блок виведення бітового потоку може вказувати правильний

синтаксис заголовка для указання цього переключення або поточного кодування, використовуюваного для поточного кадру разом з відповідним одним з бітових потоків 21.

[242] У ряді випадків, різні аспекти методів також можуть передбачати, що пристрій 20 аудіокодування визначає, чи генеруються коефіцієнти 11 НОА із синтетичного аудіооб'єкта. Ці аспекти методів можуть передбачати, що пристрій 20 аудіокодування виконаний з можливістю одержання указання, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

[243] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта.

[244] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю виключення першого вектора з матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, які представляють звукове поле, для одержання приведеної матриці кадрованих коефіцієнтів сферичних гармонік.

[245] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю виключення першого вектора з матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, які представляють звукове поле, для одержання приведеної матриці кадрованих коефіцієнтів сферичних гармонік і прогнозування вектора приведеної матриці кадрованих коефіцієнтів сферичних гармонік на основі векторів приведеної матриці, що залишилися, кадрованих коефіцієнтів сферичних гармонік.

[246] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю виключення першого вектора з матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, які представляють звукове поле, для одержання приведеної матриці кадрованих коефіцієнтів сферичних гармонік і прогнозування вектора приведеної матриці кадрованих коефіцієнтів сферичних гармонік, щонайменше частково, на основі суми векторів приведеної матриці, що залишилися, кадрованих коефіцієнтів сферичних гармонік.

[247] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю прогнозування вектора матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, щонайменше частково, на основі суми векторів матриці, що залишилися, кадрованих коефіцієнтів сферичних гармонік.

[248] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю прогнозування вектора матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, щонайменше частково, на основі суми векторів матриці, що залишилися, кадрованих коефіцієнтів сферичних гармонік і обчислення помилки на основі прогнозованого вектора.

[249] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю прогнозування вектора матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, щонайменше частково, на основі суми векторів матриці, що залишилися, кадрованих коефіцієнтів сферичних гармонік і обчислення помилки на основі прогнозованого вектора і відповідного вектора матриці кадрованих коефіцієнтів сферичних гармонік.

[250] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю прогнозування вектора матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, щонайменше частково, на основі суми векторів матриці, що залишилися, кадрованих коефіцієнтів сферичних гармонік і обчислення помилки як суми абсолютних значень різниці прогнозованого вектора і відповідного вектора матриці кадрованих коефіцієнтів сферичних гармонік.

[251] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю прогнозування вектора матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, щонайменше частково, на основі суми векторів матриці, що залишилися, кадрованих коефіцієнтів сферичних гармонік, обчислення помилки на основі прогнозованого вектора і відповідного вектора матриці кадрованих коефіцієнтів сферичних гармонік, обчислення відношення на основі енергії відповідного вектора матриці кадрованих коефіцієнтів сферичних гармонік і помилки і порівняння відношення з порогом для визначення, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

[252] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю задавання указання в бітовому потоці 21, де зберігаються стиснена версія коефіцієнтів сферичних гармонік.

[253] У ряді випадків, різні методи можуть передбачати, що пристрій 20 аудіокодування здійснює перетворення відносно коефіцієнтів 11 НОА. У цих і інших випадках, пристрій 20 аудіокодування може бути виконаний з можливістю одержання одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів 11 сферичних гармонік.

[254] У цих і інших випадках, у пристрої 20 аудіокодування перетворення включає розкладання на сингулярні значення, що генерує матрицю U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матрицю S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матрицю V , яка представляє праві сингулярні вектори з множини коефіцієнтів 11 сферичних гармонік.

[255] У цих і інших випадках, у пристрої 20 аудіокодування один або більше перших векторів містять один або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S , і при цьому матриця U і матриця S генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік.

[256] У цих і інших випадках, у пристрої 20 аудіокодування один або більше перших векторів містять один або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , і при цьому матриця U , матриця S і матриця V генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів 11 сферичних гармонік.

[257] У цих і інших випадках, у пристрої 20 аудіокодування один або більше перших векторів містять один або більше векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому пристрій 20 аудіокодування додатково виконаний з можливістю одержання значення D , яке вказує кількість векторів, що підлягають витягання з бітового потоку, для формування одного або більше векторів $U_{DIST} * S_{DIST}$ і одного або більше векторів V_{DIST}^T .

[258] У цих і інших випадках, у пристрої 20 аудіокодування один або більше перших векторів містять один або більше векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому пристрій 20 аудіокодування додатково виконаний з можливістю одержання значення D на основі кадрів аудіо, яке вказує кількість векторів, що підлягають витягання з бітового потоку, для формування одного або більше векторів $U_{DIST} * S_{DIST}$ і одного або більше векторів V_{DIST}^T .

[259] У цих і інших випадках, у пристрої 20 аудіокодування перетворення включає аналіз головних компонентів для ідентифікації характерних компонентів звукового поля і фонових компонентів звукового поля.

[260] Різні аспекти методів, описаних у цьому винаході, можуть передбачати пристрій 20 аудіокодування, виконаний з можливістю компенсації помилки квантування.

[261] У ряді випадків, пристрій 20 аудіокодування може бути виконаний з можливістю квантування одного або більше перших векторів, що представляють один або більше компонентів звукового поля, і компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля.

[262] У цих і інших випадках, пристрій аудіокодування виконаний з можливістю квантування одного або більше векторів із транспонованої матриці V , генерованої, щонайменше частково, шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[263] У цих і інших випадках, пристрій аудіокодування додатково виконаний з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє

сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і виконаний з можливістю квантування одного або більше векторів із транспонованої матриці V .

[264] У цих і інших випадках, пристрій аудіокодування додатково виконаний з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, виконаний з можливістю квантування одного або більше векторів з транспонованої матриці V і виконаний з можливістю компенсації помилки, що вноситься внаслідок квантування в один або більше векторів U^*S , обчислених шляхом множення одного або більше векторів U матриці U на один або більше векторів S матриці S .

[265] У цих і інших випадках, пристрій аудіокодування додатково виконаний з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів U_{DIST} матриці U , кожний з яких відповідає характерному компоненту звукового поля, визначення одного або більше векторів S_{DIST} матриці S , кожний з яких відповідає одному і тому ж характерному компоненту звукового поля, і визначення одного або більше векторів V_{DIST}^T транспонованої матриці V , кожний з яких відповідає одному і тому ж характерному компоненту звукового поля, виконаний з можливістю квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{Q_DIST}^T$ і виконаний з можливістю компенсації помилки, що вноситься внаслідок квантування в один або більше векторів $U_{DIST}^*S_{DIST}$, обчислених шляхом множення одного або більше векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S , для генерації одного або більше векторів $U_{DIST}^*S_{DIST}$ з компенсацією помилки.

[266] У цих і інших випадках, пристрій аудіокодування виконаний з можливістю визначення характерних коефіцієнтів сферичних гармонік на основі одного або більше векторів U_{DIST} , одного або більше векторів S_{DIST} і одного або більше векторів V_{DIST}^T і здійснення псевдообернення відносно векторів $V_{Q_DIST}^T$ для поділу характерних коефіцієнтів сферичних гармонік на один або більше векторів $V_{Q_DIST}^T$ і, таким чином, генерації одного або більше векторів $U_{C_DIST}^*S_{C_DIST}$ з компенсацією помилки, які компенсують, щонайменше частково, помилку, внесену за допомогою квантування векторів V_{DIST}^T .

[267] У цих і інших випадках, пристрій аудіокодування додатково виконаний з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів U_{BG} матриці U , які описують один або більше фонових компонентів звукового поля, і одного або більше векторів U_{DIST} матриці U , які описують один або більше характерних компонентів звукового поля, визначення одного або більше векторів S_{BG} матриці S , які описують один або більше фонових компонентів звукового поля, і одного або більше векторів S_{DIST} матриці S , які описують один або більше характерних компонентів звукового поля, і визначення одного або більше векторів V_{DIST}^T і одного або більше векторів V_{BG}^T транспонованої матриці V , причому вектори V_{DIST}^T описують один або більше характерних компонентів звукового поля і V_{BG}^T описують один або більше фонових компонентів звукового поля, виконаний з можливістю квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{Q_DIST}^T$ і виконаний з можливістю компенсації помилки, що вноситься внаслідок квантування в коефіцієнти сферичних гармонік фону, сформовані шляхом множення одного або більше векторів U_{BG} на один або більше векторів S_{BG} і потім на один або більше векторів V_{BG}^T , для генерації коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[268] У цих і інших випадках, пристрій аудіокодування виконаний з можливістю визначення помилки на основі векторів V_{DIST}^T і одного або більше векторів $U_{DIST}^*S_{DIST}$, сформованих шляхом множення векторів U_{DIST} на вектори S_{DIST} , і додавання визначеної помилки до коефіцієнтів сферичних гармонік фону для генерації коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[269] У цих і інших випадках, пристрій аудіокодування виконаний з можливістю компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, які також представляють ті ж один або більше компонентів звукового поля, для генерації одного або більше других векторів з компенсацією помилки і додатково виконаний з можливістю генерації бітового потоку для включення одного або більше других векторів з компенсацією помилки і квантованих одного або більше перших векторів.

[270] У цих і інших випадках, пристрій аудіокодування виконаний з можливістю компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, які також представляють ті ж один або більше компонентів звукового поля, для генерації одного або більше других векторів з компенсацією помилки і додатково виконаний з можливістю аудіокодування одного або більше других векторів з компенсацією помилки і генерації бітового потоку для включення аудіокодованих одного або більше других векторів з компенсацією помилки і квантованих одного або більше перших векторів.

[271] Різні аспекти методів можуть додатково передбачати, що пристрій 20 аудіокодування генерує приведені коефіцієнти сферичних гармонік або їх розкладання. У ряді випадків, пристрій 20 аудіокодування може бути виконаний з можливістю здійснення, на основі цільової бітової швидкості, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань, причому множина коефіцієнтів сферичних гармонік представляє звукове поле.

[272] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю, до здійснення зниження порядку, здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік для ідентифікації одного або більше перших векторів, які описують характерні компоненти звукового поля, й одного або більше других векторів, які ідентифікують фонові компоненти звукового поля, і виконаний з можливістю здійснення зниження порядку відносно одного або більше перших векторів, одного або більше других векторів або одного або більше перших векторів і одного або більше других векторів.

[273] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю здійснення аналізу контенту відносно множини коефіцієнтів сферичних гармонік або їх розкладань і виконаний з можливістю здійснення, на основі цільової бітової швидкості й аналізу контенту, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

[274] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю здійснення просторового аналізу відносно множини коефіцієнтів сферичних гармонік або їх розкладань.

[275] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю здійснення аналізу дифузії відносно множини коефіцієнтів сферичних гармонік або їх розкладань.

[276] У цих і інших випадках, пристрій 20 аудіокодування є одним або більше процесорами, виконаними з можливістю здійснення просторового аналізу й аналізу дифузії відносно множини коефіцієнтів сферичних гармонік або їх розкладань.

[277] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю указання одного або більше порядків і/або одного або більше підпорядків сферичних базисних функцій, яким відповідають ті з приведених коефіцієнтів сферичних гармонік або їх приведених розкладань у бітовому потоці, які включає в себе приведені коефіцієнти сферичних гармонік або їх приведених розкладань.

[278] У цих і інших випадках, приведені коефіцієнти сферичних гармонік або їх приведених розкладань мають менше значень, ніж множина коефіцієнтів сферичних гармонік або їх розкладань.

[279] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю видалення тих з множини коефіцієнтів сферичних гармонік або векторів їх розкладань, які мають зазначений порядок і/або підпорядок, для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

[280] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю обнулення тих з множини коефіцієнтів сферичних гармонік або тих векторів їх розкладань, які мають зазначений порядок і/або підпорядок, для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

[281] Різні аспекти методів також можуть передбачати, що пристрій 20 аудіокодування виконаний з можливістю представлення характерних компонентів звукового поля. У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю одержання першого ненульового набору коефіцієнтів вектора, що підлягають використанню для представлення характерного

компонента звукового поля, причому вектор розкладається з множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[282] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення першого ненульового набору коефіцієнтів вектора для включення всіх коефіцієнтів.

5 [283] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення першого ненульового набору коефіцієнтів, як тих з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

10 [284] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення першого ненульового набору коефіцієнтів для включення тих з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік, і виключення щонайменше одного з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

15 [285] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення першого ненульового набору коефіцієнтів для включення всіх коефіцієнтів за винятком щонайменше одного з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

20 [286] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю указання першого ненульового набору коефіцієнтів вектора в інформації побічного каналу.

[287] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю указання першого ненульового набору коефіцієнтів вектора в інформації побічного каналу без аудіокодування першого ненульового набору коефіцієнтів вектора.

25 [288] У цих і інших випадках, вектор містить вектор, розкладений з множини коефіцієнтів сферичних гармонік з використанням векторного синтезу.

[289] У цих і інших випадках, векторний синтез включає розкладання на сингулярні значення.

[290] У цих і інших випадках, вектор містить вектор V , розкладений з множини коефіцієнтів сферичних гармонік з використанням розкладання на сингулярні значення.

30 [291] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю вибору одного з множини режимів конфігурації, що дозволяють вказувати ненульовий набір коефіцієнтів вектора, і вказування ненульового набору коефіцієнтів вектора на основі вибраного одного з множини режимів конфігурації.

35 [292] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти.

[293] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

40 [294] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік, і не включає в себе щонайменше один з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

45 [295] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти за винятком щонайменше одного з коефіцієнтів.

50 [296] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю указання вибраного одного з множини режимів конфігурації в бітовому потоці.

[297] Різні аспекти методів, описаних у цьому винаході, також можуть передбачати, що пристрій 20 аудіокодування виконаний з можливістю представлення цього характерного компонента звукового поля тим або іншим чином. У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю одержання першого ненульового набору коефіцієнтів вектора, які представляють характерний компонент звукового поля, причому вектор розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[298] У цих і інших випадках, перший ненульовий набір коефіцієнтів включає в себе всі коефіцієнти вектора.

[299] У цих і інших випадках, перший ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

5 [300] У цих і інших випадках, перший ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік, і не включає в себе щонайменше один з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

10 [301] У цих і інших випадках, перший ненульовий набір коефіцієнтів включає в себе всі коефіцієнти за винятком щонайменше одного з коефіцієнтів, ідентифікованих як такі, що не мають достатньої інформації направленості.

[302] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю витягання першого ненульового набору коефіцієнтів як першої частини вектора.

15 [303] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю витягання першого ненульового набору вектора з інформації побічного каналу і одержання реконструйованої версії з множини коефіцієнтів сферичних гармонік на основі першого ненульового набору коефіцієнтів вектора.

[304] У цих і інших випадках, вектор містить вектор, розкладений з множини коефіцієнтів сферичних гармонік з використанням векторного синтезу.

20 [305] У цих і інших випадках, векторний синтез включає розкладання на сингулярні значення.

[306] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення одного з множини режимів конфігурації, що дозволяє витягати ненульовий набір коефіцієнтів вектора відповідно до одного з множини режимів конфігурації, і витягання ненульового набору коефіцієнтів вектора на основі одержаного одного з множини режимів конфігурації.

25 [307] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти.

[308] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

[309] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік, і не включає в себе щонайменше один з коефіцієнтів, які відповідають порядку, що перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

30 [310] У цих і інших випадках, один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти за винятком щонайменше одного з коефіцієнтів.

[311] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення одного з множини режимів конфігурації на основі значення, сигналізованого в бітовому потоці.

35 [312] Різні аспекти методів також можуть, у ряді випадків, передбачати, що пристрій 20 аудіокодування ідентифікує один або більше характерних аудіооб'єктів (або, іншими словами, переважних аудіооб'єктів). У ряді випадків, пристрій 20 аудіокодування може бути виконаний з можливістю ідентифікації одного або більше характерних аудіооб'єктів з одного або більше коефіцієнтів сферичних гармонік (SHC), пов'язаних з аудіооб'єктами, на основі направленості, визначеної для одного або більше з аудіооб'єктів.

40 [313] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення направленості одного або більше аудіооб'єктів на основі коефіцієнтів сферичних гармонік, пов'язаних з аудіооб'єктами.

[314] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю здійснювати розкладання на сингулярні значення відносно коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і представляти множину коефіцієнтів сферичних гармонік як функцію щонайменше частини однієї або більше з матриці U , матриці S і матриці V , причому пристрій 20

аудіокодування виконаний з можливістю визначення відповідної направленості одного або більше аудіооб'єктів на основі, щонайменше частково, матриці V .

[315] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю преупорядковування одного або більше векторів матриці V таким чином, що вектори, які мають більш високий показник направленості, розташовуються над векторами, які мають більш низький показник направленості, у переупорядкованій матриці V .

[316] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення, що вектори, які мають більш високий показник направленості, включають в себе більше інформації направленості, ніж вектори, які мають більш низький показник направленості.

[317] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю множення матриці V на матрицю S для генерації матриці VS , причому матриця VS включає в себе один або більше векторів.

[318] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю вибору елементів кожного рядка матриці VS , що пов'язані з порядком, більшим 14, піднесення до квадрата кожного з вибраних елементів для формування відповідних квадратів елементів i , для кожного рядка матриці VS , підсумовування квадратів всіх елементів для визначення показника направленості для відповідного вектора.

[319] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю вибору елементів кожного рядка матриці VS , пов'язаних з порядком, більшим 14, який включає вибір всіх елементів, починаючи з 18-го елемента кожного рядка матриці VS і закінчуючи 38-им елементом кожного рядка матриці VS .

[320] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю вибору піднабору векторів матриці VS для представлення характерних аудіооб'єктів.

У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю вибору чотирьох векторів матриці VS , і при цьому вибрані чотири вектори мають чотири найбільших показники направленості з усіх векторів матриці VS .

[321] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення, що вибраний піднабір векторів представляє характерні аудіооб'єкти, на основі направленості й енергії кожного вектора.

[322] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю здійснення порівняння енергії між одним або більше з перших векторів і одним або більше з других векторів, що представляють характерні аудіооб'єкти, для визначення переупорядкованих одного або більше перших векторів, причому один або більше перших векторів описують характерні аудіооб'єкти в першій частині аудіоданих і один або більше других векторів описують характерні аудіооб'єкти в другій частині аудіоданих.

[323] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю здійснення крос-кореляції між одним або більше з перших векторів і одним або більше з других векторів, що представляють характерні аудіооб'єкти, для визначення переупорядкованих одного або більше перших векторів, причому один або більше перших векторів описують характерні аудіооб'єкти в першій частині аудіоданих і один або більше других векторів описують характерні аудіооб'єкти в другій частині аудіоданих.

[324] Різні аспекти методів також можуть, у ряді випадків, передбачати, що пристрій 20 аудіокодування виконаний з можливістю здійснення компенсації енергії відносно розкладань коефіцієнтів 11 НОА. У цих і інших випадках, пристрій 20 аудіокодування може бути виконаний з можливістю здійснення синтезу на векторній основі відносно множини коефіцієнтів сферичних гармонік для генерації розкладених представлень з множини коефіцієнтів сферичних гармонік, що представляють один або більше аудіооб'єктів і відповідну інформацію направленості, причому коефіцієнти сферичних гармонік пов'язані з порядком і описують звукове поле, визначення окремої і фонові інформації направленості з інформації направленості, зниження порядку інформації направленості, пов'язаної з фоновими аудіооб'єктами, для генерації перетвореної інформації направленості фону, застосування компенсації для збільшення значень перетвореної інформації направленості для збереження повної енергії звукового поля.

[325] У цих і інших випадках, пристрій 20 аудіокодування може бути виконаний з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік для генерації матриці U і матриці S , які представляють аудіооб'єкти, і матриці V , яка представляє інформацію направленості, визначення характерних векторів-стовпців матриці V і векторів-стовпців фону матриці V , зниження порядку векторів-стовпців фону матриці V для генерації перетворених векторів-стовпців фону матриці V і застосування компенсації для збільшення значень перетворених векторів-стовпців фону матриці V для збереження повної енергії звукового поля.

[326] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення кількості помітних сингулярних значень матриці S , причому кількість характерних векторів-стовпців матриці V дорівнює кількості помітних сингулярних значень матриці S .

5 [327] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення зниженого порядку для коефіцієнтів сферичних гармонік і нульових значень для рядків векторів-стовпців фону матриці V , пов'язаних з порядком, що перевищує знижений порядок.

10 [328] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю об'єднання фонових стовпців матриці U , фонових стовпців матриці S і транспонованих перетворених фонових стовпців матриці V для генерації модифікованих коефіцієнтів сферичних гармонік.

[329] У цих і інших випадках, модифіковані коефіцієнти сферичних гармонік описують один або більше фонових компонентів звукового поля.

15 [330] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення першої енергії вектора векторів-стовпців фону матриці V і другої енергії вектора перетворених векторів-стовпців фону матриці V і застосування значення посилення до кожного елемента вектора перетворених векторів-стовпців фону матриці V , причому значення посилення містить відношення першої енергії до другої енергії.

20 [331] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення першої середньоквадратичної енергії вектора векторів-стовпців фону матриці V і другої середньоквадратичної енергії вектора перетворених векторів-стовпців фону матриці V і застосування значення посилення до кожного елемента вектора перетворених векторів-стовпців фону матриці V , причому значення посилення містить відношення першої енергії до другої енергії.

25 [332] Різні аспекти методів, описаних у цьому винаході, також можуть передбачати, що пристрій 20 аудіокодування здійснює інтерполяцію відносно розкладених версій коефіцієнтів 11 НОА. У ряді випадків, пристрій 20 аудіокодування може бути виконаний з можливістю одержання розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, щонайменше частково, шляхом здійснення інтерполяції відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік.

[333] У цих і інших випадках, перше розкладання містить першу матрицю V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

35 [334] У цих і інших випадках, друге розкладання містить другу матрицю V , яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[335] У цих і інших випадках, перше розкладання містить першу матрицю V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік, і друге розкладання містить другу матрицю V , яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

40 [336] У цих і інших випадках, відрізок часу містить підкадр кадру аудіо.

[337] У цих і інших випадках, відрізок часу містить часову вибірку кадру аудіо.

[338] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю одержання інтерпольованого розкладання першого розкладання і другого розкладання для коефіцієнта сферичної гармоніки з першої множини коефіцієнтів сферичних гармонік.

45 [339] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю одержання інтерпольованих розкладань першого розкладання для першої частини першої множини коефіцієнтів сферичних гармонік, включених у перший кадр, і другого розкладання для другої частини другої множини коефіцієнтів сферичних гармонік, включених у другий кадр, і пристрій 20 аудіокодування додатково виконаний з можливістю застосування інтерпольованих розкладань до першого часового компонента першої частини першої множини коефіцієнтів сферичних гармонік, включених у перший кадр, для генерації першого штучного часового компонента першої множини коефіцієнтів сферичних гармонік і застосування відповідних інтерпольованих розкладань до другого часового компонента другої частини другої множини коефіцієнтів сферичних гармонік, включених у другий кадр, для генерації другого штучного часового компонента другої множини коефіцієнтів сферичних гармонік.

[340] У цих і інших випадках, перший часовий компонент генерується шляхом здійснення синтезу на векторній основі відносно першої множини коефіцієнтів сферичних гармонік.

[341] У цих і інших випадках, другий часовий компонент генерується шляхом здійснення синтезу на векторній основі відносно другої множини коефіцієнтів сферичних гармонік.

[342] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю прийому першого штучного часового компонента і другого штучного часового компонента, обчислення інтерпольованих розкладань першого розкладання для першої частини першої множини коефіцієнтів сферичних гармонік і другого розкладання для другої частини другої множини коефіцієнтів сферичних гармонік і застосування обернених інтерпольованих розкладань до першого штучного часового компонента для відновлення першого часового компонента і до другого штучного часового компонента для відновлення другого часового компонента.

[343] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю інтерполяції першого просторового компонента першої множини коефіцієнтів сферичних гармонік і другого просторового компонента другої множини коефіцієнтів сферичних гармонік.

[344] У цих і інших випадках, перший просторовий компонент містить першу матрицю U , яка представляє ліві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

[345] У цих і інших випадках, другий просторовий компонент містить другу матрицю U , яка представляє ліві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[346] У цих і інших випадках, перший просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для першої множини коефіцієнтів сферичних гармонік і другий просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для другої множини коефіцієнтів сферичних гармонік.

[347] У цих і інших випадках, перший просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для першої множини коефіцієнтів сферичних гармонік і другий просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для другої множини коефіцієнтів сферичних гармонік, і пристрій 20 аудіокодування виконаний з можливістю інтерполяції останніх N елементів першого просторового компонента і перших N елементів другого просторового компонента.

[348] У цих і інших випадках, друга множина коефіцієнтів сферичних гармонік іде за першою множиною коефіцієнтів сферичних гармонік у часовій області.

[349] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю розкладання першої множини коефіцієнтів сферичних гармонік для генерації першого розкладання першої множини коефіцієнтів сферичних гармонік.

[350] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю розкладання другої множини коефіцієнтів сферичних гармонік для генерації другого розкладання другої множини коефіцієнтів сферичних гармонік.

[351] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю здійснення розкладання на сингулярні значення відносно першої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори першої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення першої множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

[352] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю здійснення розкладання на сингулярні значення відносно другої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори другої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення другої множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[353] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік представляє плоскостерпне представлення звукового поля.

[354]

[355] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік представляють один або більше змішаних один з одним монофонічних аудіооб'єктів.

[356] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік містять відповідні перший і другий коефіцієнти сферичних гармонік, що представляють тривимірне звукове поле.

[357] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, що має порядок, який більше одиниці.

[358] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, що має порядок, який дорівнює чотирьом.

[359] У цих і інших випадках, інтерполяція є зваженою інтерполяцією першого розкладання і другого розкладання, причому вагові коефіцієнти зваженої інтерполяції, застосовувані до першого розкладання, зворотно пропорційні часу, представленому векторами першого і другого розкладання, і при цьому вагові коефіцієнти зваженої інтерполяції, застосовувані до другого розкладання, пропорційні часу, представленому векторами першого і другого розкладання.

[360] У цих і інших випадках, розкладені інтерпольовані коефіцієнти сферичних гармонік згладжують щонайменше один із просторових компонентів і часових компонентів першої множини коефіцієнтів сферичних гармонік і другої множини коефіцієнтів сферичних гармонік.

[361] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю обчислення $Us[n]=HOA(n)*(V_vec[n])-1$ для одержання скаляра.

[362] У цих і інших випадках, інтерполяція включає лінійну інтерполяцію. У цих і інших випадках, інтерполяція включає нелінійну інтерполяцію. У цих і інших випадках, інтерполяція включає косинусну інтерполяцію. У цих і інших випадках, інтерполяція включає зважену косинусну інтерполяцію. У цих і інших випадках, інтерполяція включає кубічну інтерполяцію. У цих і інших випадках, інтерполяція включає адаптивну сплайнову інтерполяцію. У цих і інших випадках, інтерполяція включає інтерполяцію мінімальної кривизни.

[363] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю генерації бітового потоку, що включає в себе представлення розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, і указання типу інтерполяції.

[364] У цих і інших випадках, указання включає один або більше бітів, що відображаються в тип інтерполяції.

[365] Таким чином, різні аспекти методів, описаних у цьому винаході, можуть передбачати, що пристрій 20 аудіокодування виконаний з можливістю одержання бітового потоку, що включає в себе представлення розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, і указання типу інтерполяції.

[366] У цих і інших випадках, указання включає один або більше бітів, що відображаються в тип інтерполяції.

[367] У цьому відношенні, пристрій 20 аудіокодування може представляти один варіант здійснення методів, згідно з якими пристрій 20 аудіокодування можемо, у ряді випадків, бути виконаний з можливістю генерації бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[368] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю генерації бітового потоку для включення поля, яке вказує режим прогнозування, використовуваний при стисненні просторового компонента.

[369] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю генерації бітового потоку для включення інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.

[370] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю генерації бітового потоку для включення поля, яке вказує значення, що виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.

[371] У цих і інших випадках, значення містить значення $nbits$.

[372] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю генерації бітового потоку для включення стисненої версії з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, де значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.

[373] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю генерації бітового потоку для включення коду Хаффмана для представлення ідентифікатора категорії, який ідентифікує категорію стиснення, якій відповідає просторовий компонент.

[374] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю генерації бітового потоку для включення знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

[375] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю генерації бітового потоку для включення коду Хаффмана для представлення залишкового значення просторового компонента.

[376] У цих і інших випадках, векторний синтез включає розкладання на сингулярні значення.

[377] У цьому відношенні, пристрій 20 аудіокодування може додатково реалізувати різні аспекти методів, згідно з якими пристрій 20 аудіокодування можемо, у ряді випадків, бути виконаний з можливістю ідентифікації кодової книги Хаффмана для використання при стисненні

просторового компонента з множини просторових компонентів на основі порядку просторового компонента відносно просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

5 [378] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю ідентифікації кодової книги Хаффмана на основі режиму прогнозування, використовуваного при стисненні просторового компонента.

[379] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, інформації таблиці Хаффмана, що ідентифікує кодову книгу Хаффмана.

10 [380] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, яке вказує значення, що виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.

15 [381] У цих і інших випадках, значення містить значення nbits.

[382] У цих і інших випадках, бітовий потік містить стиснену версію з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.

20 [383] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана, вибраного з ідентифікованої кодової книги Хаффмана, для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.

25 [384] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

[385] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана, вибраного з ідентифікованої кодової книги Хаффмана, для представлення залишкового значення просторового компонента.

30 [386] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю стиснення просторового компонента на основі ідентифікованої кодової книги Хаффмана для генерації стисненої версії просторового компонента і генерації бітового потоку для включення стисненої версії просторового компонента.

35 [387] Крім того, пристрій 20 аудіокодування може, у ряді випадків, реалізувати різні аспекти методів, згідно з якими пристрій 20 аудіокодування може бути виконаний з можливістю визначення розміру кроку квантування, що підлягає використанню при стисненні просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

40 [388] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення розміру кроку квантування на основі цільової бітової швидкості.

[389] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, і визначення розміру кроку квантування на основі різниці між оцінкою і цільовою бітовою швидкістю.

45 [390] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, визначення різниці між оцінкою і цільовою бітовою швидкістю і визначення розміру кроку квантування шляхом додавання різниці до цільової бітової швидкості.

50 [391] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю обчислення оціненої кількості бітів, що підлягають генерації для просторового компонента, за умови, що кодова книга відповідає цільовій бітій швидкості.

[392] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю обчислення оціненої кількості бітів, що підлягають генерації для просторового компонента, при даному режимі кодування, використовуваному при стисненні просторового компонента.

55 [393] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю обчислення першої оцінки кількості бітів, що підлягають генерації для просторового компонента, за умови, що перший режим кодування підлягає використанню при стисненні просторового компонента, обчислення другої оцінки кількості бітів, що підлягають генерації для просторового компонента, за умови, що другий режим кодування підлягає використанню при стисненні

просторового компонента, вибору однієї з першої оцінки і другої оцінка, що має найменшу кількість бітів, що підлягають використанню, як визначеної оцінки кількості бітів.

[394] У цих і інших випадках, пристрій 20 аудіокодування виконаний з можливістю ідентифікації ідентифікатора категорії, що ідентифікує категорію, якій відповідає просторовий компонент, ідентифікації бітової довжини залишкового значення для просторового компонента, яке було б одержане при стисненні просторового компонента, що відповідає категорії, і визначення оцінки кількості бітів шляхом, щонайменше частково, додавання кількості бітів, використовуваних для представлення ідентифікатора категорії, до бітової довжини залишкового значення.

[395] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю вибору однієї з множини кодових книг, що підлягають використанню при стисненні просторового компонента.

[396] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, з використанням кожної з множини кодових книг і вибору однієї з множини кодових книг, яка привела до визначеної оцінки, що має найменшу кількість бітів.

[397] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, з використанням однієї або більше з множини кодових книг, причому одна або більше з множини кодових книг вибрана на основі порядку елементів просторового компонента, що підлягають стисненню, відносно інших елементів просторового компонента.

[398] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент не прогнозується з наступного просторового компонента.

[399] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент прогнозується з наступного просторового компонента.

[400] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент представляє синтетичний аудіооб'єкт у звуковому полі.

[401] У цих і інших випадках, синтетичний аудіооб'єкт містить аудіооб'єкт, підданий імпульсно-кодовій модуляції (IKM).

[402] У цих і інших випадках, пристрій 20 аудіокодування додатково виконаний з можливістю визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент представляє записаний аудіооб'єкт у звуковому полі.

[403] У кожному з різних вищеписаних випадків, варто розуміти, що пристрій 20 аудіокодування може здійснювати спосіб або іншим чином містити засіб для здійснення кожного етапу способу, який має здійснювати пристрій 20 аудіокодування. У ряді випадків, цей засіб може містити один або більше процесорів. У ряді випадків, один або більше процесорів можуть представляти процесор спеціального призначення, що виконує інструкції, які зберігаються на довгостроковому зчитуваному комп'ютером носії даних. Іншими словами, різні аспекти методів у кожному з наборів прикладів кодування можуть передбачати довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати спосіб, для виконання якого призначений пристрій 20 аудіокодування.

[404] На фіг. 5 більш детально показана блок-схема, що демонструє пристрій 24 аудіодекодування, представлений на фіг. 3. Як показано в прикладі, наведеному на фіг. 5, пристрій 24 аудіодекодування може включати в себе блок 72 витягання, блок 90 реконструкції на основі направленості і блок 92 реконструкції на векторній основі.

[405] Блок 72 витягання може представляти блок, виконаний з можливістю прийому бітового потоку 21 і витягання різних кодованих версій (наприклад, кодової версії на основі напрямку або кодової версії на векторній основі) коефіцієнтів 11 НОА. Блок 72 витягання може визначати з вищезгаданого елемента синтаксису (наприклад, елемента синтаксису ChannelType, показаного в прикладах, наведених на фіг. 10E і 10H(i)-10O(ii)), чи кодовані коефіцієнти 11 НОА за допомогою різних версій. Коли кодування на основі напрямку здійснене, блок 72 витягання може витягати версію на основі напрямку коефіцієнтів 11 НОА і елементи

синтаксису, пов'язані з цією кодовою версією (яка позначається як інформація 91 на основі напрямку в прикладі, наведеному на фіг. 5), передавати цю інформацію 91 на основі направленості на блок 90 реконструкції на основі напрямку. Цей блок 90 реконструкції на основі напрямку може представляти блок, виконаний з можливістю реконструкції коефіцієнтів НОА у формі коефіцієнтів 11' НОА на основі інформації 91 на основі напрямку. Бітовий потік і розміщення елементів синтаксису в бітовому потоці описані нижче більш детально відносно прикладу, наведеного на фіг. 10-10O(ii) і 11.

[406] Коли елемент синтаксису вказує, що коефіцієнти 11 НОА були кодовані з використанням синтезу на векторній основі, блок 72 витягання може витягати кодовані вектори 57 переднього плану $V[k]$, кодовані коефіцієнти 59 НОА оточення і кодовані сигнали 59 nFG. Блок 72 витягання може передавати кодовані вектори 57 переднього плану $V[k]$ на блок 74 квантування і кодовані коефіцієнти 59 НОА оточення разом з кодованими сигналами 61 nFG на блок 80 психоакустичного декодування.

[407] Для витягання кодованих векторів 57 переднього плану $V[k]$, кодованих коефіцієнтів 59 НОА оточення і кодованих сигналів 59 nFG, блок 72 витягання може одержувати інформацію 57 побічного каналу, що включає в себе елемент синтаксису, позначений codedVVecLength. Блок 72 витягання може виділяти codedVVecLength з інформації 57 побічного каналу. Блок 72 витягання може бути виконаний з можливістю діяти в будь-якому з вищеописаних режимів конфігурації на основі елемента синтаксису codedVVecLength.

[408] Потім блок 72 витягання діє відповідно до будь-якого з режимів конфігурації для виділення стисненої форми приведених векторів 55_k переднього плану $V[k]$ з інформації 57 побічного каналу. Блок 72 витягання може діяти відповідно до оператора переключення, представленого у наступному псевдокоді, відповідно до синтаксису, представленому в нижченаведеній таблиці синтаксису для VVectorData:

```

switch CodedVVecLength {
  case 0:
    VVecLength=NumOfHoaCoeffs;
    for(m=0; m<VVecLength; ++m){
      VVecCoeffIdx[m]=m;
    }
    break;
  case 1:
    VVecLength=NumOfHoaCoeffs-MinNumOfCoeffsForAmbHOA-NumOfContAddHoaChans;
    n=0;
    for(m=MinNumOfCoeffsForAmbHOA; m<NumOfHoaCoeffs; ++m){
      CoeffIdx=m+1;
      if(CoeffIdx isNotMemberOf ContAddHoaCoeff){
        VVecCoeffIdx[n]=CoeffIdx-1;
        n++;
      }
    }
    break;
  case 2:
    VVecLength=NumOfHoaCoeffs-MinNumOfCoeffsForAmbHOA;
    for(m=0; m<VVecLength; ++m){
      VVecCoeffIdx[m]=m+MinNumOfCoeffsForAmbHOA;
    }
    break;
  case 3:
    VVecLength=NumOfHoaCoeffs-NumOfContAddHoaChans;
    n=0;
    for(m=0; m<NumOfHoaCoeffs; ++m){
      c=m+1;
      if(3 isNotMemberOf ContAddHoaCoeff){
        VVecCoeffIdx[n]=c-1;
        n++;
      }
    }
  }
}

```

Синтаксис	Кількість бітів	Мнемоніка
<pre> VVectorData(i) { if (NbbitsQ(k)[i] == 5){ for (m=0; m< VVecLength; ++m){ VVec[i][VVecCoeffId[m]](k) = (VecVal / 128.0) – 1.0; } } elseif(NbbitsQ(k)[i] >= 6){ for (m=0; m< VVecLength; ++m){ huffIdx = huffSelect(VVecCoeffId[m], PFlag[i], CbFlag[i]); cid = huffDecode(NbbitsQ[i], huffIdx, huffVal); aVal[i][m] = 0.0; if (cid > 0) { aVal[i][m] = sgn = (sgnVal * 2) - 1; if (cid > 1) { aVal[i][m] = sgn * (2.0^(cid - 1) + intAddVal); } } VVec[i][VVecCoeffId[m]](k) = aVal[i][m] *(2^(16 – NbbitsQ(k)[i])*aVal[i][m])/2^15; if (PFlag(k)[i] == 1) { VVec[i][VVecCoeffId[m]](k)+= VVec[i][VVecCoeffId[m]](k-1) } } } } </pre>		
	8	uimbsbf
	динамічний	huffDecode
	1	bslbf
	cid-1	uimbsbf

- [409] У вищенаведеній таблиці синтаксису, перший оператор переключення з чотирма випадками (випадками 0-3) забезпечує можливість визначення довжини вектора V_{DIST}^T відносно числа (VVecLength) і індексів коефіцієнтів (VVecCoeffId). Перший випадок, case 0, указує, що задані всі коефіцієнти для векторів V_{DIST}^T (NumOfHoaCoeffs). Другий випадок, case 1, указує, що задаються тільки коефіцієнти вектора V_{DIST}^T , що відповідають кількості, яка більше MinNumOfCoeffsForAmbHOA, які можуть позначати те, що іменується вище $(N_{DIST}+1)^2 - (N_{BG}+1)^2$. Крім того, коефіцієнти NumOfContAddAmbHoaChan, ідентифіковані в ContAddAmbHoaChan, віднімаються. Список ContAddAmbHoaChan указує додаткові канали (де "канали" означають конкретний коефіцієнт, що відповідає визначеному порядку, комбінації підпорядків), які відповідають порядку, що перевищує порядок MinAmbHoaOrder. Третій випадок, case 2, указує, що задаються коефіцієнти вектора V_{DIST}^T , що відповідають кількості, яка більше MinNumOfCoeffsForAmbHOA, які можуть позначати те, що іменується вище $(N_{DIST}+1)^2 - (N_{BG}+1)^2$. Четвертий випадок, case 3, указує, що задаються коефіцієнти вектора V_{DIST}^T , які залишилися після видалення коефіцієнтів, ідентифікованих як NumOfContAddAmbHoaChan. Обидва VVecLength, а також список VVecCoeffId дійсні для всіх VVectors у HOAFrame.
- [410] Після цього оператора переключення, рішення, чи здійснювати однорідне деквантування, може керуватися NbbitsQ (або, як позначено вище, nbits), яке, якщо дорівнює 5, однорідне 8-бітове скалярне деквантування здійснюється. Навпаки, значення NbbitsQ, яке більше або дорівнює 6, може приводити до застосування хатфманівського декодування. Вищезгадане значення cid може дорівнювати двом молодшим бітам значення NbbitsQ. Розглянутий вище режим прогнозування позначається у вищенаведеній таблиці синтаксису як PFlag, тоді як біт інформації НТ позначається у вищенаведеній таблиці синтаксису як CbFlag.
- Синтаксис, що залишився, указує, як декодування виникає по суті, аналогічним чином

описаному вище. Різні приклади бітового потоку 21, що узгоджується з кожним з різних вищезгаданих випадків, більш докладно описані нижче з посиланням на фіг. 10H(i)-10O(ii).

[411] Блок 92 реконструкції на векторній основі представляє блок, виконаний з можливістю здійснення операцій, зворотних описаним вище відносно блока 27 синтезу на векторній основі, для реконструкції коефіцієнтів 11' НОА. Блок 92 реконструкції на векторній основі може включати в себе блок 74 квантування, блок 76 просторово-часової інтерполяції, блок 78 формування переднього плану, блок 80 психоакустичного декодування, блок 82 формування коефіцієнтів НОА і блок 84 переупорядковування.

[412] Блок 74 квантування може представляти блок, виконаний з можливістю діяти в режимі, зворотному блоку 52 квантування, показаному в прикладі, наведеному на фіг. 4, щоб деквантувати кодовані вектори 57 переднього плану $V[k]$ і, таким чином, генерувати приведені вектори 55_k переднього плану $V[k]$. Блок 74 деквантування може, у деяких прикладах, здійснювати форму ентропійного декодування і скалярного деквантування в режимі, зворотному описаному вище відносно блока 52 квантування. Блок 74 деквантування може пересилати приведені вектори 55_k переднього плану $V[k]$ на блок 84 переупорядковування.

[413] Блок 80 психоакустичного декодування може діяти в режимі, зворотному блоку 40 психоакустичного аудіокодування, показаному в прикладі, наведеному на фіг. 4, щоб декодувати кодовані коефіцієнти 59 НОА оточення і кодовані сигнали 61 nFG і, таким чином, генерувати коефіцієнти 47' НОА оточення з компенсацією енергії й інтерпольованих сигналів 49' nFG (які також можуть іменуватися інтерпольованими аудіооб'єктами 49' nFG). Блок 80 психоакустичного декодування може передавати коефіцієнти 47' НОА оточення з компенсацією енергії на блок 82 формування коефіцієнтів НОА і сигнали 49' nFG на блок 84 переупорядковування.

[414] Блок 84 переупорядковування може представляти блок, виконаний з можливістю діяти аналогічно зворотному описаному вище відносно блока 34 переупорядковування. Блок 84 переупорядковування може приймати елементи синтаксису, що указують вихідний порядок компонентів переднього плану коефіцієнтів 11 НОА. Блок 84 переупорядковування може, на основі цих елементів синтаксису переупорядковування, переупорядковувати інтерпольовані сигнали 49' nFG і приведені вектори 55_k переднього плану $V[k]$ для генерації переупорядкованих сигналів 49" nFG і переупорядкованих векторів $55_k'$ переднього плану $V[k]$. Блок 84 переупорядковування може виводити переупорядковані сигнали 49" nFG на блок 78 формування переднього плану і переупорядковані вектори $55_k'$ переднього плану $V[k]$ на блок 76 просторово-часової інтерполяції.

[415] Блок 76 просторово-часової інтерполяції може діяти аналогічно описаному вище відносно блока 50 просторово-часової інтерполяції. Блок 76 просторово-часової інтерполяції може приймати переупорядковані вектори $55_k'$ переднього плану $V[k]$ і здійснювати просторово-часову інтерполяцію відносно переупорядкованих векторів $55_k'$ переднього плану $V[k]$ і переупорядкованих векторів $55_{k-1}'$ переднього плану $V[k-1]$ для генерації інтерпольованих векторів $55_k''$ переднього плану $V[k]$. Блок 76 просторово-часової інтерполяції може пересилати інтерпольовані вектори $55_k''$ переднього плану $V[k]$ на блок 78 формування переднього плану.

[416] Блок 78 формування переднього плану може представляти блок, виконаний з можливістю здійснення матричного множення відносно інтерпольованих векторів $55_k''$ переднього плану $V[k]$ і переупорядкованих сигналів 49" nFG для генерації коефіцієнтів 65 НОА переднього плану. Блок 78 формування переднього плану може здійснювати матричне множення переупорядкованих сигналів 49" nFG на інтерпольовані вектори $55_k''$ переднього плану $V[k]$.

[417] Блок 82 формування коефіцієнтів НОА може представляти блок, виконаний з можливістю додавання коефіцієнтів 65 НОА переднього плану до каналів 47' НОА оточення для одержання коефіцієнтів 11' НОА, де штрих указує, що ці коефіцієнти 11' НОА можуть бути аналогічні, але не дорівнюють коефіцієнтам 11 НОА. Різниця між коефіцієнтами 11 і 11' НОА можуть бути обумовлені втратами внаслідок передачі по втратних середовищах передачі, квантування або інших втратних операцій.

[418] Таким чином, методи можуть передбачати, що пристрій аудіодекодування, наприклад пристрій 24 аудіодекодування, визначає, з бітового потоку, квантовану інформацію направленості, кодований аудіооб'єкт переднього плану і кодовані коефіцієнти амбіофонії більш високого порядку оточення (НОА), причому квантована інформація направленості і кодований аудіооб'єкт переднього плану представляють коефіцієнти НОА переднього плану, що описують компонентів переднього плану звукового поля, і при цьому кодовані коефіцієнти НОА оточення описують оточуючий компонент звукового поля, деквантує квантовану інформацію направленості для генерації інформації направленості, здійснює просторово-часову

інтерполяцію відносно інформації направленості для генерації інтерпольованої інформації направленості, аудіодекодує кодований аудіооб'єкт переднього плану для генерації аудіооб'єкта переднього плану і кодовані коефіцієнти HOA оточення для генерації коефіцієнтів HOA оточення, визначає коефіцієнти HOA переднього плану як функцію інтерпольованої інформації направленості й аудіооб'єкта переднього плану і визначає коефіцієнти HOA як функцію коефіцієнтів HOA переднього плану і коефіцієнтів HOA оточення.

[419] Таким чином, різні аспекти методів можуть передбачати, що уніфікований пристрій 24 аудіодекодування переключається між двома різними схемами зняття стиснення. У ряді випадків, пристрій 24 аудіодекодування може бути виконаний з можливістю вибору однієї з множини схем зняття стиснення на основі указання, чи генеруються стиснена версія коефіцієнтів сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта і зняття стиснення стисненої версії коефіцієнтів сферичних гармонік з використанням вибраної однієї з множини схем зняття стиснення.

У цих і інших випадках, пристрій 24 аудіодекодування містить вбудований декодер.

[420] У ряді випадків, пристрій 24 аудіодекодування може бути виконаний з можливістю одержання указання, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

[421] У цих і інших випадках, пристрій 24 аудіодекодування виконаний з можливістю одержання указання з бітового потоку, де зберігається стиснена версія коефіцієнтів сферичних гармонік.

[422] Таким чином, різні аспекти методів можуть передбачати, що пристрій 24 аудіодекодування одержує вектори, що описують характерні і фонові компоненти звукового поля. У ряді випадків, пристрій 24 аудіодекодування може бути виконаний з можливістю визначення одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів сферичних гармонік.

[423] У цих і інших випадках, у пристрої 24 аудіодекодування перетворення включає розкладання на сингулярні значення, що генерує матрицю U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матрицю S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матрицю V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік.

[424] У цих і інших випадках, у пристрої 24 аудіодекодування один або більше перших векторів містять один або більше аудіокодованих векторів $U_{\text{DIST}} * S_{\text{DIST}}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S , і при цьому матриця U і матриця S генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік.

[425] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{\text{DIST}} * S_{\text{DIST}}$ для генерації аудіодекодованої версії одного або більше аудіокодованих векторів $U_{\text{DIST}} * S_{\text{DIST}}$.

[426] У цих і інших випадках, у пристрої 24 аудіодекодування один або більше перших векторів містять один або більше аудіокодованих векторів $U_{\text{DIST}} * S_{\text{DIST}}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , і при цьому матриця U , матриця S і матриця V генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік.

[427] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{\text{DIST}} * S_{\text{DIST}}$ для генерації аудіодекодованої версії одного або більше аудіокодованих векторів $U_{\text{DIST}} * S_{\text{DIST}}$.

[428] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з можливістю множення векторів $U_{\text{DIST}} * S_{\text{DIST}}$ на вектори V_{DIST}^T для відновлення тих з множини сферичних гармонік, які представляють характерні компоненти звукового поля.

[429] У цих і інших випадках, у пристрої 24 аудіодекодування один або більше других векторів містять один або більше аудіокодованих векторів $U_{\text{BG}} * S_{\text{BG}} * V_{\text{BG}}^T$, які, до аудіокодування, генерувалися шляхом множення векторів U_{BG} , включених у матрицю U , на вектори S_{BG} , включені в матрицю S , і потім на вектори V_{BG}^T , включені в транспоновану матрицю V , і при

цьому матриця S і кожна з матриці U і матриці V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік.

[430] У цих і інших випадках, у пристрої 24 аудіодекодування один або більше других векторів містять один або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$, які, до аудіокодування, генерувалися шляхом множення векторів U_{BG} , включених у матрицю U , на вектори S_{BG} , включені в матрицю S , і потім на вектори V_{BG}^T , включені в транспоновану матрицю V , причому матриця S , матриця U і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому пристрій 24 аудіодекодування додатково виконаний з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$ для генерації одного або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$.

[431] У цих і інших випадках, у пристрої 24 аудіодекодування один або більше перших векторів містять один або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому пристрій 24 аудіодекодування додатково виконаний з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$ для генерації одного або більше векторів $U_{DIST} * S_{DIST}$ і множення векторів $U_{DIST} * S_{DIST}$ на вектори V_{DIST}^T для відновлення тих з множини коефіцієнтів сферичних гармонік, які описують характерні компоненти звукового поля, причому один або більше других векторів містять один або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$, які, до аудіокодування, генерувалися шляхом множення векторів U_{BG} , включених у матрицю U , на вектори S_{BG} , включені в матрицю S , і потім на вектори V_{BG}^T , включені в транспоновану матрицю V , і при цьому пристрій 24 аудіодекодування додатково виконаний з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$ для відновлення щонайменше частини з множини коефіцієнтів сферичних гармонік, які описують фонові компоненти звукового поля, і додавання множини коефіцієнтів сферичних гармонік, які описують характерні компоненти звукового поля, до щонайменше частини з множини коефіцієнтів сферичних гармонік, які описують фонові компоненти звукового поля, для генерації реконструйованої версії з множини коефіцієнтів сферичних гармонік.

[432] У цих і інших випадках, у пристрої 24 аудіодекодування один або більше перших векторів містять один або більше векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому пристрій аудіодекодування 20 додатково виконаний з можливістю одержання значення D , яке вказує кількість векторів, що підлягають витягання з бітового потоку, для формування одного або більше векторів $U_{DIST} * S_{DIST}$ і одного або більше векторів V_{DIST}^T .

[433] У цих і інших випадках, у пристрої 24 аудіодекодування один або більше перших векторів містять один або більше векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому пристрій 24 аудіодекодування додатково виконаний з можливістю одержання значення D на основі кадрів аудіо, що вказує кількість векторів, що підлягають витягання з бітового потоку, для формування одного або більше векторів $U_{DIST} * S_{DIST}$ і одного або більше векторів V_{DIST}^T .

[434] У цих і інших випадках, у пристрої 24 аудіодекодування перетворення включає аналіз головних компонентів для ідентифікації характерних компонентів звукового поля і фонових компонентів звукового поля.

[435] Різні аспекти методів, описаних у цьому винаході, також можуть передбачати, що пристрій 24 аудіокодування здійснює інтерполяцію відносно розкладених версій коефіцієнтів НОА. У ряді випадків, пристрій 24 аудіодекодування може бути виконаний з можливістю одержання розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, щонайменше частково, шляхом здійснення інтерполяції відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік.

[436] У цих і інших випадках, перше розкладання містить першу матрицю V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

[437] У цих і інших прикладах, друге розкладання містить другу матрицю V , яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

5 [438] У цих і інших випадках, перше розкладання містить першу матрицю V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік, і друге розкладання містить другу матрицю V , яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[439] У цих і інших випадках, відрізок часу містить підкадр кадру аудіо.

10 [440] У цих і інших випадках, відрізок часу містить часову вибірку кадру аудіо.

[441] У цих і інших випадках, пристрій 24 аудіодекодування виконаний з можливістю одержання інтерпольованого розкладання першого розкладання і другого розкладання для коефіцієнта сферичної гармоніки з першої множини коефіцієнтів сферичних гармонік.

15 [442] У цих і інших випадках, пристрій 24 аудіодекодування виконаний з можливістю одержання інтерпольованих розкладань першого розкладання для першої частини першої множини коефіцієнтів сферичних гармонік, включених у перший кадр, і другого розкладання для другої частини другої множини коефіцієнтів сферичних гармонік, включених у другий кадр, і пристрій 24 аудіодекодування додатково виконаний з можливістю застосування інтерпольованих розкладань до першого часового компонента першої частини першої множини коефіцієнтів сферичних гармонік, включених у перший кадр, для генерації першого штучного часового компонента першої множини коефіцієнтів сферичних гармонік, і застосування відповідних інтерпольованих розкладань до другого часового компонента другої частини другої множини коефіцієнтів сферичних гармонік, включених у другий кадр, для генерації другого штучного часового компонента другої множини коефіцієнтів сферичних гармонік, включених.

25 [443] У цих і інших випадках, перший часовий компонент генерується шляхом здійснення синтезу на векторній основі відносно першої множини коефіцієнтів сферичних гармонік.

[444] У цих і інших випадках, другий часовий компонент генерується шляхом здійснення синтезу на векторній основі відносно другої множини коефіцієнтів сферичних гармонік.

30 [445] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з можливістю прийому першого штучного часового компонента і другого штучного часового компонента, обчислення інтерпольованих розкладань першого розкладання для першої частини першої множини коефіцієнтів сферичних гармонік і другого розкладання для другої частини другої множини коефіцієнтів сферичних гармонік і застосування обернених інтерпольованих розкладань до першого штучного часового компонента для відновлення першого часового компонента і до другого штучного часового компонента для відновлення другого часового компонента.

[446] У цих і інших випадках, пристрій 24 аудіодекодування виконаний з можливістю інтерполяції першого просторового компонента першої множини коефіцієнтів сферичних гармонік і другого просторового компонента другої множини коефіцієнтів сферичних гармонік.

40 [447] У цих і інших випадках, перший просторовий компонент містить першу матрицю U , яка представляє ліві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

[448] У цих і інших випадках, другий просторовий компонент містить другу матрицю U , яка представляє ліві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

45 [449] У цих і інших випадках, перший просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для першої множини коефіцієнтів сферичних гармонік і другий просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для другої множини коефіцієнтів сферичних гармонік.

[450] У цих і інших випадках, перший просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для першої множини коефіцієнтів сферичних гармонік і другий просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для другої множини коефіцієнтів сферичних гармонік, і пристрій 24 аудіодекодування виконаний з можливістю інтерполяції останніх N елементів першого просторового компонента і перших N елементів другого просторового компонента.

55 [451] У цих і інших випадках, друга множина коефіцієнтів сферичних гармонік іде за першою множиною коефіцієнтів сферичних гармонік у часовій області.

[452] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з можливістю розкладання першої множини коефіцієнтів сферичних гармонік для генерації першого розкладання першої множини коефіцієнтів сферичних гармонік.

[453] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з можливістю розкладання другої множини коефіцієнтів сферичних гармонік для генерації другого розкладання другої множини коефіцієнтів сферичних гармонік.

[454] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з
5
можливістю здійснення розкладання на сингулярні значення відносно першої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори першої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення першої множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

[455] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з
10
можливістю здійснення розкладання на сингулярні значення відносно другої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори другої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення другої множини коефіцієнтів сферичних гармонік, і матриці V , яка
15
представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[456] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік представляють плоскохвильове представлення звукового поля.

[457] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік представляють один або більше змішаних один з одним монофонічних аудіооб'єктів.

[458] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік
20
містять відповідні перший і другий коефіцієнти сферичних гармонік, що представляють тривимірне звукове поле.

[459] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, що має порядок, який більше
25
одиниці.

[460] У цих і інших випадках, перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, що має порядок, який дорівнює чотирьом.

[461] У цих і інших випадках, інтерполяція є зваженою інтерполяцією першого розкладання і
30
другого розкладання, причому вагові коефіцієнти зваженої інтерполяції, застосовувані до першого розкладання, зворотно пропорційні часу, представленому векторами першого і другого розкладання, і при цьому вагові коефіцієнти зваженої інтерполяції, застосовувані до другого розкладання, пропорційні часу, представленому векторами першого і другого розкладання.

[462] У цих і інших випадках, розкладені інтерпольовані коефіцієнти сферичних гармонік
35
згладжують щонайменше один із просторових компонентів і часових компонентів першої множини коефіцієнтів сферичних гармонік і другої множини коефіцієнтів сферичних гармонік.

[463] У цих і інших випадках, пристрій 24 аудіодекодування виконаний з можливістю обчислення $Us[n] = HOA(n) * (V_vec[n]) - 1$ для одержання скаляра.

[464] У цих і інших випадках, інтерполяція включає лінійну інтерполяцію. У цих і інших
40
випадках, інтерполяція включає нелінійну інтерполяцію. У цих і інших випадках, інтерполяція включає косинусну інтерполяцію. У цих і інших випадках, інтерполяція включає зважену косинусну інтерполяцію. У цих і інших випадках, інтерполяція включає кубічну інтерполяцію. У цих і інших випадках, інтерполяція включає адаптивну сплайнову інтерполяцію. У цих і інших випадках, інтерполяція включає інтерполяцію мінімальної кривизни.

[465] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з
45
можливістю генерації бітового потоку, що включає в себе представлення розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, і указання типу інтерполяції.

[466] У цих і інших випадках, указання містить один або більше бітів, що відображаються в
50
тип інтерполяції.

[467] У цих і інших випадках, пристрій 24 аудіодекодування додатково виконаний з
можливістю одержання бітового потоку, що включає в себе представлення розкладених
інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу, і указання типу
інтерполяції.

[468] У цих і інших випадках, указання містить один або більше бітів, що відображаються в
55
тип інтерполяції.

[469] Різні аспекти методів, у ряді випадків, можуть додатково передбачати, що пристрій 24 аудіодекодування виконаний з можливістю одержання бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується
60
шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[470] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента.

5 [471] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.

10 [472] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.

[473] У цих і інших випадках, значення містить значення nbits.

15 [474] У цих і інших випадках, бітовий потік містить стиснену версію з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.

[475] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.

20 [476] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

25 [477] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення залишкового значення просторового компонента.

[478] У цих і інших випадках, пристрій містить пристрій аудіодекодування.

30 [479] Різні аспекти методів також можуть передбачати, що пристрій 24 аудіодекодування ідентифікує кодову книгу Хаффмана для використання при знятті стиснення стисненої версії просторового компонента з множини стиснених просторових компонентів на основі порядку стисненої версії просторового компонента відносно стиснених просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

35 [480] У цих і інших випадках, пристрій 24 аудіодекодування виконаний з можливістю одержання бітового потоку, що містить стиснену версію просторового компонента звукового поля, і зняття стиснення стисненої версії просторового компонента з використанням, щонайменше частково, ідентифікованої кодової книги Хаффмана для одержання просторового компонента.

40 [481] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента, і пристрій 24 аудіодекодування виконаний з можливістю зняття стиснення стисненої версії просторового компонента, щонайменше частково, на основі режиму прогнозування для одержання просторового компонента.

45 [482] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента, і пристрій 24 аудіодекодування виконаний з можливістю зняття стиснення стисненої версії просторового компонента, щонайменше частково, на основі інформації таблиці Хаффмана.

50 [483] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента, і пристрій 24 аудіодекодування виконаний з можливістю зняття стиснення стисненої версії просторового компонента, щонайменше частково, на основі значення.

55 [484] У цих і інших випадках, значення містить значення nbits.

[485] У цих і інших випадках, бітовий потік містить стиснену версію з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів, і пристрій 24 аудіодекодування виконаний з можливістю

зняття стиснення множини стисненої версії просторового компонента, щонайменше частково, на основі значення.

[486] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент, і пристрій 24 аудіодекодування виконаний з можливістю зняття стиснення стисненої версії просторового компонента, щонайменше частково, на основі коду Хаффмана.

[487] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення, і пристрій 24 аудіодекодування виконаний з можливістю зняття стиснення стисненої версії просторового компонента, щонайменше частково, на основі знакового біта.

[488] У цих і інших випадках, стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення залишкового значення просторового компонента, і пристрій 24 аудіодекодування виконаний з можливістю зняття стиснення стисненої версії просторового компонента, щонайменше частково, на основі коду Хаффмана, включеного в ідентифіковану кодову книгу Хаффмана.

[489] У кожному з різних вищеописаних випадків, варто розуміти, що пристрій 24 аудіодекодування може здійснювати спосіб або іншим чином містити засіб для здійснення кожного етапу способу, для якого пристрій 24 аудіодекодування виконаний з можливістю здійснення в ряді випадків, цей засіб може містити один або більше процесорів. У ряді випадків, один або більше процесорів можуть представляти процесор спеціального призначення, що виконує інструкції, які зберігаються на довгостроковому зчитуваному комп'ютером носії даних. Іншими словами, різні аспекти методів у кожному з наборів прикладів кодування можуть передбачати довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати спосіб, для виконання якого призначений пристрій 24 аудіодекодування.

[490] На фіг. 6 показана блок-схема операцій, що демонструє роботу блока аналізу контенту пристрою аудіокодування, наприклад блока 26 аналізу контенту, показаного в прикладі, наведеному на фіг. 4, при здійсненні різних аспектів методів, описаних у цьому винаході.

[491] Блок 26 аналізу контенту може, при визначенні, чи генеруються коефіцієнти 11 НОА, що представляють звукове поле, із синтетичного аудіооб'єкта, одержувати кадровані коефіцієнти НОА (93), що можуть мати розмір 25 на 1024, для представлення четвертого порядку (тобто $N=4$). Після одержання кадрованих коефіцієнтів НОА (які також можуть бути позначені тут як матриця 11 кадрованих SHC, і наступні матриці кадрованих SHC можуть бути позначені як матриці 27B, 27C і т. д. кадрованих SHC), блок 26 аналізу контенту може виключити перший вектор кадрованих коефіцієнтів 11 НОА для генерації приведених кадрованих коефіцієнтів НОА (94).

[492] Потім блок 26 аналізу контенту може прогнозувати перший ненульовий вектор приведених кадрованих коефіцієнтів НОА з векторів, що залишилися, приведених кадрованих коефіцієнтів НОА (95). Після прогнозування першого ненульового вектора, блок 26 аналізу контенту може одержувати помилку на основі прогнозованого першого ненульового вектора і фактичного ненульового вектора (96). Одержавши помилку, блок 26 аналізу контенту може обчислювати відношення на основі енергії фактичного першого ненульового вектора і помилки (97). Потім блок 26 аналізу контенту може порівнювати це відношення з порогом (98). Коли відношення не перевищує поріг ("Ні" 98), блок 26 аналізу контенту може визначити, що матриця 11 кадрованих SHC генерується з запису, і вказати в бітовому потоці, що відповідне кодоване представлення матриці 11 SHC згенероване з запису (100, 101). Коли відношення перевищує поріг ("Так" 98), блок 26 аналізу контенту може визначити, що матриця 11 кадрованих SHC генерується із синтетичного аудіооб'єкта, і вказати в бітовому потоці, що відповідне кодоване представлення матриці 11 SHC згенероване із синтетичного аудіооб'єкта (102, 103). У ряді випадків, коли матриця 11 кадрованих SHC згенерована з запису, блок 26 аналізу контенту передає матрицю 11 кадрованих SHC на блок 27 синтезу на векторній основі (101). У ряді випадків, коли матриця 11 кадрованих SHC згенерована із синтетичного аудіооб'єкта, блок 26 аналізу контенту передає матрицю 11 кадрованих SHC на блок 28 синтезу на основі напрямку (104).

[493] На фіг. 7 показана блок-схема операцій, що демонструє роботу пристрою аудіокодування, наприклад пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4, при здійсненні різних аспектів методів синтезу на векторній основі, описаних у цьому винаході. Початково пристрій 20 аудіокодування приймає коефіцієнти 11 НОА (106). Пристрій 20

аудіокодування може викликати блок 30 LIT, який може застосовувати LIT відносно коефіцієнтів HOA для виведення перетворених коефіцієнтів HOA (наприклад, у випадку SVD, перетворені коефіцієнти HOA можуть містити вектори 33 US[k] і вектори 35 V[k]) (107).

5 [494] Потім пристрій 20 аудіокодування може викликати блок 32 обчислення параметрів для здійснення вищеописаного аналізу відносно будь-якої комбінації векторів 33 US[k], векторів 33 US[k-1], векторів 35 V[k] і/або V[k-1] для ідентифікації різних параметрів вищеописаним чином. Таким чином, блок 32 обчислення параметрів може визначати щонайменше один параметр на основі аналізу перетворених коефіцієнтів 33/35 HOA (108).

10 [495] Потім пристрій 20 аудіокодування може викликати блок 34 переупорядкування, який може переупорядковувати перетворені коефіцієнти HOA (які, знову ж, у контексті SVD, можуть означати вектори 33 US[k] і вектори 35 V[k]) на основі параметра для генерації переупорядкованих перетворених коефіцієнтів 33'/35' HOA (або, іншими словами, векторів 33' US[k] і векторів 35' V[k]), як описано вище (109). Пристрій 20 аудіокодування може, при виконанні будь-якої з вищенаведених операцій або наступних операцій, також викликати блок 15 44 аналізу звукового поля. Блок 44 аналізу звукового поля може, як описано вище, здійснювати аналіз звукового поля відносно коефіцієнтів 11 HOA і/або перетворених коефіцієнтів 33/35 HOA для визначення сумарної кількості каналів 45 переднього плану (nFG), порядку фонового звукового поля (N_{BG}) і кількості (nBGa) і індексів (i) додаткових каналів HOA BG для відправлення (які можуть спільно позначатися як інформація 43 фонового каналу в прикладі, 20 наведеному на фіг. 4) (109).

[496] Пристрій 20 аудіокодування може також викликати блок 48 вибору фону. Блок 48 вибору фону може визначати коефіцієнти 47 HOA фону або оточення на основі інформації 43 фонового каналу (110). Пристрій 20 аудіокодування може додатково викликати блок 36 вибору переднього плану, який може вибирати ті з переупорядкованих векторів 33' US[k] і 25 переупорядкованих векторів 35' V[k], які представляють компоненти переднього плану або характерні компоненти звукового поля на основі nFG 45 (який може представляти один або більше індексів, що ідентифікують ці вектори переднього плану) (112).

[497] Пристрій 20 аудіокодування може викликати блок 38 компенсації енергії. Блок 38 компенсації енергії може здійснювати компенсацію енергії відносно коефіцієнтів 47 HOA оточення для компенсації втрати енергії внаслідок видалення різних каналів HOA блоком 48 вибору фону (114) і, таким чином, генерувати коефіцієнти 47' HOA оточення з компенсацією енергії.

[498] Потім пристрій 20 аудіокодування також може викликати блок 50 просторово-часової інтерполяції. Блок 50 просторово-часової інтерполяції може здійснювати просторово-часову 35 інтерполяцію відносно переупорядкованих перетворених коефіцієнтів 33'/35' HOA для одержання інтерпольованих сигналів 49' переднього плану (які також можуть іменуватися "інтерпольовані сигнали 49' nFG") і інформації 53, що залишилися, направленості переднього плану (яка також може іменуватися "вектори 53 V[k]") (116). Потім пристрій 20 аудіокодування може викликати блок 46 приведення коефіцієнтів. Блок 46 приведення коефіцієнтів може 40 здійснювати приведення коефіцієнтів відносно векторів 53, що залишилися, переднього плану V[k] на основі інформації 43 фонового каналу для одержання скороченої інформації 55 направленості переднього плану (яка також може іменуватися приведеними векторами 55 переднього плану V[k]) (118).

[499] Потім пристрій 20 аудіокодування може викликати блок 52 квантування для стиснення, 45 вищеописаним чином, приведених векторів 55 переднього плану V[k] і генерації кодованих векторів 57 переднього плану V[k] (120).

[500] Пристрій 20 аудіокодування може також викликати блок 40 психоакустичного аудіокодера. Блок 40 психоакустичного аудіокодера може психоакустично кодувати кожен вектор коефіцієнтів 47' HOA оточення з компенсацією енергії й інтерпольованих сигналів 49' 50 nFG для генерації кодованих коефіцієнтів 59 HOA оточення і кодованих сигналів 61 nFG. Потім пристрій аудіокодування може викликати блок 42 генерації бітового потоку. Блок 42 генерації бітового потоку може генерувати бітовий потік 21 на основі кодової інформації 57 направленості переднього плану, кодованих коефіцієнтів 59 HOA оточення, кодованих сигналів 61 nFG і інформації 43 фонового каналу.

55 [501] На фіг. 8 показана блок-схема операцій, що демонструє роботу пристрою аудіодекодування, наприклад пристрою 24 аудіодекодування, показаного на фіг. 5, при здійсненні різних аспектів методів, описаних у цьому винаході. Спочатку, пристрій 24 аудіодекодування може приймати бітовий потік 21 (130). Після прийому бітового потоку, пристрій 24 аудіодекодування може викликати блок 72 витягання. Передбачаючи з метою 60 розгляду, що бітовий потік 21 указує, що реконструкція на векторній основі підлягає здійсненню,

пристрій 72 витягання може виділяти цей бітовий потік для витягання вищезгаданої інформації, передаючи цю інформацію на блок 92 реконструкції на векторній основі.

[502] Іншими словами, блок 72 витягання може витягати кодовану інформацію 57 направленості переднього плану (яка, знову ж, також може іменуватися кодованими векторами 57 переднього плану $V[k]$), кодовані коефіцієнти 59 HOA оточення і кодовані сигнали переднього плану (які також можуть іменуватися кодованими сигналами 59 nFG переднього плану або кодованими аудіооб'єктами 59 переднього плану) з бітового потоку 21 вищеописаним чином (132).

[503] Пристрій 24 аудіодекодування може додатково викликати блок 74 квантування. Блок 74 квантування може ентропійно декодувати і деквантувати кодовану інформацію 57 направленості переднього плану для одержання скороченої інформації 55_k направленості переднього плану (136). Пристрій 24 аудіодекодування також може викликати блок 80 психоакустичного декодування. Блок психоакустичного аудіокодування 80 може декодувати кодовані коефіцієнти 59 HOA оточення і кодовані сигнали 61 переднього плану для одержання коефіцієнтів 47' HOA оточення з компенсацією енергії й інтерпольованих сигналів 49' переднього плану (138). Блок 80 психоакустичного декодування може передавати коефіцієнти 47' HOA оточення з компенсацією енергії на блок 82 формування коефіцієнтів HOA і сигнали 49' nFG на блок 84 переупорядковування.

[504] Блок 84 переупорядковування може приймати елементи синтаксису, що указують вихідний порядок компонентів переднього плану коефіцієнтів 11 HOA. Блок 84 переупорядковування може, на основі цих елементів синтаксису переупорядковування, переупорядковувати інтерпольовані сигнали 49' nFG і приведені вектори 55_k переднього плану $V[k]$ для генерації переупорядкованих сигналів 49" nFG і переупорядкованих векторів 55_k' переднього плану $V[k]$ (140). Блок 84 переупорядковування може виводити переупорядковані сигнали 49" nFG на блок 78 формування переднього плану і переупорядковані вектори 55_k' переднього плану $V[k]$ на блок 76 просторово-часової інтерполяції.

[505] Потім пристрій 24 аудіодекодування може викликати блок 76 просторово-часової інтерполяції. Блок 76 просторово-часової інтерполяції може приймати переупорядковану інформацію 55_k' направленості переднього плану і здійснювати просторово-часову інтерполяцію відносно скороченої інформації 55_k/55_{k-1} направленості переднього плану для генерації інтерпольованої інформації 55_k" направленості переднього плану (142). Блок 76 просторово-часової інтерполяції може пересилати інтерпольовані вектори 55_k" переднього плану $V[k]$ на блок 718 формування переднього плану.

[506] Пристрій 24 аудіодекодування може викликати блок 78 формування переднього плану. Блок 78 формування переднього плану може здійснювати матричне множення інтерпольованих сигналів 49" переднього плану на інтерпольовану інформацію 55_k" направленості переднього плану для одержання коефіцієнтів 65 HOA переднього плану (144). Пристрій 24 аудіодекодування також може викликати блок 82 формування коефіцієнтів HOA. Блок 82 формування коефіцієнтів HOA може додавати коефіцієнти 65 HOA переднього плану до каналів 47' HOA оточення для одержання коефіцієнтів 11' HOA (146).

[507] На фіг. 9A-9L показані блок-схеми, що більш детально демонструють різні аспекти пристрою 20 аудіокодування прикладу, наведеного на фіг. 4. На фіг. 9A більш детально показана блок-схема, що демонструє блок 30 LIT пристрою 20 аудіокодування. Як показано в прикладі, наведеному на фіг. 9A, блок 30 LIT може включати в себе множинні різні лінійні оборотні перетворення 200-200N. Блок 30 LIT може включати в себе, для забезпечення декількох прикладів, перетворення 200A розкладання на сингулярні значення (SVD) ("SVD 200A"), перетворення 200B аналізу основних компонентів (PCA) ("PCA 200B"), перетворення 200C Крунена-Лева (KLT) ("KLT 200C"), швидке перетворення Фур'є (FFT) 200D ("FFT 200D") і дискретне косинусне перетворення (DCT) 200N ("DCT 200N"). Блок 30 LIT може викликати будь-яке з цих лінійних оборотних перетворень 200 для застосування відповідного перетворення відносно коефіцієнтів 11 HOA і генерувати відповідні перетворені коефіцієнти 33/35 HOA.

[508] Хоча вони описані як здійснювані безпосередньо відносно коефіцієнтів 11 HOA, блок 30 LIT може застосовувати лінійні оборотні перетворення 200 до похідних коефіцієнтів 11 HOA. Наприклад, блок 30 LIT може застосовувати SVD 200 відносно матриці спектральної щільності потужності, виведеної з коефіцієнтів 11 HOA. Матриця спектральної щільності потужності може бути позначена як PSD і одержана шляхом матричного множення транспонованої $hoaFrame$ на $hoaFrame$, що викладено в нижченаведеному псевдокоді. Позначення $hoaFrame$ означає кадр коефіцієнтів 11 HOA.

[509] Блок 30 LIT може, після застосування SVD 200 (svd) до PSD, одержувати матрицю $S[k]^2$ ($S_squared$) і матрицю $V[k]$. Матриця $S[k]^2$ може позначати квадрат матриці $S[k]$, у зв'язку з чим

блок 30 LIT (або, альтернативно, блок 200 SVD у порядку одного прикладу) може застосовувати операцію витягання квадратного кореня з матриці $S[k]^2$ для одержання матриці $S[k]$. Блок 200 SVD може, у ряді випадків, здійснювати квантування відносно матриці $V[k]$ для одержання квантованої матриці $V[k]$ (яка може бути позначена як матриця $V[k]'$). Блок 30 LIT може

5 одержувати матрицю $U[k]$, спочатку множачи матрицю $S[k]$ на квантовану матрицю $V[k]'$ для одержання матриці $SV[k]'$. Потім блок 30 LIT може одержувати псевдообернену (pinv) матрицю $SV[k]'$ і потім множити коефіцієнти 11 HOA на псевдообернену матрицю $SV[k]'$ для одержання матриці $U[k]$. Вищенаведений опис може бути представлений наступним псевдокодом:

```
PSD=hoaFrame*hoaFrame;
```

```
10 [V, S_squared]=svd(PSD, 'econ');
```

```
S=sqrt(S_squared);
```

```
U=hoaFrame*pinv(S*V');
```

[510] Шляхом здійснення SVD відносно спектральної щільності потужності (PSD) коефіцієнтів HOA замість самих коефіцієнтів, блок 30 LIT може потенційно знижувати

15 обчислювальну складність здійснення SVD відносно одного або більше з циклів процесора і простору збереження, досягаючи при цьому такої ж ефективності аудіокодування джерела, як якби SVD застосовувалося безпосередньо до коефіцієнтів HOA. Таким чином, вищеописане SVD типу PSD може потенційно вимагати менше обчислювальної потужності, оскільки SVD здійснюється на матриці F^*F (де F - кількість коефіцієнтів HOA). У порівнянні з матрицею M^*F , де M - довжина кадру, тобто 1024 або більш вибірок. Тепер складність SVD, завдяки застосуванню до PSD замість коефіцієнтів 11 HOA, може бути близько $O(L^3)$ у порівнянні з $O(M^*L^2)$ у випадку застосування до коефіцієнтів 11 HOA (де $O(*)$ позначає символ великого O обчислювальної складності, звичайно використовуваний у комп'ютерній техніці).

[511] На фіг. 9B показана блок-схема, що більш детально демонструє блок 32 обчислення

25 параметрів пристрою 20 аудіокодування. Блок 32 обчислення параметрів може включати в себе блок 202 аналізу енергії і блок 204 крос-кореляції. Блок 202 аналізу енергії може здійснювати вищеописаний аналіз енергії відносно одного або більше з векторів 33 $US[k]$ і векторів 35 $V[k]$ для генерації одного або більше з параметра кореляції (R), параметрів властивостей

30 направленості (θ , ϕ , γ) і властивості енергії (e) для одного або більше з поточного кадру (k) і попереднього кадру ($k-1$). Аналогічно, блок 204 крос-кореляції може здійснювати вищеописану крос-кореляцію відносно одного або більше з векторів 33 $US[k]$ і векторів 35 $V[k]$ для генерації одного або більше з параметра кореляції (R), параметрів властивостей направленості (θ , ϕ , γ) і властивості енергії (e) для одного або більше з поточного кадру (k) і попереднього кадру ($k-1$). Блок 32 обчислення параметрів може виводити параметри 37 поточного кадру і параметри 39

35 попереднього кадру.

[512] На фіг. 9C показана блок-схема, що більш детально демонструє блок 34

переупорядковування пристрою 20 аудіокодування. Блок 34 переупорядковування включає в себе блок 206 оцінювання параметрів і блок 208 переупорядковування векторів. Блок 206

40 оцінювання параметрів представляє блок, виконаний з можливістю оцінювання параметрів 39 попереднього кадру і параметрів 37 поточного кадру вищеописаним чином для генерації індексів 205 переупорядковування. Індеси 205 переупорядковування включають в себе індекси, що ідентифікують, як потрібно переупорядковувати вектори з векторів 33 $US[k]$ і з векторів 35 $V[k]$ (наприклад, парами індексів, де перший індекс пари ідентифікує індекс поточного положення вектора і другий індекс пари ідентифікує положення переупорядкованого вектора).

45 Блок 208 переупорядковування векторів представляє блок, виконаний з можливістю преупорядковування векторів 33 $US[k]$ і векторів 35 $V[k]$ відповідно до індексів 205 переупорядковування. Блок 34 переупорядковування може виводити переупорядковані вектори 33' $US[k]$ і переупорядковані вектори 35' $V[k]$, також передаючи індекси 205 переупорядковування як один або більше елементів синтаксису на блок 42 генерації бітового

50 потоку.

[513] На фіг. 9D показана блок-схема, що більш детально демонструє блок 44 аналізу

звукowego поля пристрою 20 аудіокодування. Як показано в прикладі, наведеному на фіг. 9D, блок 44 аналізу звукowego поля може включати в себе блок 210A аналізу сингулярних значень, блок 210B аналізу енергії, блок 210C просторового аналізу, блок 210D аналізу просторового

55 маскувння, блок 210E аналізу дифузії і блок 210F аналізу направленості. Блок 210A аналізу сингулярних значень може представляти блок, виконаний з можливістю аналізу нахилу кривої, створеної зменшуваними діагональними значеннями векторів S (утворюючих частину векторів 33 $US[k]$), де великі сингулярні значення представляють звуки переднього плану або характерні звуки і низькі сингулярні значення представляють фонові компоненти звукowego поля, як

описано вище. Блок 210B аналізу енергії може представляти блок, виконаний з можливістю визначення енергії векторів $35 V[k]$ для кожного вектора.

[514] Блок 210C просторового аналізу може представляти блок, виконаний з можливістю здійснення вищеописаного просторового аналізу енергії шляхом перетворення коефіцієнтів 11 HOA у просторову область і ідентифікації областей високої енергії, що представляють направлені компоненти звукового поля, які потрібно зберегти. Блок 210D аналізу просторового маскування може представляти блок, виконаний з можливістю здійснення аналізу просторового маскування аналогічно блоку просторового аналізу енергії, за винятком того, що блок 210D аналізу просторового маскування може ідентифікувати просторові області, які маскуються просторово близькими звуками більш високої енергії. Блок 210E аналізу дифузії може представляти блок, виконаний з можливістю здійснення вищеописаного аналізу дифузії відносно коефіцієнтів 11 HOA для ідентифікації областей дифузної енергії, що можуть представляти фонові компоненти звукового поля. Блок 210F аналізу направленості може представляти блок, виконаний з можливістю здійснення вищезгаданого аналізу направленості, що передбачає обчислення векторів $VS[k]$ і піднесення до квадрата кожного елемента кожного з цих векторів $VS[k]$ і їх підсумовування для ідентифікації показника направленості. Блок 210F аналізу направленості може забезпечувати цей показник направленості для кожного з векторів $VS[k]$ на блок 212 ідентифікації (ID) фону/переднього плану (BG/FG).

[515] Блок 44 аналізу звукового поля також може включати в себе блок 212 ID BG/FG, що може представляти блок, виконаний з можливістю визначення сумарної кількості каналів 45 переднього плану (nFG), порядку фонового звукового поля (N_{BG}) і кількості ($nBGa$) і індексів (i) додаткових каналів HOA BG для відправлення (які можуть спільно позначатися як інформація 43 фонового каналу в прикладі, наведеному на фіг. 4) на основі будь-якої комбінації аналізу, що виводиться будь-якою комбінацією блоків 210-210F аналізу. Блок 212 ID BG/FG може визначати nFG 45 і інформацію 43 фонового каналу для досягнення цільової бітової швидкості 41.

[516] На фіг. 9E показана блок-схема, що більш детально демонструє блок 36 вибору переднього плану пристрою 20 аудіокодування. Блок 36 вибору переднього плану включає в себе блок 214 виділення вектора, який може виділяти або інакше витягати вектори 49 переднього плану $US[k]$ і вектори 51_k переднього плану $V[k]$, ідентифіковані елементом 45 синтаксису nFG , з переупорядкованих векторів $33' US[k]$ і переупорядкованих векторів $35' V[k]$. Блок 214 виділення вектора може виділяти різні вектори, що представляють компоненти переднього плану звукового поля, ідентифіковані блоком 44 аналізу звукового поля і зазначені елементом 45 синтаксису nFG (які також можуть іменуватися інформацією 45 каналів переднього плану). Як показано в прикладі, наведеному на фіг. 9E, блок 214 виділення вектора може вибирати, у ряді випадків, непослідовні вектори серед векторів 49 переднього плану $US[k]$ і векторів 51_k переднього плану $V[k]$ для представлення компонентів переднього плану звукового поля. Крім того, блок 214 виділення вектора може вибирати, у ряді випадків, одні і ті ж вектори (відносно позиції) серед векторів 49 переднього плану $US[k]$ і векторів 51_k переднього плану $V[k]$ для представлення компонентів переднього плану звукового поля.

[517] На фіг. 9F показана блок-схема, що більш детально демонструє блок 48 вибору фону пристрою 20 аудіокодування. Блок 48 вибору фону може визначати коефіцієнти 47 HOA фону або оточення на основі інформації фонового каналу (наприклад, фонового звукового поля (N_{BG}) і кількості ($nBGa$) і індексів (i) додаткових каналів HOA BG для відправлення). Наприклад, коли N_{BG} дорівнює одиниці, блок 48 вибору фону може вибирати коефіцієнти 11 HOA для кожної вибірки кадру аудіо, що має порядок, який менше або дорівнює одиниці. Потім блок 48 вибору фону може, у цьому прикладі, вибрати коефіцієнти 11 HOA, що мають індекс, ідентифікований одним з індексів (i) як додаткові коефіцієнти HOA BG, де $nBGa$ надходить на блок 42 генерації бітового потоку для указання в бітовому потоці 21, щоб пристрій аудіодекодування, наприклад пристрій 24 аудіодекодування, показаний в прикладі, наведеному на фіг. 5, виділяв коефіцієнти 47 HOA BG з бітового потоку 21. Потім блок 48 вибору фону може виводити коефіцієнти 47 HOA оточення на блок 38 компенсації енергії. Коефіцієнти 47 HOA оточення можуть мати розміри $D: M \times [(N_{BG}+1)^2 + nBGa]$.

[518] На фіг. 9G показана блок-схема, що більш детально демонструє блок 38 компенсації енергії пристрою 20 аудіокодування. Блок 38 компенсації енергії може представляти блок, виконаний з можливістю здійснення компенсації енергії відносно коефіцієнтів 47 HOA оточення для компенсації втрати енергії внаслідок видалення різних каналів HOA блоком 48 вибору фону. Блок 38 компенсації енергії може включати в себе блок 218 визначення енергії, блок 220 аналізу енергії і блок 222 посилення енергії.

[519] Блок 218 визначення енергії може представляти блок, виконаний з можливістю ідентифікації RMS для кожного рядка і/або стовпця однієї або більше з переупорядкованої

матриці 33' US[k] і переупорядкованої матриці 35' V[k]. Блок 38 визначення енергії також може ідентифікувати RMS для кожного рядка і/або стовпця одного або більше з вибраних каналів переднього плану, що може включати в себе сигнали 49 nFG і вектори 51_k переднього плану V[k], і коефіцієнти 47 HOA оточення зниженого порядку. RMS для кожного рядка і/або стовпця однієї або більше з переупорядкованої матриці 33' US[k] і переупорядкованої матриці 35' V[k] може зберігатися у вектор, позначений RMS_{FULL}, тоді як RMS для кожного рядка і/або стовпця одного або більше із сигналів 49 nFG, векторів 51_k переднього плану V[k] і коефіцієнтів 47 HOA оточення зниженого порядку можуть зберігатися у вектор, позначений RMS_{REDUCED}.

[520] У деяких прикладах, для визначення кожного RMS відповідних рядків і/або стовпців однієї або більше з переупорядкованої матриці 33' US[k], переупорядкованої матриці 35' V[k], сигналів 49 nFG, векторів 51_k переднього плану V[k] і коефіцієнтів 47 HOA оточення зниженого порядку, блок 218 визначення енергії може спочатку застосовувати рендеризатор опорних коефіцієнтів сферичних гармонік (SHC) до стовпців. Застосування рендеризатора опорних SHC блоком 218 визначення енергії дозволяє визначати RMS в області SHC для визначення енергії загального звукового поля, описаного кожним рядком і/або стовпцем кадру, представленого рядками і/або стовпцями однієї або більше з переупорядкованої матриці 33' US[k], переупорядкованої матриці 35' V[k], сигналів 49 nFG, векторів 51_k переднього плану V[k] і коефіцієнтів 47 HOA оточення зниженого порядку. Блок 38 визначення енергії може передавати ці вектори RMS_{FULL} і RMS_{REDUCED} на блок 220 аналізу енергії.

[521] Блок 220 аналізу енергії може представляти блок, виконаний з можливістю обчислення вектора Z значень посилення, згідно з наступним рівнянням: $Z = RMS_{FULL}/RMS_{REDUCED}$.

Блок 220 аналізу енергії потім може передавати цей вектор Z значень посилення на блок 222 посилення енергії. Блок 222 посилення енергії може представляти блок, виконаний з можливістю застосування цього вектора Z значень посилення або різних його ділянок до одного або більше із сигналів 49 nFG, векторів 51_k переднього плану V[k] і коефіцієнтів 47 HOA оточення зниженого порядку. У ряді випадків, вектор Z значень посилення застосовується тільки до коефіцієнтів 47 HOA оточення зниженого порядку відповідно до наступного рівняння $HOA_{BG-RED}' = HOA_{BG-RED} Z^T$, де HOA_{BG-RED} позначає коефіцієнти 47 HOA оточення зниженого порядку, HOA_{BG-RED}' позначає приведені коефіцієнти 47' HOA оточення з компенсацією енергії, і Z^T позначає транспонування вектора Z.

[522] На фіг. 9H показана блок-схема, що демонструє, більш детально, блок 50 просторово-часової інтерполяції пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. Блок 50 просторово-часової інтерполяції може представляти блок, виконаний з можливістю прийому векторів 51_k переднього плану V[k] для k-го кадру і векторів 51_{k-1} переднього плану V[k-1] для попереднього кадру (тому позначених k-1) і здійснення просторово-часової інтерполяції для генерації інтерпольованих векторів переднього плану V[k]. Блок 50 просторово-часової інтерполяції може включати в себе блок 224 інтерполяції V і блок 226 адаптації переднього плану.

[523] Блок 224 інтерполяції V може вибирати частину поточних векторів 51_k переднього плану V[k] для інтерполяції на основі ділянок, що залишилися, поточних векторів 51_k переднього плану V[k] і попередніх векторів 51_{k-1} переднього плану V[k-1]. Блок 224 інтерполяції V може вибирати частину як один або більше з вищезгаданих підкадрів або одну єдину невизначену частину, яка може змінюватися на основі кадрів. Блок 224 інтерполяції V може, у ряді випадків, вибрати одиничну частину 128 вибірок з 1024 вибірок поточних векторів 51_k переднього плану V[k] для інтерполяції. Потім блок 224 інтерполяції V може перетворювати кожний з векторів серед поточних векторів 51_k переднього плану V[k] і попередніх векторів 51_{k-1} переднього плану V[k-1] для поділу просторових карт шляхом проектування векторів на сферу (з використанням проекційної матриці, наприклад Т-подібної матриці). Потім блок 224 інтерполяції V може інтерпретувати вектори в V як форми на сфері. Для інтерполяції матриць V для частини 256 вибірок, блок 224 інтерполяції V потім може інтерпольовати ці просторові форми і потім перетворювати їх зворотню у вектори області сферичних гармонік за допомогою оберненої проекційної матриці. Методи цього винаходу, таким чином, можуть забезпечувати плавний перехід між матрицями V. Потім блок 224 інтерполяції V може генерувати вектори 53, що залишилися, V[k], які представляють вектори 51_k переднього плану V[k] після модифікації, для видалення інтерпольованої частини векторів 51_k переднього плану V[k]. Потім блок 224 інтерполяції V може передавати інтерпольовані вектори 51_k' переднього плану V[k] на блок 226 адаптації nFG.

[524] При виборі одиничної частини для інтерполяції, блок 224 інтерполяції V може генерувати елемент синтаксису, позначений CodedSpatialInterpolationTime 254, що ідентифікує

тривалість або, іншими словами, час інтерполяції (наприклад, відносно кількості вибірок). При виборі одиначної частини здійснення інтерполяції підкадру, блок 224 інтерполяції V також може генерувати інший елемент синтаксису, позначений SpatialInterpolationMethod 255, що може ідентифікувати тип здійснюваної інтерполяції (або, у ряді випадків, чи була здійснена інтерполяція). Блок 50 просторово-часової інтерполяції може виводити ці елементи 254 і 255 синтаксису на блок 42 генерації бітового потоку.

[525] Блок 226 адаптації nFG може представляти блок, виконаний з можливістю генерації адаптованих сигналів 49' nFG. Блок 226 адаптації nFG може генерувати адаптовані сигнали 49' nFG, спочатку одержуючи коефіцієнти НОА переднього плану шляхом множення сигналів 49 nFG на вектори 51_k переднього плану V[k]. Після одержання коефіцієнтів НОА переднього плану, блок 226 адаптації nFG може ділити коефіцієнти НОА переднього плану на інтерпольовані вектори 53 переднього плану V[k] для одержання адаптованих сигналів 49' nFG (які можуть іменуватися інтерпольованими сигналами 49' nFG, за умови, що ці сигнали виведені з інтерпольованих векторів 51_k переднього плану V[k]).

[526] На фіг. 9I показана блок-схема, що демонструє, більш детально, блок 46 приведення коефіцієнтів пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. Блок 46 приведення коефіцієнтів може представляти блок, виконаний з можливістю здійснення приведення коефіцієнтів відносно векторів 53, що залишилися, переднього плану V[k] на основі інформації 43 фонового каналу для виведення приведених векторів 55 переднього плану V[k] на блок 52 квантування. Приведені вектори 55 переднього плану V[k] можуть мати розміри D: $[(N+1)^2 - (N_{BG}+1)^2 - nBGa] \times nFG$.

[527] Блок 46 приведення коефіцієнтів може включати в себе блок 228 мінімізації коефіцієнтів, що може представляти блок, виконаний з можливістю зниження або іншої мінімізації розміру кожного з векторів 53, що залишилися, переднього плану V[k] шляхом видалення будь-яких коефіцієнтів, що враховуються в коефіцієнтах 47 НОА фону (ідентифікованих інформацією 43 фонового каналу). Блок 228 мінімізації коефіцієнтів може видаляти коефіцієнти, ідентифіковані інформацією 43 фонового каналу, для одержання приведених векторів 55 переднього плану V[k].

[528] На фіг. 9J показана блок-схема, що демонструє, більш детально, блок 40 психоакустичного аудіокодера пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. Блок 40 психоакустичного аудіокодера може представляти блок, виконаний з можливістю здійснення психоакустичного кодування відносно коефіцієнтів 47' НОА фону з компенсацією енергії й інтерпольованих сигналів 49' nFG. Як показано в прикладі, наведеному на фіг. 9H, блок 40 психоакустичного аудіокодера може викликати множинні екземпляри психоакустичних аудіокодерів 40A-40N для аудіокодування кожного з каналів коефіцієнтів 47' НОА фону з компенсацією енергії (де канал у цьому контексті означає коефіцієнти для усіх вибірок у кадрі, що відповідає сферичній базисній функції конкретного порядку/підпорядку) і кожного з сигналів з інтерпольованих сигналів 49' nFG. У деяких прикладах, блок 40 психоакустичного аудіокодера представляє або інакше включає в себе (при реалізації апаратними засобами) аудіокодери 40A-40N у достатній кількості, щоб окремо кодувати кожен канал коефіцієнтів 47' НОА фону з компенсацією енергії (або nBGa плюс сумарна кількість індексів (i)) і кожен сигнал з інтерпольованих сигналів 49' nFG (або nFG) для усього nBGa плюс сумарна кількість індексів (i) додаткових каналів НОА оточення плюс nFG. Аудіокодери 40A-40N можуть виводити кодовані коефіцієнти 59 НОА фону і кодовані сигнали 61 nFG.

[529] На фіг. 9K показана блок-схема, що демонструє, більш детально, блок 52 квантування пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. У прикладі, наведеному на фіг. 9K, блок 52 квантування включає в себе блок 230 однорідного квантування, блок 232 nbits, блок 234 прогнозування, блок 236 режиму прогнозування ("блок 236 режиму прогнозування"), блок 238 кодування категорії і залишку і блок 240 вибору таблиці Хаффмана. Блок 230 однорідного квантування представляє блок, виконаний з можливістю здійснення однорідного квантування, описаного вище відносно одного з просторових компонентів (який може представляти будь-який з приведених векторів 55 переднього плану V[k]). Блок 232 nbits представляє блок, виконаний з можливістю визначення параметра або значення nbits.

[530] Блок 234 прогнозування представляє блок, виконаний з можливістю здійснення прогнозування відносно квантованого просторового компонента. Блок 234 прогнозування може здійснювати прогнозування шляхом здійснення поелементного віднімання поточного одного з приведених векторів 55 переднього плану V[k] з наступного в часі відповідного одного з приведених векторів 55 переднього плану V[k] (які можуть бути позначені як приведені вектори 55 переднього плану V[k-1]). Результат цього прогнозування може іменуватися прогнозованим просторовим компонентом.

[531] Блок 236 режиму прогнозування може представляти блок, виконаний з можливістю вибору режиму прогнозування. Блок 240 вибору таблиці Хаффмана може представляти блок, виконаний з можливістю вибору належної таблиці Хаффмана для кодування *cid*. Блок 236 режиму прогнозування і блок 240 вибору таблиці Хаффмана можуть діяти, у порядку одного

5 прикладу, згідно з наступним псевдокодом:

для даного *nbits*, витягнути всі таблиці Хаффмана, що мають *nbits*
B00=0; B01=0; B10=0; B11=0; // ініціалізувати для обчислення очікуваних бітів для кожного
 режиму кодування

for *m=1:(кількість елементів у векторі)*

10 // обчислювати очікувану кількість бітів для елемента вектора *v(m)*

// без прогнозування і з використанням таблиці Хаффмана 5

B00=B00+calculate_bits(v(m), HT5);

// без прогнозування і з використанням таблиці Хаффмана {1,2,3}

B01=B01+calculate_bits(v(m), HTq); q у {1,2,3}

15 // обчислювати очікувану кількість бітів для залишку прогнозування *e(m)*

e(m)=v(m)-vp(m); // vp(m): елемент вектора попереднього кадру

// з прогнозуванням і з використанням таблиці Хаффмана 4

B10=B10+calculate_bits(e(m), HT4);

// з прогнозуванням і з використанням таблиці Хаффмана 5

20 *B11=B11+calculate_bits(e(m), HT5);*

end

// знайти найкращий режим прогнозування і таблицю Хаффмана, що дають мінімальну кількість бітів

// найкращі режим прогнозування і таблиця Хаффмана позначені прапорами *pflag* і *Htflag*,

25 відповідно

[Be, id]=min([B00 B01 B10 B11]);

Switch *id*

case 1: *pflag=0; HTflag=0;*

case 2: *pflag=0; HTflag=1;*

30 case 3: *pflag=1; HTflag=0;*

case 4: *pflag=1; HTflag=1;*

end

[532] Блок 238 кодування категорії і залишку може представляти блок, виконаний з
 35 можливістю здійснення категоризації і залишкового кодування прогнозованого просторового компонента або квантованого просторового компонента (коли прогнозування заборонене), як більш детально описано вище.

[533] Як показано в прикладі, наведеному на фіг. 9K, блок 52 квантування може виводити різні параметри або значення для включення або в бітовий потік 21, або в побічну інформацію (яка сама може бути бітовим потоком, окремим від бітового потоку 21). Передбачаючи, що
 40 інформація вказується в інформації побічного каналу, блок 50 скалярного/ентропійного квантування може виводити значення *nbits* як значення 233 *nbits*, режим прогнозування як режим 237 прогнозування й інформацію таблиці Хаффмана як інформацію 241 таблиці Хаффмана на блок 42 генерації бітового потоку спільно зі стисненою версією просторового компонента (показаною як кодовані вектори 57 переднього плану *V[k]* у прикладі, наведеному на
 45 фіг. 4), який у цьому прикладі може означати код Хаффмана, вибраний для кодування *cid*, знакового біта і блоково кодованого залишку. Значення *nbits* може вказуватися один раз в інформації побічного каналу для всіх кодованих векторів 57 переднього плану *V[k]*, тоді як режим прогнозування й інформація таблиці Хаффмана можуть указуватися для кожного з кодованих векторів 57 переднього плану *V[k]*. Частина бітового потоку, що вказує стиснену
 50 версію просторового компонента, додатково показана в прикладі, наведеному на фіг. 10B і/або 10C.

[534] На фіг. 9L показана блок-схема, що демонструє, більш детально, блок 42 генерації бітового потоку пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. Блок 42 генерації бітового потоку може включати в себе блок 242 генерації інформації головного каналу і блок 244 генерації інформації побічного каналу. Блок 242 генерації інформації головного каналу може генерувати головний бітовий потік 21, що включає в себе один або
 55 більше, якщо не всі, з індексів 205 переупорядковування, елемента 254 синтаксису *CodedSpatialInterpolationTime*, елемента 255 синтаксису *SpatialInterpolationMethod*, кодованих коефіцієнтів 59 НОА фону і кодованих сигналів 61 *nFG*. Блок 244 генерації інформації побічного каналу може представляти блок, виконаний з можливістю генерації бітового потоку 21В

побічного каналу, що може включати в себе одне або більше, якщо не всі, зі значення 233 nbits, режиму 237 прогнозування, інформації 241 таблиці Хаффмана і кодованих векторів 57 переднього плану V[k]. Бітові потоки 21 і 21B можуть спільно іменуватися бітовим потоком 21. У деяких контекстах, бітовий потік 21 може означати тільки бітовий потік 21 головного каналу, тоді як бітовий потік 21B може іменуватися інформацією 21B побічного каналу.

[535] На фіг. 10A-10O(ii) показані схеми, що демонструють ділянки бітового потоку або інформації побічного каналу, що може більш детально вказувати стиснені просторові компоненти. У прикладі, наведеному на фіг. 10A, частина 250 включає в себе поле 251 ідентифікатора рендеризатора ("ID рендеризатора") і поле 252 HOADecoderConfig. Поле 251 ID рендеризатора може представляти поле, де зберігається ID рендеризатора, використаний для мікшування контенту HOA. Поле 252 HOADecoderConfig може представляти поле, виконане з можливістю збереження інформації для ініціалізації просторового декодера HOA.

[536] Поле 252 HOADecoderConfig додатково включає в себе поле 253 інформації направленості ("інформація напрямку"), поле 254 CodedSpatialInterpolationTime, поле 255 SpatialInterpolationMethod, поле 256 CodedVVecLength і поле 257 інформації коефіцієнта підсилення. Поле 253 інформації направленості може представляти поле, де зберігається інформація для конфігурування декодера синтезу на основі напрямку. Поле 254 CodedSpatialInterpolationTime може представляти поле, де зберігається час просторово-часової інтерполяції сигналів на векторній основі. Поле 255 SpatialInterpolationMethod може представляти поле, де зберігається указання типу інтерполяції, застосовуваного при виконанні просторово-часової інтерполяції сигналів на векторній основі. Поле 256 CodedVVecLength може представляти поле, де зберігається довжина переданого вектора даних, використовуваного для синтезу сигналів на векторній основі. Поле 257 інформації коефіцієнта підсилення представляє поле, де зберігається інформація, що вказує корекцію коефіцієнта підсилення, застосовувану до сигналів.

[537] У прикладі, наведеному на фіг. 10B, частина 258A представляє частину каналу побічної інформації, де частина 258A включає в себе заголовок 259 кадру, що включає в себе поле 260 кількості байтів і поле 261 Nbits. Поле 260 кількості байтів може представляти поле для вираження кількості байтів, включених у кадр для указання просторових компонентів з v1 по vp, які включають у себе нулі для поля 264 побайтового вирівнювання. Поле 261 Nbits представляє поле, що може вказувати значення nbits, ідентифіковане для використання при знятті стиснення просторових компонентів v1-vp.

[538] Як показано далі в прикладі, наведеному на фіг. 10B, частина 258A може включати в себе бітові підпотоки для v1-vp, кожний з яких включає в себе поле 262 режиму прогнозування, інформаційне поле 263 таблиці Хаффмана і відповідний один зі стиснених просторових компонентів v1-vp. Поле 262 режиму прогнозування може представляти поле для збереження указання, чи здійснювалося прогнозування відносно відповідного одного зі стиснених просторових компонентів v1-vp. Інформаційне поле 263 таблиці Хаффмана представляє поле для указання, щонайменше частково, яка таблиця Хаффмана підлягає використанню для декодування різних аспектів відповідного одного зі стиснених просторових компонентів v1-vp.

[539] У цьому відношенні, методи можуть передбачати, що пристрій 20 аудіокодування одержує бітовий потік, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[540] На фіг. 10C показана схема, що демонструє альтернативний приклад частини 258B інформації побічного каналу, яка може більш детально вказувати стиснені просторові компоненти. У прикладі, наведеному на фіг. 10C, частина 258B включає в себе заголовок 259 кадру, що включає в себе поле 261 Nbits. Поле 261 Nbits представляє поле, що може вказувати значення nbits, ідентифіковане для використання при знятті стиснення просторових компонентів v1-vp.

[541] Як показано далі в прикладі, наведеному на фіг. 10C, частина 258B може включати в себе бітові підпотоки для v1-vp, кожний з яких включає в себе поле 262 режиму прогнозування, інформаційне поле 263 таблиці Хаффмана і відповідний один зі стиснених просторових компонентів v1-vp. Поле 262 режиму прогнозування може представляти поле для збереження указання, чи здійснювалося прогнозування відносно відповідного одного зі стиснених просторових компонентів v1-vp. Інформаційне поле 263 таблиці Хаффмана представляє поле для указання, щонайменше частково, яка таблиця Хаффмана підлягає використанню для декодування різних аспектів відповідного одного зі стиснених просторових компонентів v1-vp.

[542] Поле 261 Nbits у ілюстрованому прикладі включає в себе підполя A 265, B 266 і C 267. У цьому прикладі, A 265 і B 266 є 1-бітовими підполями, тоді як C 267 є 2-бітовим підполем. Інші

приклади можуть включати в себе підполя 265, 266 і 267 різних розмірів. Поле 265 A і поле 266 B можуть представляти поля, де зберігаються перший і другий старші біти поля 261 Nbits, тоді як поле 267 C може представляти поле, де зберігаються молодші біти поля 261 Nbits.

[543] Частина 258B також може включати в себе поле 268 AddAmbHoalInfoChannel. Поле 268 AddAmbHoalInfoChannel може представляти поле, де зберігається інформація для додаткових коефіцієнтів НОА оточення. Як показано в прикладі, наведеному на фіг. 10C, AddAmbHoalInfoChannel 268 включає в себе поле 246 CodedAmbCoeffldx, поле 247 AmbCoeffldxTransition. Поле 246 CodedAmbCoeffldx може представляти поле, де зберігається індекс додаткового коефіцієнта НОА оточення. Поле 247 AmbCoeffldxTransition може представляти поле, виконане з можливістю збереження даних, що вказують, чи збільшується, у цьому кадрі, додатковий коефіцієнт НОА оточення або зменшується.

[544] На фіг. 10C(i) показана схема, що демонструє альтернативний приклад частини 258B' інформації побічного каналу, яка може більш детально вказувати стиснені просторові компоненти. У прикладі, наведеному на фіг. 10C(i), частина 258B' включає в себе заголовок 259 кадру, що включає в себе поле 261 Nbits. Поле 261 Nbits представляє поле, що може вказувати значення nbits, ідентифіковане для використання при знятті стиснення просторових компонентів v1-vn.

[545] Як показано далі в прикладі, наведеному на фіг. 10C(i), частина 258B' може включати в себе бітові підпотоки для v1-vn, кожний з яких включає в себе інформаційне поле 263 таблиці Хаффмана і відповідний один зі стиснених направлених компонентів v1-vn, що не включають у себе поле 262 режиму прогнозування. У всіх інших відношеннях, частина 258B' може бути аналогічна частині 258B.

[546] На фіг. 10D показана схема, що більш детально демонструє частину 258C бітового потоку 21. Частина 258C аналогічна частині 258, за винятком того, що заголовок 259 кадру і побайтове вирівнювання 264 нуля видалені, тоді як поле 261 Nbits додане перед кожним з бітових потоків для v1-vn, як показано в прикладі, наведеному на фіг. 10D.

[547] На фіг. 10D(i) показана схема, що демонструє частину 258C' бітового потоку 21 більш детально. Частина 258C' аналогічна частині 258C, за винятком того, що частина 258C' не включає в себе поле 262 режиму прогнозування для кожного з векторів V v1-vn.

[548] На фіг. 10E показана схема, що більш детально демонструє частину 258D бітового потоку 21. Частина 258D аналогічна частині 258B, за винятком того, що заголовок 259 кадру і побайтове вирівнювання 264 нуля видалені, тоді як поле 261 Nbits додане перед кожним з бітових потоків для v1-vn, як показано в прикладі, наведеному на фіг. 10E.

[549] На фіг. 10E(i) показана схема, що демонструє частину 258D' бітового потоку 21 більш детально. Частина 258D' аналогічна частині 258D, за винятком того, що частина 258D' не включає в себе поле 262 режиму прогнозування для кожного з векторів V v1-vn. У цьому відношенні, пристрій 20 аудіокодування може генерувати бітовий потік 21, що не включає в себе поле 262 режиму прогнозування для кожного стисненого вектора V, як показано відносно прикладів, наведених на фіг. 10C(i), 10D(i) і 10E(i).

[550] На фіг. 10F показана схема, що демонструє, іншим чином, частину 250 бітового потоку 21, показаного в прикладі, наведеному на фіг. 10A. Частина 250, показана в прикладі, наведеному на фіг. 10D, включає в себе поле HOAOrder (яке не показане в прикладі, наведеному на фіг. 10F, для спрощення ілюстрації), поле MinAmbHoaOrder (яке, знову ж, не показане в прикладі, наведеному на фіг. 10, для спрощення ілюстрації), поле 253 інформації напрямку, поле 254 CodedSpatialInterpolationTime, поле 255 SpatialInterpolationMethod, поле 256 CodedVVecLength і поле 257 інформації коефіцієнта підсилення. Як показано в прикладі, наведеному на фіг. 10F, поле 254 CodedSpatialInterpolationTime може містити трибітове поле, поле 255 SpatialInterpolationMethod може містити однобітове поле і поле 256 CodedVVecLength може містити двобітове поле.

[551] На фіг. 10G показана схема, що більш детально демонструє частину 248 бітового потоку 21. Частина 248 представляє тривимірне (3D) корисне навантаження уніфікованого мовного/аудіокодера (USAC), що включає в себе поле 249 HOAFrame (яке також може бути позначене як побічна інформація, інформація побічного каналу або бітовий потік побічного каналу). Як показано в прикладі, наведеному на фіг. 10E, збільшений вигляд поля 249 HOAFrame може бути аналогічний частині 258B бітового потоку 21, показаного в прикладі, наведеному на фіг. 10C. "ChannelSideInfoData" включає в себе поле 269 ChannelType, що не показане в прикладі, наведеному на фіг. 10C, для спрощення ілюстрації, поле 265 A, позначене як "ba" у прикладі, наведеному на фіг. 10E, поле 266 B, позначене як "bb" у прикладі, наведеному на фіг. 10E, і поле 267 C позначене як "unitC" у прикладі, наведеному на фіг. 10E. Поле ChannelType указує, чи є канал сигналом на основі напрямку, сигналом на векторній

основі або додатковим коефіцієнтом HOA оточення. Між різними ChannelSideInfoData присутні поля 268 AddAmbHoaInfoChannel, де різні бітові потоки вектора V позначені сірим (наприклад, "бітовий потік для v1" і "бітовий потік для v2").

[552] На фіг. 10H-10O(ii) показані схеми, що більш детально демонструють інші різні зразкові ділянки 248H-248O бітового потоку 21 разом із супутніми ділянками 250H-250O HOAconfig. Фіг. 10H(i) і 10H(ii) ілюструють перший приклад бітового потоку 248H і супутню частину 250H конфігурації HOA, згенерованої згідно з випадком 0 у вищенаведеному псевдокоді. У прикладі, наведеному на фіг. 10H(i), частина 250H HOAconfig включає в себе елемент 256 синтаксису CodedVVecLength, заданий для вказання, що кодовані всі елементи вектора V, наприклад, всі 16 елементів вектора V. Частина 250H HOAconfig також включає в себе елемент 255 синтаксису SpatialInterpolationMethod, заданий для вказання, що інтерполяційною функцією просторово-часової інтерполяції є піднятий косинус. Частина 250H HOAconfig, крім того, включає в себе CodedSpatialInterpolationTime 254, заданий для вказання інтерпольованої тривалості вибірки, що дорівнює 256. Частина 250H HOAconfig додатково включає в себе елемент 150 синтаксису MinAmbHoaOrder, заданий для вказання, що порядок MinimumHOA контенту HOA оточення дорівнює одиниці, де пристрій 24 аудіодекодування може виводити елемент синтаксису MinNumOfCoeffsForAmbHOA рівним $(1+1)^2$ або чотирьом. Частина 250H HOAconfig включає в себе елемент 152 синтаксису HoaOrder, заданий для вказання порядку HOA контенту рівним трьом (або, іншими словами, $N=3$), де пристрій 24 аудіодекодування може виводити NumOfHoaCoeffs рівним $(N+1)^2$ або 16.

[553] Як показано далі в прикладі, наведеному на фіг. 10H(i), частина 248H включає в себе уніфікований тривимірний (USAC-3D) кадр аудіо мовного й аудіокодування (USAC), у якому два кадри HOA 249A і 249B зберігаються в корисному навантаженні розширення USAC, за умови, що два кадри аудіо зберігаються в одному кадрі USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрій 24 аудіодекодування може виводити кількість адаптивних транспортних каналів як функцію елемента синтаксису numHOATransportChannels і елемента синтаксису MinNumOfCoeffsForAmbHOA. У нижченаведених прикладах, передбачається, що елемент синтаксису numHOATransportChannels дорівнює 7 і елемент синтаксису MinNumOfCoeffsForAmbHOA дорівнює чотирьом, де кількість адаптивних транспортних каналів дорівнює елементу синтаксису numHOATransportChannels мінус елемент синтаксису MinNumOfCoeffsForAmbHOA (або трьом).

[554] Фіг. 10H(ii) більш детально демонструє кадри 249A і 249B. Як показано в прикладі, наведеному на фіг. 10H(ii), кадр 249A включає в себе поля 154-154C ChannelSideInfoData (CSID), поля HOAGainCorrectionData (HOAGCD), поля 156 і 156B VVectorData і поля HOAPredictionInfo. Поле 154 CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10H(i). Поле 154B CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10H(ii). Поле 154C CSID включає в себе поле 269 ChannelType, що має значення 3. Кожне з полів 154-154C CSID відповідає відповідному одному з транспортних каналів 1, 2 і 3. Фактично, кожне поле 154-154C CSID указує, чи є відповідне корисне навантаження 156 і 156B сигналами на основі напрямку (коли відповідний ChannelType дорівнює нулю), сигналами на векторній основі (коли відповідний ChannelType дорівнює одиниці), додатковим коефіцієнтом HOA оточення (коли відповідний ChannelType дорівнює двом) або порожнім (коли ChannelType дорівнює трьом).

[555] У прикладі, наведеному на фіг. 10H(ii), кадр 249A включає в себе два сигнали на векторній основі (за умови, що ChannelType 269 дорівнює 1 у полях 154 і 154B CSID) і порожній (за умови, що ChannelType 269 дорівнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 250H HOAconfig, пристрій 24 аудіодекодування може визначити, що кодовані всі 16 елементів вектора V. Тому кожен VVectorData 156 і 156B включає в себе всі 16 елементів вектора, кожний з яких однорідно квантується 8 бітами. Як зазначено вище, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0. Крім того, як зазначено однією зірочкою (*), схема кодування сигналізується за допомогою NbitsQ=5 у полі CSID для відповідного транспортного каналу.

[556] У кадрі 249B, поля 154 і 154B CSID такі ж, як у кадрі 249, тоді як поле 154C CSID кадру 249B переключається на ChannelType, що дорівнює одиниці. Таким чином, поле 154C CSID кадру 249B включає в себе Cbflag 267, Pflag 267 (що вказує хаффманівське кодування) і Nbits 261 (що дорівнює дванадцяти). В результаті, кадр 249B включає в себе третє поле 156C VVectorData, що включає в себе 16 елементів вектора V, кожний з яких однорідно квантується 12 бітами і кодується по Хаффману. Як згадано вище, кількість і індекси кодованих елементів

VVectorData задаються параметром CodedVVecLength=0, тоді як схема кодування Хаффмана сигналізується за допомогою NbitsQ=12, CbFlag=0 і Pflag=0 у полі 154C CSID для цього конкретного транспортного каналу (наприклад, транспортного каналу № 3).

[557] Приклади, приведені на фіг. 10I(i) і 10I(ii) ілюструють другий приклад бітового потоку 248I і супутньої частини 250I конфігурації HOA, згенерованої згідно з випадком 0 у вищенаведеному псевдокоді. У прикладі, наведеному на фіг. 10I(i), частина 250I HOAconfig включає в себе елемент 256 синтаксису CodedVVecLength, заданий для указання, що кодовані всі елементи вектора V, наприклад всі 16 елементів вектора V. Частина 250I HOAconfig також включає в себе елемент 255 синтаксису SpatialInterpolationMethod, заданий для указання, що інтерполяційною функцією часової-тимчасової-просторово-часової інтерполяції є піднятий косинус. Частина 250I HOAconfig, крім того, включає в себе CodedSpatialInterpolationTime 254, заданий для указання інтерпольованої тривалості вибірки, що дорівнює 256.

[558] Частина 250I HOAconfig додатково включає в себе елемент 150 синтаксису MinAmbHoaOrder, заданий для указання, що порядок MinimumHOA контенту HOA оточення дорівнює одиниці, де пристрій 24 аудіодекодування може виводити елемент синтаксису MinNumOfCoeffsForAmbHOA рівним $(1+1)^2$ або чотирьом. Пристрій 24 аудіодекодування також може виводити елемент синтаксису MaxNoOfAddActiveAmbCoeffs як заданий рівним різниці між елементом синтаксису NumOfHoaCoeff і MinNumOfCoeffsForAmbHOA, який передбачається в цьому прикладі рівним 16-4 або 12. Пристрій 24 аудіодекодування також може виводити елемент синтаксису AmbAssignBits як заданий рівним $\text{ceil}(\log_2(\text{MaxNoOfAddActiveAmbCoeffs})) = \text{ceil}(\log_2(12)) = 4$. Частина 250H HOAconfig включає в себе елемент 152 синтаксису HoaOrder, заданий для указання порядку HOA контенту рівним трьом (або, іншими словами, N=3), де пристрій 24 аудіодекодування може виводити NumOfHoaCoeffs рівним $(N+1)^2$ або 16.

[559] Як показано далі в прикладі, наведеному на фіг. 10I(i), частина 248H включає в себе кадр аудіо USAC-3D, у якому два кадри HOA 249C і 249D зберігаються в корисному навантаженні розширення USAC, за умови, що два кадри аудіо зберігаються в одному кадрі USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрій 24 аудіодекодування може виводити кількість адаптивних транспортних каналів як функцію елемента синтаксису numHOATransportChannels і елемента синтаксису MinNumOfCoeffsForAmbHOA. У нижченаведених прикладах, передбачається, що елемент синтаксису numHOATransportChannels дорівнює 7 і елемент синтаксису MinNumOfCoeffsForAmbHOA дорівнює чотирьом, де кількість адаптивних транспортних каналів дорівнює елементу синтаксису numHOATransportChannels мінус елемент синтаксису MinNumOfCoeffsForAmbHOA (або трьом).

[560] Фіг. 10I(ii) більш детально демонструє кадри 249C і 249D. Як показано в прикладі, наведеному на фіг. 10I(ii), кадр 249C включає в себе поля 154-154C CSID і поля 156 VVectorData. Поле 154 CSID включає в себе CodedAmbCoeffIdx 246, AmbCoeffIdxTransition 247 (де подвійна зірочка (**)) указує, що, для адаптивного транспортного каналу № 1, внутрішній стан декодера передбачається тут рівним AmbCoeffIdxTransitionState=2, у результаті чого бітове поле CodedAmbCoeffIdx сигналізується або інакше вказується в бітовому потоці), і ChannelType 269 (який дорівнює двом, що сигналізує, що відповідне корисне навантаження є додатковим коефіцієнтом HOA оточення). Пристрій 24 аудіодекодування може виводити AmbCoeffIdx рівним CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA або 5 у цьому прикладі. Поле 154B CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10I(ii). Поле 154C CSID включає в себе поле 269 ChannelType, що має значення 3.

[561] У прикладі, наведеному на фіг. 10I(ii), кадр 249C включає в себе одиничний сигнал на векторній основі (за умови, що ChannelType 269 дорівнює 1 у полях CSID 154B) і порожній (за умови, що ChannelType 269 дорівнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 250I HOAconfig, пристрій 24 аудіодекодування може визначити, що кодовані всі 16 елементів вектора V. Тому VVectorData 156 включає в себе всі 16 елементів вектора, кожний з яких однорідно квантується 8 бітами. Як зазначено виноскою 1, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0. Крім того, як зазначено виноскою 2, схема кодування сигналізується за допомогою NbitsQ=5 у полі CSID для відповідного транспортного каналу.

[562] У кадрі 249D, поле 154 CSID включає в себе AmbCoeffIdxTransition 247, що вказує, що переходу не відбулося, і, таким чином, CodedAmbCoeffIdx 246 можна вивести з попереднього кадру і не потрібно повторно сигналізувати або інакше вказувати. Поля 154B і 154C CSID кадру 249D такі ж, як для кадру 249C, і, таким чином, як і кадр 249C, кадр 249D включає в себе

одиночне поле 156 VVectorData, що включає в себе всі 16 елементів вектора, кожний з яких однорідно квантується 8 бітами.

[563] Фіг. 10J(i) і 10J(ii) ілюструють перший приклад бітового потоку 248J і супутньої частини 250J конфігурації HOA, згенерованої згідно з випадком 1 у вищенаведеному псевдокоді. У прикладі, наведеному на фіг. 10J(i), частина 250J HOAconfig включає в себе елемент 256 синтаксису CodedVVecLength, заданий для вказання, що кодовані всі елементи вектора V, за винятком елементів від 1 до елементів синтаксису MinNumOfCoeffsForAmbHOA і елементів, зазначених в елементі синтаксису ContAddAmbHoaChan (передбачається рівним нулю в цьому прикладі). Частина 250J HOAconfig також включає в себе елемент 255 синтаксису SpatialInterpolationMethod, заданий для вказання, що інтерполяційною функцією просторово-часової інтерполяції є піднятий косинус. Частина 250J HOAconfig, крім того, включає в себе CodedSpatialInterpolationTime 254, заданий для вказання інтерпольованої тривалості вибірки, що дорівнює 256. Частина 250J HOAconfig додатково включає в себе елемент 150 синтаксису MinAmbHoaOrder, заданий для вказання, що порядок MinimumHOA контенту HOA оточення дорівнює одиниці, де пристрій 24 аудіодекодування може виводити елемент синтаксису MinNumOfCoeffsForAmbHOA рівним $(1+1)^2$ або чотирьом. Частина 250J HOAconfig включає в себе елемент 152 синтаксису HoaOrder, заданий для вказання порядку HOA контенту рівним трьом (або, іншими словами, $N=3$), де пристрій 24 аудіодекодування може виводити NumOfHoaCoeffs рівним $(N+1)^2$ або 16.

[564] Як показано далі в прикладі, наведеному на фіг. 10J(i), частина 248J включає в себе кадр аудіо USAC-3D, у якому два кадри HOA 249E і 249F зберігаються в корисному навантаженні розширення USAC, за умови, що два кадри аудіо зберігаються в одному кадрі USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрій 24 аудіодекодування може виводити кількість адаптивних транспортних каналів як функцію елемента синтаксису numHOATransportChannels і елемента синтаксису MinNumOfCoeffsForAmbHOA. У нижченаведених прикладах, передбачається, що елемент синтаксису numHOATransportChannels дорівнює 7 і елемент синтаксису MinNumOfCoeffsForAmbHOA дорівнює чотирьом, де кількість адаптивних транспортних каналів дорівнює елементу синтаксису numHOATransportChannels мінус елемент синтаксису MinNumOfCoeffsForAmbHOA (або трьом).

[565] Фіг. 10J(ii) більш детально демонструє кадри 249E і 249F. Як показано в прикладі, наведеному на фіг. 10J(ii), кадр 249E включає в себе поля 154-154C CSID і поля 156 і 156B VVectorData. Поле 154 CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10J(i). Поле 154B CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10J(ii). Поле 154C CSID включає в себе поле 269 ChannelType, що має значення 3. Кожне з полів 154-154C CSID відповідає відповідному одному з транспортних каналів 1, 2 і 3.

[566] У прикладі, наведеному на фіг. 10J(ii), кадр 249E включає в себе два сигнали на векторній основі (за умови, що ChannelType 269 дорівнює 1 у полях 154 і 154B CSID) і порожній (за умови, що ChannelType 269 дорівнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 250H HOAconfig, пристрій 24 аудіодекодування може визначити, що всі 12 елементів вектора V кодуються (де 12 виводиться як $(\text{HoaOrder}+1)^2 - (\text{MinNumOfCoeffsForAmbHOA}) - (\text{ContAddAmbHoaChan}) = 16 - 4 - 0 = 12$). Тому кожен VVectorData 156 і 156B включає в себе всі 12 елементів вектора, кожний з яких однорідно квантується 8 бітами. Як зазначено вище, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0. Крім того, як зазначено однією зірочкою (*), схема кодування сигналізується за допомогою NbitsQ=5 у полі CSID для відповідного транспортного каналу.

[567] У кадрі 249F, полючи 154 і 154B CSID такі ж, як у кадрі 249E, тоді як поле 154C CSID кадру 249F переключається на ChannelType, що дорівнює одиниці. Таким чином, поле 154C CSID кадру 249B включає в себе Cbflag 267, Pflag 267 (що вказує хаффманівське кодування) і Nbits 261 (що дорівнює дванадцяти). У результаті, кадр 249F включає в себе третє поле 156C VVectorData, що включає в себе 12 елементів вектора V, кожний з яких однорідно квантується 12 бітами і кодується по Хаффману. Як згадано вище, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0, тоді як схема кодування Хаффмана сигналізується за допомогою NbitsQ=12, CbFlag=0 і Pflag=0 у полі 154C CSID для цього конкретного транспортного каналу (наприклад, транспортного каналу № 3).

[568] Приклади, наведені на фіг. 10K(i) і 10K(ii), ілюструють другий приклад бітового потоку 248K і супутньої частини 250K конфігурації HOA, згенерованої згідно з випадком 1 у

вищенаведеному псевдокодi. У прикладi, наведеному на фiг. 10K(i), дiлянки 250K HOAconfig включають в себе елемент 256 синтаксису CodedVVecLength, заданий для указання, що кодованi всi елементи вектора V, за винятком елементiв вiд 1 до елементiв синтаксису MinNumOfCoeffsForAmbHOA i елементiв, зазначених в елементi синтаксису ContAddAmbHoaChan (передбачається рiвним одиницi в цьому прикладi). Частина 250K HOAconfig також включає в себе елемент 255 синтаксису SpatialInterpolationMethod, заданий для указання, що iнтерполяцiйною функцiєю просторово-часової iнтерполяцiї є пiднятий косинус. Частина 250K HOAconfig, крiм того, включає в себе CodedSpatialInterpolationTime 254, заданий для указання iнтерпольованої тривалостi вибiрки, що дорiвнює 256.

[569] Частина 250K HOAconfig додатково включає в себе елемент 150 синтаксису MinAmbHoaOrder, заданий для указання, що порядок MinimumHOA контенту HOA оточення дорiвнює одиницi, де пристрiй 24 аудiодекодування може виводити елемент синтаксису MinNumOfCoeffsForAmbHOA рiвним $(1+1)^2$ або чотирьом. Пристрiй 24 аудiодекодування також може виводити елемент синтаксису MaxNoOfAddActiveAmbCoeffs як заданий рiвним рiзницi мiж елементом синтаксису NumOfHoaCoeff i MinNumOfCoeffsForAmbHOA, що передбачається в цьому прикладi рiвним 16-4 або 12. Пристрiй 24 аудiодекодування також може виводити елемент синтаксису AmbAssignBits як заданий рiвним $\text{ceil}(\log_2(\text{MaxNoOfAddActiveAmbCoeffs})) = \text{ceil}(\log_2(12)) = 4$. Частина 250K HOAconfig включає в себе елемент 152 синтаксису HoaOrder, заданий для указання порядку HOA контенту рiвним трьом (або, iншими словами, $N=3$), де пристрiй 24 аудiодекодування може виводити NumOfHoaCoeffs рiвним $(N+1)^2$ або 16.

[570] Як показано далi в прикладi, наведеному на фiг. 10K(i), частина 248K включає в себе кадр аудiо USAC-3D, у якому два кадри 249G i 249H HOA зберiгаються в корисному навантаженнi розширення USAC, за умови, що два кадри аудiо зберiгаються в одному кадрi USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрiй 24 аудiодекодування може виводити кiлькiсть адаптивних транспортних каналiв як функцiю елемента синтаксису numHOATransportChannels i елемента синтаксису MinNumOfCoeffsForAmbHOA. У нижченаведених прикладах, передбачається, що елемент синтаксису numHOATransportChannels дорiвнює 7 i елемент синтаксису MinNumOfCoeffsForAmbHOA дорiвнює чотирьом, де кiлькiсть адаптивних транспортних каналiв дорiвнює елементу синтаксису numHOATransportChannels мiнус елемент синтаксису MinNumOfCoeffsForAmbHOA (або трьом).

[571] Фiг. 10K(ii) бiльш детально демонструє кадри 249G i 249H. Як показано в прикладi, наведеному на фiг. 10K(ii), кадр 249G включає в себе поля 154-154C CSID i поля 156 VVectorData. Поле 154 CSID включає в себе CodedAmbCoeffIdx 246, AmbCoeffIdxTransition 247 (де подвiйна зiрочка (**)) указує, що, для адаптивного транспортного каналу № 1, внутрiшнiй стан декодера передбачається тут рiвним AmbCoeffIdxTransitionState=2, у результатi чого бiтове поле CodedAmbCoeffIdx сигналізується або iнакше вказується в бiтовому потоцi) i ChannelType 269 (який дорiвнює двом, що сигналізує, що вiдповiдне корисне навантаження є додатковим коефiцiєнтом HOA оточення). Пристрiй 24 аудiодекодування може виводити AmbCoeffIdx рiвним $\text{CodedAmbCoeffIdx} + 1 + \text{MinNumOfCoeffsForAmbHOA}$ або 5 у цьому прикладi. Поле 154B CSID включає в себе unitC 267, bb 266 i ba 265 разом з ChannelType 269, кожне з яких задане рiвним вiдповiдним значенням 01, 1, 0 i 01, показаним у прикладi, наведеному на фiг. 10K(ii). Поле 154C CSID включає в себе поле 269 ChannelType, що має значення 3.

[572] У прикладi, наведеному на фiг. 10K(ii), кадр 249G включає в себе одиничний сигнал на векторнiй основi (за умови, що ChannelType 269 дорiвнює 1 у полях CSID 154B) i порожнiй (за умови, що ChannelType 269 дорiвнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 250K HOAconfig, пристрiй 24 аудiодекодування може визначити, що 11 елементiв вектора V кодуються (де 12 виводиться як $(\text{HOAOrder} + 1)^2 - (\text{MinNumOfCoeffsForAmbHOA}) - (\text{ContAddAmbHoaChan}) = 16 - 4 - 1 = 11$). Тому VVectorData 156 включає в себе всi 11 елементiв вектора, кожний з яких однорiдно квантується 8 бiтами. Як зазначено виноскою 1, кiлькiсть i iндекси кодованих елементiв VVectorData задаються параметром CodedVVecLength=0. Крiм того, як зазначено виноскою 2, схема кодування сигналізується за допомогою NbitsQ=5 у полi CSID для вiдповiдного транспортного каналу.

[573] У кадрi 249H, поле 154 CSID включає в себе AmbCoeffIdxTransition 247, що вказує, що переходу не вiдбулося, i, таким чином, CodedAmbCoeffIdx 246 можна вивести з попереднього кадру i не потрiбно повторно сигналізувати або iнакше вказувати. Поля 154B i 154C CSID кадру 249H такi ж, як для кадру 249G, i, таким чином, як i кадр 249G, кадр 249H включає в себе одиничне поле 156 VVectorData, що включає в себе 11 елементiв вектора, кожний з яких однорiдно квантується 8 бiтами.

[574] Фіг. 10L(i) і 10L(ii) ілюструють перший приклад бітового потоку 248L і супутньої частини 250L конфігурації HOA, згенерованої згідно з випадком 2 у вищенаведеному псевдокоді. У прикладі, наведеному на фіг. 10L(i), частина 250L HOAconfig включає в себе елемент 256 синтаксису CodedVVecLength, заданий для вказання, що кодовані всі елементи вектора V, за винятком елементів від нульового порядку до порядку, зазначеного елементом 150 синтаксису MinAmbHoaOrder (який дорівнює $(HoaOrder+1)^2 - (MinAmbHoaOrder+1)^2 = 16 - 4 = 12$ у цьому прикладі). Частина 250L HOAconfig також включає в себе елемент 255 синтаксису SpatialInterpolationMethod, заданий для вказання, що інтерполяційною функцією просторово-часової інтерполяції є піднятий косинус. Частина 250L HOAconfig, крім того, включає в себе CodedSpatialInterpolationTime 254, заданий для вказання інтерпольованої тривалості вибірки, що дорівнює 256. Частина 250L HOAconfig додатково включає в себе елемент 150 синтаксису MinAmbHoaOrder, заданий для вказання, що порядок MinimumHOA контенту HOA оточення дорівнює одиниці, де пристрій 24 аудіодекодування може виводити елемент синтаксису MinNumOfCoeffsForAmbHOA рівним $(1+1)^2$ або чотирьом. Частина 250L HOAconfig включає в себе елемент 152 синтаксису HoaOrder, заданий для вказання порядку HOA контенту рівним трьом (або, іншими словами, $N=3$), де пристрій 24 аудіодекодування може виводити NumOfHoaCoeffs рівним $(N+1)^2$ або 16.

[575] Як показано далі в прикладі, наведеному на фіг. 10L(i), частина 248L включає в себе кадр аудіо USAC-3D, у якому два кадри 249I і 249J HOA зберігаються в корисному навантаженні розширення USAC, за умови, що два кадри аудіо зберігаються в одному кадрі USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрій 24 аудіодекодування може виводити кількість адаптивних транспортних каналів як функцію елемента синтаксису numHOATransportChannels і елемента синтаксису MinNumOfCoeffsForAmbHOA. У нижченаведених прикладах, передбачається, що елемент синтаксису numHOATransportChannels дорівнює 7 і елемент синтаксису MinNumOfCoeffsForAmbHOA дорівнює чотирьом, де кількість адаптивних транспортних каналів дорівнює елементу синтаксису numHOATransportChannels мінус елемент синтаксису MinNumOfCoeffsForAmbHOA (або трьом).

[576] Фіг. 10L(ii) більш детально демонструє кадри 249I і 249J. Як показано в прикладі, наведеному на фіг. 10L(ii), кадр 249I включає в себе поля 154-154C CSID і поля 156 і 156B VVectorData. Поле 154 CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10J(i). Поле 154B CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10L(ii). Поле 154C CSID включає в себе поле 269 ChannelType, що має значення 3. Кожне з полів 154-154C CSID відповідає відповідному одному з транспортних каналів 1, 2 і 3.

[577] У прикладі, наведеному на фіг. 10L(ii), кадр 249I включає в себе два сигнали на векторній основі (за умови, що ChannelType 269 дорівнює 1 у полях 154 і 154B CSID) і порожній (за умови, що ChannelType 269 дорівнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 250H HOAconfig, пристрій 24 аудіодекодування може визначити, що 12 елементів вектора V кодуються. Тому кожен VVectorData 156 і 156B включає в себе 12 елементів вектора, кожний з яких однорідно квантується 8 бітами. Як зазначено вище, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0. Крім того, як зазначено однією зірочкою (*), схема кодування сигналізується за допомогою NbitsQ=5 у полі CSID для відповідного транспортного каналу.

[578] У кадрі 249J, поля 154 і 154B CSID такі ж, як у кадрі 249I, тоді як поле 154C CSID кадру 249F переключається на ChannelType, що дорівнює одиниці. Таким чином, поле 154C CSID кадру 249B включає в себе Cbflag 267, Pflag 267 (що вказує хаффманівське кодування) і Nbits 261 (що дорівнює дванадцяти). У результаті, кадр 249F включає в себе третє поле 156C VVectorData, що включає в себе 12 елементів вектора V, кожний з яких однорідно квантується 12 бітами і кодується по Хаффману. Як згадано вище, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0, тоді як схема кодування Хаффмана сигналізується за допомогою NbitsQ=12, CbFlag=0 і Pflag=0 у полі 154C CSID для цього конкретного транспортного каналу (наприклад, транспортного каналу № 3).

[579] Приклади, наведені на фіг. 10M(i) і 10M(ii), ілюструють другий приклад бітового потоку 248M і супутньої частини 250M конфігурації HOA, згенерованої згідно з випадком 2 у вищенаведеному псевдокоді. У прикладі, наведеному на фіг. 10M(i), частина 250M HOAconfig включає в себе елемент 256 синтаксису CodedVVecLength, заданий для вказання, що кодовані всі елементи вектора V, за винятком елементів від нульового порядку до порядку, зазначеного

елементом 150 синтаксису MinAmbHoaOrder (який дорівнює $(\text{HoaOrder}+1)^2 - (\text{MinAmbHoaOrder}+1)^2 = 16-4=12$ у цьому прикладі). Частина 250M HOAconfig також включає в себе елемент 255 синтаксису SpatialInterpolationMethod, заданий для указання, що інтерполяційною функцією просторово-часової інтерполяції є піднятий косинус. Частина 250M HOAconfig крім того, включає в себе CodedSpatialInterpolationTime 254, заданий для указання інтерпольованої тривалості вибірки, що дорівнює 256.

[580] Частина 250M HOAconfig додатково включає в себе елемент 150 синтаксису MinAmbHoaOrder заданий для указання, що порядок MinimumHOA контенту HOA оточення дорівнює одиниці, де пристрій 24 аудіодекодування може виводити елемент синтаксису MinNumOfCoeffsForAmbHOA рівним $(1+1)^2$ або чотирьом. Пристрій 24 аудіодекодування також може виводити елемент синтаксису MaxNoOfAddActiveAmbCoeffs як заданий рівним різниці між елементом синтаксису NumOfHoaCoeff і MinNumOfCoeffsForAmbHOA, що передбачається в цьому прикладі рівним 16-4 або 12. Пристрій 24 аудіодекодування також може виводити елемент синтаксису AmbAssignBits як заданий рівним $\text{ceil}(\log_2(\text{MaxNoOfAddActiveAmbCoeffs})) = \text{ceil}(\log_2(12)) = 4$. Частина 250M HOAconfig включає в себе елемент 152 синтаксису HoaOrder, заданий для указання порядку HOA контенту рівним трьом (або, іншими словами, $N=3$), де пристрій 24 аудіодекодування може виводити NumOfHoaCoeffs рівним $(N+1)^2$ або 16.

[581] Як показано далі в прикладі, наведеному на фіг. 10M(i), частина 248M включає в себе кадр аудіо USAC-3D, у якому два кадри 249K і 249L HOA зберігаються в корисному навантаженні розширення USAC, за умови, що два кадри аудіо зберігаються в одному кадрі USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрій 24 аудіодекодування може виводити кількість адаптивних транспортних каналів як функцію елемента синтаксису numHOATransportChannels і елемента синтаксису MinNumOfCoeffsForAmbHOA. У нижченаведених прикладах, передбачається, що елемент синтаксису numHOATransportChannels дорівнює 7 і елемент синтаксису MinNumOfCoeffsForAmbHOA дорівнює чотирьом, де кількість адаптивних транспортних каналів дорівнює елементу синтаксису numHOATransportChannels мінус елемент синтаксису MinNumOfCoeffsForAmbHOA (або трьом).

[582] Фіг. 10M(ii) більш детально демонструє кадри 249K і 249L. Як показано в прикладі, наведеному на фіг. 10M(ii), кадр 249K включає в себе поля 154-154C CSID і поле 156 VVectorData. Поле 154 CSID включає в себе CodedAmbCoeffIdx 246, AmbCoeffIdxTransition 247 (де подвійна зірочка (**)) указує, що, для адаптивного транспортного каналу № 1, внутрішній стан декодера передбачається тут рівним AmbCoeffIdxTransitionState=2, у результаті чого бітове поле CodedAmbCoeffIdx сигналізується або інакше вказується в бітовому потоці) і ChannelType 269 (який дорівнює двом, що сигналізує, що відповідне корисне навантаження є додатковим коефіцієнтом HOA оточення). Пристрій 24 аудіодекодування може виводити AmbCoeffIdx рівним CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA або 5 у цьому прикладі. Поле 154B CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10M(ii). Поле 154C CSID включає в себе поле 269 ChannelType, що має значення 3.

[583] У прикладі, наведеному на фіг. 10M(ii), кадр 249K включає в себе одиничний сигнал на векторній основі (за умови, що ChannelType 269 дорівнює 1 у полях CSID 154B) і порожній (за умови, що ChannelType 269 дорівнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 250M HOAconfig, пристрій 24 аудіодекодування може визначити, що 12 елементів вектора V кодуються. Тому VVectorData 156 включає в себе 12 елементів вектора, кожний з яких однорідно квантується 8 бітами. Як зазначено виноскою 1, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0. Крім того, як зазначено виноскою 2, схема кодування сигналізується за допомогою NbitsQ=5 у полі CSID для відповідного транспортного каналу.

[584] У кадрі 249L, поле 154 CSID включає в себе AmbCoeffIdxTransition 247, що вказує, що переходу не відбулося, і, таким чином, CodedAmbCoeffIdx 246 можна вивести з попереднього кадру і не потрібно повторно сигналізувати або інакше вказувати. Поля 154B і 154C CSID кадру 249L такі ж, як для кадру 249K, і, таким чином, як і кадр 249K, кадр 249L включає в себе одиничне поле 156 VVectorData, що включає в себе 12 елементів вектора, кожний з яких однорідно квантується 8 бітами.

[585] Фіг. 10N(i) і 10N(ii) ілюструють перший приклад бітового потоку 248N і супутньої частини 250N конфігурації HOA, згенерованої згідно з випадком 3 у вищенаведеному псевдокоді. У прикладі, наведеному на фіг. 10N(i), частина 250N HOAconfig включає в себе елемент 256 синтаксису CodedVVecLength, заданий для указання, що кодовані всі елементи

вектора V , за винятком елементів, зазначених в елементі синтаксису `ContAddAmbHoaChan` (який передбачається рівним нулю в цьому прикладі). Частина 250N `HOAconfig` також включає в себе елемент 255 синтаксису `SpatialInterpolationMethod`, заданий для указання, що інтерполяційною функцією просторово-часової інтерполяції є піднятий косинус. Частина 250N `HOAconfig`, крім того, включає в себе `CodedSpatialInterpolationTime` 254, заданий для указання інтерпольованої тривалості вибірки, що дорівнює 256. Частина 250N `HOAconfig` додатково включає в себе елемент 150 синтаксису `MinAmbHoaOrder`, заданий для указання, що порядок `MinimumHOA` контенту HOA оточення дорівнює одиниці, де пристрій 24 аудіодекодування може виводити елемент синтаксису `MinNumOfCoeffsForAmbHOA` рівним $(1+1)^2$ або чотирьом. Частина 250N `HOAconfig` включає в себе елемент 152 синтаксису `HoaOrder`, заданий для указання порядку HOA контенту рівним трьом (або, іншими словами, $N=3$), де пристрій 24 аудіодекодування може виводити `NumOfHoaCoeffs` рівним $(N+1)^2$ або 16.

[586] Як показано далі в прикладі, наведеному на фіг. 10N(i), частина 248N включає в себе кадр аудіо USAC-3D, у якому два кадри HOA 249M і 249N зберігаються в корисному навантаженні розширення USAC, за умови, що два кадри аудіо зберігаються в одному кадрі USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрій 24 аудіодекодування може виводити кількість адаптивних транспортних каналів як функцію елемента синтаксису `numHOATransportChannels` і елемента синтаксису `MinNumOfCoeffsForAmbHOA`. У нижченаведених прикладах, передбачається, що елемент синтаксису `numHOATransportChannels` дорівнює 7 і елемент синтаксису `MinNumOfCoeffsForAmbHOA` дорівнює чотирьом, де кількість адаптивних транспортних каналів дорівнює елементу синтаксису `numHOATransportChannels` мінус елемент синтаксису `MinNumOfCoeffsForAmbHOA` (або трьом).

[587] Фіг. 10N(ii) більш детально демонструє кадри 249M і 249N. Як показано в прикладі, наведеному на фіг. 10N(ii), кадр 249M включає в себе поля 154-154C CSID і поля 156 і 156B `VVectorData`. Поле 154 CSID включає в себе `unitC` 267, `bb` 266 і `ba` 265 разом з `ChannelType` 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10J(i). Поле 154B CSID включає в себе `unitC` 267, `bb` 266 і `ba` 265 разом з `ChannelType` 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 10N(ii). Поле 154C CSID включає в себе поле 269 `ChannelType`, що має значення 3. Кожне з полів 154-154C CSID відповідає відповідному одному з транспортних каналів 1, 2 і 3.

[588] У прикладі, наведеному на фіг. 10N(ii), кадр 249M включає в себе два сигнали на векторній основі (за умови, що `ChannelType` 269 дорівнює 1 у полях 154 і 154B CSID) і порожній (за умови, що `ChannelType` 269 дорівнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 250M `HOAconfig`, пристрій 24 аудіодекодування може визначити, що 16 елементів вектора V кодуються. Тому кожен `VVectorData` 156 і 156B включає в себе 16 елементів вектора, кожний з яких однорідно квантується 8 бітами. Як зазначено вище, кількість і індекси кодованих елементів `VVectorData` задаються параметром `CodedVVecLength=0`. Крім того, як зазначено однією зірочкою (*), схема кодування сигналізується за допомогою `NbitsQ=5` у полі CSID для відповідного транспортного каналу.

[589] У кадрі 249N, поля 154 і 154B CSID такі ж, як у кадрі 249M, тоді як поле 154C CSID кадру 249F переключається на `ChannelType`, що дорівнює одиниці. Таким чином, поле 154C CSID кадру 249B включає в себе `Cbflag` 267, `Pflag` 267 (що вказує хатфманівське кодування) і `Nbits` 261 (що дорівнює дванадцяти). У результаті, кадр 249F включає в себе третє поле 156C `VVectorData`, що включає в себе 16 елементів вектора V , кожний з яких однорідно квантується 12 бітами і кодується по Хаффману. Як згадано вище, кількість і індекси кодованих елементів `VVectorData` задаються параметром `CodedVVecLength=0`, тоді як схема кодування Хаффмана сигналізується за допомогою `NbitsQ=12`, `CbFlag=0` і `Pflag=0` у полі 154C CSID для цього конкретного транспортного каналу (наприклад, транспортного каналу № 3).

[590] Приклади, наведені на фіг. 10O(i) і 10O(ii), ілюструють другий приклад бітового потоку 248O і супутньої частини 250O конфігурації HOA, згенерованої згідно з випадком 3 у вищенаведеному псевдокоді. У прикладі, наведеному на фіг. 10O(i), частина 250O `HOAconfig` включає в себе елемент 256 синтаксису `CodedVVecLength`, заданий для указання, що кодовані всі елементи вектора V , за винятком елементів, зазначених в елементі синтаксису `ContAddAmbHoaChan` (який передбачається рівним одиниці в цьому прикладі). Частина 250O `HOAconfig` також включає в себе елемент 255 синтаксису `SpatialInterpolationMethod`, заданий для указання, що інтерполяційною функцією просторово-часової інтерполяції є піднятий косинус. Частина 250O `HOAconfig`, крім того, включає в себе `CodedSpatialInterpolationTime` 254, заданий для указання інтерпольованої тривалості вибірки, що дорівнює 256.

[591] Частина 2500 HOAconfig додатково включає в себе елемент 150 синтаксису MinAmbHoaOrder заданий для вказання, що порядок MinimumHOA контенту HOA оточення дорівнює одиниці, де пристрій 24 аудіодекодування може виводити елемент синтаксису MinNumOfCoeffsForAmbHOA рівним $(1+1)^2$ або чотирьом. Пристрій 24 аудіодекодування також може виводити елемент синтаксису MaxNoOfAddActiveAmbCoeffs як заданий рівним різниці між елементом синтаксису NumOfHoaCoeff і MinNumOfCoeffsForAmbHOA, що передбачається в цьому прикладі рівним 16-4 або 12. Пристрій 24 аудіодекодування також може виводити елемент синтаксису AmbAssignBits як заданий рівним $\text{ceil}(\log_2(\text{MaxNoOfAddActiveAmbCoeffs})) = \text{ceil}(\log_2(12)) = 4$. Частина 2500 HOAconfig включає в себе елемент 152 синтаксису HoaOrder, заданий для вказання порядку HOA контенту рівним трьом (або, іншими словами, $N=3$), де пристрій 24 аудіодекодування може виводити NumOfHoaCoeffs рівним $(N+1)^2$ або 16.

[592] Як показано далі в прикладі, наведеному на фіг. 100(i), частина 2480 включає в себе кадр аудіо USAC-3D, у якому два кадри 2490 і 249P HOA зберігаються в корисному навантаженні розширення USAC, за умови, що два кадри аудіо зберігаються в одному кадрі USAC-3D, коли дозволене дублювання спектральної смуги (SBR). Пристрій 24 аудіодекодування може виводити кількість адаптивних транспортних каналів як функцію елемента синтаксису numHOATransportChannels і елемента синтаксису MinNumOfCoeffsForAmbHOA. У нижченаведених прикладах, передбачається, що елемент синтаксису numHOATransportChannels дорівнює 7 і елемент синтаксису MinNumOfCoeffsForAmbHOA дорівнює чотирьом, де кількість адаптивних транспортних каналів дорівнює елементу синтаксису numHOATransportChannels мінус елемент синтаксису MinNumOfCoeffsForAmbHOA (або трьом).

[593] Фіг. 100(ii) більш детально демонструє кадри 2490 і 249P. Як показано в прикладі, наведеному на фіг. 100(ii), кадр 2490 включає в себе поля 154-154C CSID і поле 156 VVectorData. Поле 154 CSID включає в себе CodedAmbCoeffIdx 246, AmbCoeffIdxTransition 247 (де подвійна зірочка (**)) вказує, що, для адаптивного транспортного каналу № 1, внутрішній стан декодера передбачається тут рівним AmbCoeffIdxTransitionState=2, у результаті чого бітове поле CodedAmbCoeffIdx сигналізується або інакше вказується в бітовому потоці) і ChannelType 269 (який дорівнює двом, що сигналізує, що відповідне корисне навантаження є додатковим коефіцієнтом HOA оточення). Пристрій 24 аудіодекодування може виводити AmbCoeffIdx рівним CodedAmbCoeffIdx+1+MinNumOfCoeffsForAmbHOA або 5 у цьому прикладі. Поле 154B CSID включає в себе unitC 267, bb 266 і ba 265 разом з ChannelType 269, кожне з яких задане рівним відповідним значенням 01, 1, 0 і 01, показаним у прикладі, наведеному на фіг. 100(ii). Поле 154C CSID включає в себе поле 269 ChannelType, що має значення 3.

[594] У прикладі, наведеному на фіг. 100(ii), кадр 2490 включає в себе одиничний сигнал на векторній основі (за умови, що ChannelType 269 дорівнює 1 у полях 154C CSID) і порожній (за умови, що ChannelType 269 дорівнює 3 у полях 154C CSID). З урахуванням вищевказаної частини 2500 HOAconfig, пристрій 24 аудіодекодування може визначити, що 16 мінус один, зазначений елементом синтаксису ContAddAmbHoaChan (наприклад, коли елемент вектора, пов'язаний з індексом 6, вказується як елемент синтаксису ContAddAmbHoaChan), або 15 елементів вектора V кодуються. Тому VVectorData 156 включає в себе 15 елементів вектора, кожний з яких однорідно квантується 8 бітами. Як зазначено вище, кількість і індекси кодованих елементів VVectorData задаються параметром CodedVVecLength=0. Крім того, як зазначено вище, схема кодування сигналізується за допомогою NbitsQ=5 у полі CSID для відповідного транспортного каналу.

[595] У кадрі 249P, поле 154 CSID включає в себе AmbCoeffIdxTransition 247, що вказує, що переходу не відбулося, і, таким чином, CodedAmbCoeffIdx 246 можна вивести з попереднього кадру і не потрібно повторно сигналізувати або інакше вказувати. Поля 154B і 154C CSID кадру 249P такі ж, як для кадру 2490, і, таким чином, як і кадр 2490, кадр 249P включає в себе одиничне поле 156 VVectorData, що включає в себе 15 елементів вектора, кожний з яких однорідно квантується 8 бітами.

[596] На фіг. 11A-11G показані блок-схеми, що демонструють, більш детально, різні блоки пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. На фіг. 11A показана блок-схема, що демонструє, більш детально, блок 72 витягання пристрою 24 аудіодекодування. Як показано в прикладі, наведеному на фіг. 11A, блок 72 витягання може включати в себе блок 270 виділення режиму, блок 272 конфігурування режиму ("блок 272 конфігурування режиму"), і блок 274 конфігурованого витягання.

[597] Блок 270 виділення режиму може представляти блок, виконаний з можливістю виділення вищезгаданого елемента синтаксису, що вказує режим кодування (наприклад,

елемент синтаксису ChannelType, показаний у прикладі, наведеному на фіг. 10E), використовуваний для кодування коефіцієнтів 11 НОА, для формування бітового потоку 21. Блок 270 виділення режиму може передавати визначений елемент синтаксису на блок 272 конфігурування режиму. Блок 272 конфігурування режиму може представляти блок, виконаний з

5 можливістю конфігурування блока 274 конфігурованого витягання на основі виділеного елемента синтаксису. Блок 272 конфігурування режиму може конфігурувати блок 274 конфігурованого витягання для витягання кодованого представлення коефіцієнтів 11 НОА на основі напрямку з бітового потоку 21 або витягання кодованого представлення коефіцієнтів 11 НОА на векторній основі з бітового потоку 21 на основі виділеного елемента синтаксису.

10 [598] Коли кодування на основі напрямку здійснено, блок 274 конфігурованого витягання може витягати версію на основі напрямку коефіцієнтів 11 НОА і елементи синтаксису, пов'язані з цією кодовою версією (яка позначається як інформація 91 на основі напрямку в прикладі, наведеному на фіг. 11A). Ця інформація 91 на основі напрямку може включати в себе інформацію 253 направленості, показану в прикладі, наведеному на фіг. 10D, і

15 SideChannelInfoData на основі напрямку, показаний у прикладі, наведеному на фіг. 10E, заданий за допомогою ChannelType, що дорівнює нулю.

[599] Коли елемент синтаксису вказує, що коефіцієнти 11 НОА були кодовані з використанням синтезу на векторній основі (наприклад, коли елемент синтаксису ChannelType дорівнює одиниці), блок 274 конфігурованого витягання може витягати кодовані вектори 57 переднього плану $V[k]$, кодовані коефіцієнти 59 НОА оточення і кодовані сигнали 59 pFG. Блок 274 конфігурованого витягання також може, визначивши, що елемент синтаксису вказує, що

20 коефіцієнти 11 НОА були кодовані з використанням синтезу на векторній основі, витягати елемент 254 синтаксису CodedSpatialInterpolationTime і елемент 255 синтаксису SpatialInterpolationMethod з бітового потоку 21, передаючи ці елементи 254 і 255 синтаксису на

25 блок 76 просторово-часової інтерполяції.

[600] На фіг. 11B показана блок-схема, що демонструє, більш детально, блок 74 квантування пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. Блок 74 квантування може представляти блок, виконаний з можливістю діяти в режимі, зворотному

30 блоку 52 квантування, показаному в прикладі, наведеному на фіг. 4, щоб ентропійно декодувати і деквантувати кодовані вектори 57 переднього плану $V[k]$ і, таким чином, генерувати приведені вектори 55_k переднього плану $V[k]$. Блок скалярного/ентропійного деквантування 984 може включати в себе блок 276 декодування категорії/залишку, блок 278 прогнозування і блок 280 однорідного деквантування.

[601] Блок 276 декодування категорії/залишку може представляти блок, виконаний з

35 можливістю здійснення хаффманівського декодування відносно кодованих векторів 57 переднього плану $V[k]$ з використанням таблиці Хаффмана, ідентифікованої інформацією 241 таблиці Хаффмана (яка, як згадано вище, виражається як елемент синтаксису в бітовому потоці 21). Блок 276 декодування категорії/залишку може виводити квантовані вектори переднього плану $V[k]$ на блок 278 прогнозування. Блок 278 прогнозування може представляти блок, виконаний з

40 можливістю здійснення прогнозування відносно квантованих векторів переднього плану $V[k]$ на основі режиму 237 прогнозування, виводячи доповнені квантовані вектори переднього плану $V[k]$ на блок 280 однорідного деквантування. Блок 280 однорідного деквантування може представляти блок, виконаний з можливістю здійснення деквантування відносно доповнених квантованих векторів переднього плану $V[k]$ на основі значення 233 nbits,

45 виводячи приведені вектори 55_k переднього плану $V[k]$.

[602] На фіг. 11C показана блок-схема, що демонструє, більш детально, блок 80 психоакустичного декодування пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. Як згадано вище, блок 80 психоакустичного декодування може діяти в режимі, зворотному блоку 40 психоакустичного аудіокодування, показаному в прикладі,

50 наведеному на фіг. 4, щоб декодувати кодовані коефіцієнти 59 НОА оточення і кодовані сигнали 61 pFG і, таким чином, генерувати коефіцієнти 47' НОА оточення з компенсацією енергії й інтерпольованих сигналів 49' pFG (які також можуть іменуватися інтерпольованими аудіооб'єктами 49' pFG). Блок 80 психоакустичного декодування може передавати коефіцієнти 47' НОА оточення з компенсацією енергії на блок 82 формування коефіцієнтів НОА і сигнали 49' pFG на блок 84 переупорядковування. Блок 80 психоакустичного декодування може включати в себе множину аудіодекодерів 80-80N, за аналогією з блоком 40 психоакустичного аудіокодування. Аудіодекодери 80-80N можуть бути реалізовані або інакше включені в блок 40 психоакустичного аудіокодування в достатній кількості для підтримки, як згадано вище,

55 одночасного декодування кожного каналу коефіцієнтів 47' НОА фону і кожного сигналу із сигналів 49' pFG.

60

[603] На фіг. 11D показана блок-схема, що демонструє, більш детально, блок 84 переупорядкування пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. Блок 84 переупорядкування може представляти блок, виконаний з можливістю діяти аналогічно зворотному, описаному вище відносно блока 34 переупорядкування. Блок 84 переупорядкування може включати в себе блок 282 переупорядкування векторів, що може представляти блок, виконаний з можливістю прийому елементів 205 синтаксису, що указують вихідний порядок компонентів переднього плану коефіцієнтів 11 НОА. Блок 72 витягання може виділяти ці елементи 205 синтаксису з бітового потоку 21 і передавати елемент 205 синтаксису на блок 84 переупорядкування. Блок 282 переупорядкування векторів може, на основі цих елементів 205 синтаксису переупорядкування, переупорядковувати інтерпольовані сигнали 49' nFG і приведені вектори 55_k' переднього плану V[k] для генерації переупорядкованих сигналів 49" nFG і переупорядкованих векторів 55_k" переднього плану V[k]. Блок 84 переупорядкування може виводити переупорядковані сигнали 49" nFG на блок 78 формування переднього плану і переупорядковані вектори 55_k' переднього плану V[k] на блок 76 просторово-часової інтерполяції.

[604] На фіг. 11E показана блок-схема, що демонструє, більш детально, блок 76 просторово-часової інтерполяції пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. Блок 76 просторово-часової інтерполяції може діяти аналогічно описаному вище відносно блока 50 просторово-часової інтерполяції. Блок 76 просторово-часової інтерполяції може включати в себе блок 284 інтерполяції V, що може представляти блок, виконаний з можливістю прийому переупорядкованих векторів 55_k' переднього плану V[k] і здійснення просторово-часової інтерполяції відносно переупорядкованих векторів 55_k' переднього плану V[k] і переупорядкованих векторів 55_{k-1}' переднього плану V[k-1] для генерації інтерпольованих векторів 55_k" переднього плану V[k]. Блок 284 інтерполяції V може здійснювати інтерполяцію на основі елемента 254 синтаксису CodedSpatialInterpolationTime і елемента 255 синтаксису SpatialInterpolationMethod. У деяких прикладах, блок 285 інтерполяції V може інтерполювати вектори V протягом тривалості, зазначеної елементом 254 синтаксису CodedSpatialInterpolationTime, з використанням типу інтерполяції, ідентифікованого елементом 255 синтаксису SpatialInterpolationMethod. Блок 76 просторово-часової інтерполяції може пересилати інтерпольовані вектори 55_k" переднього плану V[k] на блок 78 формування переднього плану.

[605] На фіг. 11F показана блок-схема, що демонструє, більш детально, блок 78 формування переднього плану пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. Блок 78 формування переднього плану може включати в себе блок 286 множення, що може представляти блок, виконаний з можливістю здійснення матричного множення відносно інтерпольованих векторів 55_k" переднього плану V[k] і переупорядкованих сигналів 49" nFG для генерації коефіцієнтів 65 НОА переднього плану.

[606] На фіг. 11G показана блок-схема, що демонструє, більш детально, блок 82 формування коефіцієнтів НОА пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. Блок 82 формування коефіцієнтів НОА може включати в себе блок 288 додавання, що може представляти блок, виконаний з можливістю додавання коефіцієнтів 65 НОА переднього плану до каналів 47' НОА оточення для одержання коефіцієнтів 11' НОА.

[607] На фіг. 12 показана схема, що демонструє приклад аудіо-екосистеми, яка може здійснювати різні аспекти методів, описаних у цьому винаході. Як показано на фіг. 12, аудіо-екосистема 300 може включати в себе захоплення 301, редагування 302, кодування 303, передачу 304 і відтворення 305.

[608] Захоплення 301 може представляти методи аудіо-екосистеми 300, де захоплюється аудіоконтент. Приклади захоплення 301 включають в себе, але без обмеження, запис звуку (наприклад, живого звуку), генерацію аудіосигналу (наприклад, аудіооб'єктів, створення шумових ефектів, синтез звуку, імітації) і ін. У деяких прикладах, звук може записуватися на концертах, спортивних змаганнях і при проведенні спостереження. У деяких прикладах, аудіосигнал може генеруватися при здійсненні імітацій і авторингу/міксування (наприклад, фільмів, ігор). Аудіооб'єкти можуть використовуватися в Голлівуді (наприклад, студіях IMAX). У деяких прикладах, захоплення 301 може здійснюватися творцем контенту, наприклад творцем 12 контенту, показаним на фіг. 3.

[609] Редагування 302 може представляти методи аудіо-екосистеми 300, де аудіоконтент редагується і/або модифікується. У порядку одного прикладу, аудіоконтент може редагуватися шляхом об'єднання множинних блоків аудіоконтенту в єдиний блок аудіоконтенту. У порядку іншого прикладу, аудіоконтент може редагуватися шляхом регулювання фактичного аудіоконтенту (наприклад, регулювання рівнів одного або більше частотних компонентів

аудіоконтенту). У деяких прикладах, редагування 302 може здійснюватися системою аудіоредагування, наприклад системою 18 аудіоредагування, показаною на фіг. 3. У деяких прикладах, редагування 302 можна здійснювати на мобільному пристрої, наприклад одному або більше з мобільних пристроїв, представлених на фіг. 29.

5 [610] Кодування, 303 може представляти методи аудіо-екосистеми 300, де аудіоконтент кодується в представлення аудіоконтенту. У деяких прикладах, представленням аудіоконтенту може бути бітовий потік, наприклад бітовий потік 21, показаний на фіг. 3. У деяких прикладах, кодування 302 можна здійснювати пристроєм аудіокодування, наприклад пристроєм 20 аудіокодування, показаним на фіг. 3.

10 [611] Передача 304 може представляти елементи аудіо-екосистеми 300, де аудіоконтент переноситься від творця контенту до споживача контенту. У деяких прикладах, аудіоконтент може переноситися в реальному часі або майже в реальному часі. Наприклад, аудіоконтент може передаватися в потоці споживачу контенту. У деяких прикладах, аудіоконтент може переноситися шляхом кодування аудіоконтенту в носії, наприклад зчитуваному комп'ютером носій даних. Наприклад, аудіоконтент може зберігатися на диску, приводі і ін. (наприклад, диску Blu-ray, карті пам'яті, твердому диску і т. д.).

15 [612] Відтворення 305 може представляти методи аудіо-екосистеми 300, де аудіоконтент рендеризується і відтворюється споживачу контенту. У деяких прикладах, відтворення 305 може включати в себе рендеризацію 3D звукового поля на основі одного або більше аспектів середовища відтворення. Іншими словами, відтворення 305 може базуватися на локальному акустичному ландшафті.

20 [613] На фіг. 13 показана схема, що більш детально демонструє один приклад аудіо-екосистеми, показаної на фіг. 12. Як показано на фіг. 13, аудіо-екосистема 300 може включати в себе аудіоконтент 308, кіностудії 310, музичні студії 311, студії 312 запису звукового супроводу ігор, аудіоконтент 313 на основі каналів, машини 314 кодування, ігрові аудіосистеми 315, машини 316 ігрового аудіокодування/рендеризації і системи 317 доставки. Приклад студії 312 запису звукового супроводу ігор представлений на фіг. 26. Деякі приклади машини 316 ігрового аудіокодування/рендеризації представлені на фіг. 27.

30 [614] Як показано на фіг. 13, кіностудії 310, музичні студії 311 і студії 312 запису звукового супроводу ігор можуть приймати аудіоконтент 308. У ряді випадків, аудіоконтент 308 може представляти вихід захоплення 301, показаного на фіг. 12. Кіностудії 310 можуть виводити аудіоконтент 313 на основі каналів (наприклад, у 2.0, 5.1 і 7.1), наприклад, з використанням робочої станції цифрового аудіо (DAW). Музичні студії 310 можуть виводити аудіоконтент 313 на основі каналів (наприклад, у 2.0 і 5.1), наприклад, з використанням DAW. У будь-якому випадку, машини 314 кодування можуть приймати і кодувати аудіоконтент 313 на основі каналів за допомогою одного або більше кодеків (наприклад, AAC, AC3, Dolby True HD, Dolby Digital Plus і DTS Master Audio) для виведення системами 317 доставки. Таким чином, машини 314 кодування можуть бути прикладом кодування 303, показаного на фіг. 12. Студії 312 запису звукового супроводу ігор можуть виводити одну або більше ігрових аудіосистем 315, наприклад, з використанням DAW. Машини 316 ігрового аудіокодування/рендеризації можуть кодувати і/або рендеризувати аудіосистеми 315 в аудіоконтент на основі каналів для виведення системами 317 доставки. У деяких прикладах, вихід кіностудій 310, музичних студій 311 і студій 312 запису звукового супроводу ігор може представляти вихід редагування 302, показаного на фіг. 12. У деяких прикладах, вихід машин 314 кодування і/або машин 316 ігрового аудіокодування/рендеризації може переноситися на системи 317 доставки за допомогою методів передачі 304, показаної на фіг. 12.

45 [615] На фіг. 14 показана схема, що більш детально демонструє інший приклад аудіо-екосистеми, показаної на фіг. 12. Як показано на фіг. 14, аудіо-екосистема 300B може включати в себе аудіооб'єкти 319 широкомовного запису, професійні аудіосистеми 320, споживче захоплення 322 на пристрої, аудіоформат 323 НОА, рендеризацію 324 на пристрої, споживчі аудіо, TV і приладдя 325 і автомобільні аудіосистеми 326.

50 [616] Як показано на фіг. 14, аудіооб'єкти 319 широкомовного запису, професійні аудіосистеми 320 і споживче захоплення 322 на пристрої можуть кодувати свій вихідний сигнал з використанням аудіоформату 323 НОА. Таким чином, аудіоконтент можна кодувати з використанням аудіоформату 323 НОА у єдине представлення, що може відтворюватися з використанням рендеризації 324 на пристрої, споживчих аудіо, TV, і приладь 325 і автомобільних аудіосистем 326. Іншими словами, єдине представлення аудіоконтенту може відтворюватися на універсальній системі відтворення аудіосигналу (тобто, на відміну від потребуючої конкретної конфігурації, наприклад, 5.1, 7.1 і т. д.).

[617] На фіг. 15A і 15B показані схеми, що більш детально демонструють інші приклади аудіо-екосистеми, показаної на фіг. 12. Як показано на фіг. 15A, аудіо-екосистема 300C може включати в себе елементи 331 захоплення й елементи 336 відтворення. Елементи 331 захоплення можуть включати в себе дротові і/або бездротові пристрої 332 захоплення (наприклад, власні мікрофони), захоплення 334 об'ємного звуку на пристрої і мобільні пристрої 335 (наприклад, смартфони і планшети). У деяких прикладах, дротові і/або бездротові пристрої 332 захоплення можуть бути підключені до мобільного пристрою 335 по дротовому(их) і/або бездротовому(их) каналу(ах) 333 зв'язку.

[618] Відповідно до одного або більше методів цього винаходу, мобільний пристрій 335 можна використовувати для захоплення звукового поля. Наприклад, мобільний пристрій 335 може захоплювати звукове поле через дротові і/або бездротові пристрої 332 захоплення і/або захоплення 334 об'ємного звуку на пристрої (наприклад, множина мікрофонів, вбудованих у мобільний пристрій 335). Потім мобільний пристрій 335 може кодувати захоплене звукове поле в НОА 337 для відтворення одним або більше з елементів 336 відтворення. Наприклад, користувач мобільного пристрою 335 може записувати (захоплювати звукове поле) живу подію (наприклад, зустрічі, конфігурації, гри, концерту і т. д.) і кодувати запис у НОА.

[619] Мобільний пристрій 335 також може використовувати один або більше з елементів 336 відтворення для відтворення звукового поля, кодованого в НОА. Наприклад, мобільний пристрій 335 може декодувати звукове поле, кодоване в НОА, і виводити сигнал на один або більше з елементів 336 відтворення, що призначає одному або більше з елементів 336 відтворення відтворювати звукове поле. У порядку одного прикладу, мобільний пристрій 335 може використовувати дротові і/або бездротові канали 338 зв'язку для виведення сигналу на один або більше гучномовців (наприклад, решітки гучномовців, звукові панелі і т. д.). У порядку іншого прикладу, мобільний пристрій 335 може використовувати рішення 339 стикування для виведення сигналу на одну або більше стикувальних станцій і/або один або більше пристикованих гучномовців (наприклад, звукові системи в інтелектуальних автомобілях і/або будинках). У порядку іншого прикладу, мобільний пристрій 335 може використовувати рендеризацію 340 через головні телефони для виведення сигналу на набір головних телефонів, наприклад, для створення реалістичного стереофонічного звуку.

[620] У деяких прикладах, конкретний мобільний пристрій 335 може захоплювати 3D звукове поле і відтворювати те ж 3D звукове поле в більш пізній час. У деяких прикладах, мобільний пристрій 335 може захоплювати 3D звукове поле, кодувати 3D звукове поле в НОА і передавати кодоване 3D звукове поле на один або більше інших пристроїв (наприклад, інші мобільні пристрої і/або інші немобільні пристрої) для відтворення.

[621] Як показано на фіг. 15B, аудіо-екосистема 300D може включати в себе аудіоконтент 343, ігрові студії 344, кодований аудіоконтент 345, машини 346 рендеризації і системи 347 доставки. У деяких прикладах, ігрові студії 344 можуть включати в себе одну або більше DAW, що можуть підтримувати редагування сигналів НОА. Наприклад, одна або більше DAW можуть включати в себе змінні плати НОА і/або інструменти, що можуть бути виконані з можливістю діяти з (наприклад, працювати з) однією або більше ігровими аудіосистемами. У деяких прикладах, ігрові студії 344 можуть виводити нові формати системи, які підтримують НОА. У будь-якому випадку, ігрові студії 344 можуть виводити кодований аудіоконтент 345 на машини 346 рендеризації, які можуть рендеризувати звукове поле для відтворення системами 347 доставки.

[622] На фіг. 16 показана схема, що демонструє приклад пристрою аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході. Як показано на фіг. 16, аудіо-екосистема 300E може включати в себе вихідний 3D аудіоконтент 351, кодер 352, бітовий потік 353, декодер 354, рендеризатор 355 і елементи 356 відтворення. Як додатково проілюстровано на фіг. 16, кодер 352 може включати в себе аналіз і розкладання 357 звукового поля, витягання 358 фону, визначення 359 помітності фону, аудіокодування 360, витягання 361 передньопланового/характерного аудіосигналу й аудіокодування 362. У деяких прикладах, кодер 352 може бути виконаний з можливістю здійснення операцій, аналогічних виконуваним пристроєм 20 аудіокодування, показаним на фіг. 3 і 4. У деяких прикладах, аналіз і розкладання 357 звукового поля можуть бути виконані з можливістю здійснення операцій, аналогічних виконуваним блоком 44 аналізу звукового поля, показаним на фіг. 4. У деяких прикладах, витягання 358 фону і визначення 359 помітності фону можуть бути виконані з можливістю здійснення операцій, аналогічних виконуваним блоком 48 вибору BG, показаним на фіг. 4. У деяких прикладах, аудіокодування 360 і аудіокодування 362 можуть бути виконані з можливістю здійснення операцій, аналогічних виконуваним блоком 40 психоакустичного аудіокодера, показаним на фіг. 4. У деяких прикладах, витягання 361 передньопланового/характерного

аудіосигналу може бути виконане з можливістю здійснення операцій, аналогічних виконуваним блоком 36 вибору переднього плану, показаним на фіг. 4.

[623] У деяких прикладах, витягання 361 передньопланового/характерного аудіосигналу може аналізувати аудіоконтент, що відповідає кадру 390 відео, показаному на фіг. 33. Наприклад, витягання 361 передньопланового/характерного аудіосигналу може визначити, що аудіоконтент, що відповідає областям 391A-391C, є аудіосигналом переднього плану.

[624] Як показано на фіг. 16, кодер 352 може бути виконаний з можливістю кодування вихідного контенту 351, що може мати бітову швидкість 25-75 Мбіт/с, у бітовий потік 353, що може мати бітову швидкість 256 кбіт/с - 1,2 Мбіт/с. На фіг. 17 показана схема, що більш детально демонструє один приклад пристрою аудіокодування, показаного на фіг. 16.

[625] На фіг. 18 показана схема, що демонструє приклад пристрою аудіодекодування, який може здійснювати різні аспекти методів, описаних у цьому винаході. Як показано на фіг. 18, аудіо-екосистема 300E може включати в себе вихідний 3D аудіоконтент 351, кодер 352, бітовий потік 353, декодер 354, рендеризатор 355 і елементи 356 відтворення. Як додатково проілюстровано на фіг. 16, декодер 354 може включати в себе аудіодекодер 363, аудіодекодер 364, реконструкцію 365 переднього плану і мікшування 366. У деяких прикладах, декодер 354 може бути виконаний з можливістю здійснення операцій, аналогічних виконуваним пристроєм 24 аудіодекодування, представленим на фіг. 3 і 5. У деяких прикладах, аудіодекодер 363, аудіодекодер 364 може бути виконаний з можливістю здійснення операцій, аналогічних виконуваним блоком 80 психоакустичного декодування, представленим на фіг. 5. У деяких прикладах, реконструкція 365 переднього плану може бути виконана з можливістю здійснення операцій, аналогічних виконуваним блоком 78 формування переднього плану, представленим на фіг. 5.

[626] Як показано на фіг. 16, декодер 354 може бути виконаний з можливістю прийому і декодування бітового потоку 353 і виведення результуючого реконструйованого 3D звукового поля на рендеризатор 355, який може потім призначати одному або більше з елементів 356 відтворення виводити представлення вихідного 3D контенту 351. На фіг. 19 показана схема, що більш детально демонструє один приклад пристрою аудіодекодування, показаного на фіг. 18.

[627] На фіг. 20A-20G показані схеми, що демонструють приклад пристроїв захоплення аудіосигналу, які можуть здійснювати різні аспекти методів, описаних у цьому винаході. Фіг. 20A демонструє власний мікрофон 370, що може включати в себе множину мікрофонів, які спільно виконані з можливістю запису 3D звукового поля. У деяких прикладах, множина мікрофонів власного мікрофона 370 може розташовуватися на поверхні, по суті, сферичної кулі радіусом приблизно 4 см. У деяких прикладах, пристрій 20 аудіокодування може бути вбудований у власний мікрофон для виведення бітового потоку 17 безпосередньо з мікрофона 370.

[628] Фіг. 20B демонструє телевізійний автомобіль 372, що може бути виконаний з можливістю прийому сигналу від одного або більше мікрофонів, наприклад одного або більше власних мікрофонів 370. Телевізійний автомобіль 372 також може включати в себе аудіокодер, наприклад аудіокодер 20, показаний на фіг. 3.

[629] Фіг. 20C-20E ілюструють мобільний пристрій 374, який може включати в себе множину мікрофонів, що спільно виконані з можливістю запису 3D звукового поля. Іншими словами, множина мікрофонів може мати рознесення по X, Y, Z. У деяких прикладах, мобільний пристрій 374 може включати в себе мікрофон 376, що може повертатися для забезпечення рознесення по X, Y, Z відносно одного або більше інших мікрофонів мобільного пристрою 374. Мобільний пристрій 374 також може включати в себе аудіокодер, наприклад аудіокодер 20, показаний на фіг. 3.

[630] Фіг. 20F демонструє зносостійкий пристрій 378 захоплення відео, що може бути виконаний з можливістю запису 3D звукового поля. У деяких прикладах, зносостійкий пристрій 378 захоплення відео може бути приєднаний до шолома користувача, залученого в діяльність. Наприклад, зносостійкий пристрій 378 захоплення відео може бути приєднаний до шолома користувача, що бере участь у сплавланні по бурхливій ріці. Таким чином, зносостійкий пристрій 378 захоплення відео може захоплювати 3D звукове поле, яке представляє дію навколо користувача (наприклад, бурління води позаду користувача, мова іншого рафтера перед користувачем і т. д.).

[631] Фіг. 20G демонструє мобільний пристрій 380 з додатковими можливостями, який може бути виконаний з можливістю запису 3D звукового поля. У деяких прикладах, мобільний пристрій 380 може бути аналогічним мобільному пристрою 335, показаному на фіг. 15, з додаванням одного або більше приладів. Наприклад, власний мікрофон може бути приєднаний до мобільного пристрою 335, показаного на фіг. 15, для формування мобільного пристрою 380 з додатковими можливостями. Таким чином, мобільний пристрій 380 з додатковими

можливостями може захоплювати версію більш високої якості 3D звукового поля, а не просто використовувати компоненти захоплення звуку, вбудовані в мобільний пристрій 380 з додатковими можливостями.

[632] На фіг. 21A-21E показані схеми, що демонструють приклад пристроїв відтворення аудіосигналу, які можуть здійснювати різні аспекти методів, описаних у цьому винаході. Фіг. 21A і 21B демонструє множину гучномовців 382 і звукових панелей 384. Відповідно до одного або більше методів цього винаходу, гучномовці 382 і/або звукові панелі 384 можуть розташовуватися в будь-якій довільній конфігурації, яка дозволяє відтворювати 3D звукове поле. Фіг. 21C-21E ілюструють множину пристроїв 386-386C відтворення у вигляді головних телефонів. Пристрої 386-386C відтворення у вигляді головних телефонів можуть бути підключені до декодера за допомогою дротового або бездротового з'єднання. Відповідно до одного або більше методів цього винаходу, одиничне універсальне представлення звукового поля може використовуватися для рендеризації звукового поля на будь-якій комбінації гучномовців 382, звукових панелей 384 і пристроїв 386-386C відтворення у вигляді головних телефонів.

[633] На фіг. 22A-22H показані схеми, що демонструють приклад середовищ відтворення аудіосигналу відповідно до одного або більше методів, описаних в цьому винаході. Наприклад, фіг. 22A демонструє середовище відтворення з конфігурацією гучномовців 5.1, фіг. 22B демонструє середовище відтворення з конфігурацією гучномовців 2.0 (наприклад, стерео), фіг. 22C демонструє середовище відтворення з конфігурацією гучномовців 9.1 з передніми гучномовцями повної висоти, фіг. 22D і 22E ілюструють середовище відтворення з конфігурацією гучномовців 22.2, фіг. 22F демонструє середовище відтворення з конфігурацією гучномовців 16.0, фіг. 22G демонструє середовище відтворення з автомобільною конфігурацією гучномовців, і фіг. 22H демонструє мобільний пристрій з середовищем відтворення на основі навушників.

[634] Відповідно до одного або більше методів цього винаходу, одиничне універсальне представлення звукового поля може використовуватися для рендеризації звукового поля в будь-якому з середовищ відтворення, представлених на фіг. 22A-22H. Додатково, методи цього винаходу допускають рендеризацію для рендеризації звукового поля з універсального представлення для відтворення в середовищах відтворення, відмінних від представлених на фіг. 22A-22H. Наприклад, якщо розуміння конструкції перешкоджають правильному розміщенню гучномовців відповідно до середовища відтворення з конфігурацією гучномовців 7.1 (наприклад, якщо неможливо розмістити правий гучномовець оточення), методи цього винаходу допускають рендеризацію для компенсації з іншими 6 гучномовцями таким чином, що відтворення можна досягти в середовищі відтворення з конфігурацією гучномовців 6.1.

[635] Як показано на фіг. 23, користувач може дивитися спортивну гру при носінні головних телефонів 386. Відповідно до одного або більше методів цього винаходу, може захоплюватися 3D звукове поле спортивної гри (наприклад, один або більше власних мікрофонів можуть розташовуватися в і/або навколо бейсбольного стадіону, представленого на фіг. 24), коефіцієнти НОА, що відповідають 3D звуковому полю, можна одержувати і передавати на декодер, декодер може прийняти рішення реконструювати 3D звукове поле на основі коефіцієнтів НОА і виводити реконструйоване 3D звукове поле на рендеризатор, рендеризатор може одержувати указання відносно типу середовища відтворення (наприклад, головних телефонів) і рендеризувати реконструйоване 3D звукове поле в сигнали, які призначають головним телефонам виводити представлення 3D звукового поля спортивної гри. У деяких прикладах, рендеризатор може одержувати указання відносно типу середовища відтворення відповідно до методів, представлених на фіг. 25. Таким чином, рендеризатор може "адаптуватися" до різних положень гучномовців, типу чисел, розміру і також, в ідеалі, вирівнюватися для локального середовища.

[636] На фіг. 28 показана схема, що демонструє конфігурацію гучномовців, яку можна імітувати головними телефонами, відповідно до одного або більше методів, описаних в цьому винаході. Як показано на фіг. 28, методи цього винаходу можуть надавати можливість користувачу, що носить головні телефони 389, відчувати звукове поле, як яби звукове поле відтворювалося гучномовцями 388. Таким чином, користувач може слухати 3D звукове поле без виведення звуку на велику площу.

[637] На фіг. 30 показана схема, що демонструє кадр відео, пов'язаний з 3D звуковим полем, який може оброблятися відповідно до одного або більше методів, описаних в цьому винаході.

[638] На фіг. 31A-31M показані схеми, що демонструють графіки 400A-400M, де показані різні результати імітації здійснення синтетичної або записаної категоризації звукового поля відповідно до різних аспектів методів, описаних у цьому винаході. У прикладах, наведених на

фіг. 31A-31M, кожний із графіків 400A-400M включає в себе поріг 402, що позначається пунктирною лінією, і відповідний аудіооб'єкт 404A-404M (спільно "аудіооб'єкти 404") позначений штриховою лінією.

[639] Коли аудіооб'єкти 404 за допомогою аналізу, описаного вище відносно блока 26 аналізу контенту, визначаються такими, що знаходяться нижче порога 402, блок 26 аналізу контенту визначає, що відповідний один з аудіооб'єктів 404 представляє аудіооб'єкт, що був записаний. Як показано в прикладах, наведених на фіг. 31B, 31D-31H і 31J-31L, блок 26 аналізу контенту визначає, що аудіооб'єкти 404B, 404D-404H, 404J-404L нижче порога 402 (щонайменше +90 % часу і часто 100 % часу) і, таким чином, представляють записані аудіооб'єкти. Як показано в прикладах, наведених на фіг. 31A, 31C і 31I, блок 26 аналізу контенту визначає, що аудіооб'єкти 404A, 404C і 404I перевищують поріг 402 і, таким чином, представляють синтетичні аудіооб'єкти.

[640] У прикладі, наведеному на фіг. 31M, аудіооб'єкт 404M представляє змішаний синтетичний/записаний аудіооб'єкт, що має деякі синтетичні ділянки (наприклад, вище порога 402) і деякі синтетичні ділянки (наприклад, нижче порога 402). Блок 26 аналізу контенту в цьому випадку ідентифікує синтетичні і записані ділянки аудіооб'єкта 404M, у результаті чого пристрій 20 аудіокодування генерує бітовий потік 21 для включення кодованих аудіоданих на основі направленості і кодованих аудіоданих на векторній основі.

[641] На фіг. 32 показана схема, що демонструє графік 406 сингулярних значень з матриці S, розкладеної з коефіцієнтів амбіофонії більш високого порядку, відповідно до методів, описаних в цьому винаході. Як показано на фіг. 32, ненульових сингулярних значень, що мають великі значення, небагато. Блок 44 аналізу звукового поля, показаний на фіг. 4, може аналізувати ці сингулярні значення для визначення nFG компонентів переднього плану (або, іншими словами, переважних) (часто представлених векторами) переупорядкованих векторів 33' US[k] і переупорядкованих векторів 35' V[k].

[642] На фіг. 33A і 33B показані схеми, що демонструють відповідні графіки 410A і 410B, де показаний потенційний вплив переупорядкування при кодуванні векторів, що описують компоненти переднього плану звукового поля, відповідно до методів, описаних в цьому винаході. Графік 410A демонструє результат кодування щонайменше деяких з неупорядкованих (або, іншими словами, вихідних) векторів 33 US[k], тоді як графік 410B демонструє результат кодування відповідних упорядкованих векторів 33' US[k]. Верхня крива в кожному із графіків 410A і 410B демонструє помилку кодування, де, імовірно, існує тільки помітна помилка на графіку 410B на границях кадру. Відповідно, методи переупорядкування, описані в цьому винаході, можуть полегшувати або інакше сприяти кодуванню монофонічних аудіооб'єктів з використанням успадкованого аудіокодера.

[643] На фіг. 34 і 35 показані принципові схеми, що демонструють відмінності між ідентифікацією характерних аудіооб'єктів тільки на основі енергії і на основі направленості, відповідно до цього розкриття. У прикладі, наведеному на фіг. 34, вектори, що демонструють велику енергію, ідентифікуються як характерні аудіооб'єкти, незалежно від направленості. Як показано на фіг. 34, аудіооб'єкти, що розташовуються відповідно до більш високих значень енергії (відкладених по осі y), визначаються як "на передньому плані", незалежно від направленості (наприклад, представлені показниками направленості, відкладеними по осі x).

[644] Фіг. 35 демонструє ідентифікацію характерних аудіооб'єктів на основі направленості й енергії, наприклад, відповідно до методів, реалізованих блоком 44 аналізу звукового поля, показаним на фіг. 4. Як показано на фіг. 35, більш високі показники направленості відкладаються вліво по осі x і більш високі рівні енергії відкладаються вгору по осі y. У цьому прикладі, блок 44 аналізу звукового поля може визначити, що характерні аудіооб'єкти (наприклад, що знаходяться "на передньому плані") пов'язані з векторними даними, відкладеними відносно верхнього лівого кута графіка. У порядку одного прикладу, блок 44 аналізу звукового поля може визначити, що ті вектори, які відкладаються у верхньому лівому квадранті графіка, пов'язані з характерними аудіооб'єктами.

[645] На фіг. 36A-36F показані схеми, що демонструють проекції щонайменше частини розкладеної версії коефіцієнтів сферичних гармонік у просторову область для здійснення інтерполяції відповідно до різних аспектів методів, описаних у цьому винаході. На фіг. 36A показана схема, що демонструє проекцію одного або більше з векторів 35 V[k] на сферу 412. У прикладі, наведеному на фіг. 36A, кожне число ідентифікує окремий коефіцієнт сферичної гармоніки, проєктований на сферу (можливо, пов'язану з одним рядком і/або стовпцем матриці 19' V). Різні кольори вказують напрямок характерного аудіокомпонента, де більш світлий (і поступово більш темний) колір позначає первинний напрямок характерного компонента. Блок 50 просторово-часової інтерполяції пристрою 20 аудіокодування, показаного в прикладі,

наведеному на фіг. 4, може здійснювати просторово-часову інтерполяцію між кожною з червоних точок для генерації сфери, показаної в прикладі, наведеному на фіг. 36A.

[646] На фіг. 36B показана схема, що демонструє проекцію одного або більше з векторів 35 $V[k]$ на пелюсток. Блок 50 просторово-часової інтерполяції може проектувати один рядок і/або стовпець векторів 35 $V[k]$ або множинні рядки і/або стовпці векторів 35 $V[k]$ для генерації пелюстка 414, показаного в прикладі, наведеному на фіг. 36B.

[647] На фіг. 36C показана схема, що демонструє поперечний переріз проекції одного або більше векторів з одного або більше з векторів 35 $V[k]$ на сферу, наприклад сферу 412, показану в прикладі, наведеному на фіг. 36.

[648] На фіг. 36D-36G показані приклади знімків часу (через 1 кадр близько 20 мілісекунд), коли різні джерела звуку (бджола, вертоліт, електронна музика і люди на стадіоні) можна проілюструвати в тривимірному просторі.

[649] Методи, описані в цьому винаході, дозволяють ідентифікувати представлення цих різних джерел звуку і представляти з використанням одиничного вектора $US[k]$ і одиничного вектора $V[k]$. Часова мінливість джерел звуку представлена у векторі $US[k]$, тоді як просторовий розподіл кожного джерела звуку представлений одиничним вектором $V[k]$. Один вектор $V[k]$ може представляти ширину, положення і розмір джерела звуку. Крім того, одиничний вектор $V[k]$ може бути представлений як лінійна комбінація сферичних гармонічних базисних функцій. У графіках, показаних на фіг. 36D-36G, представлення джерел звуку базується на перетворенні одиничного вектора V у систему просторових координат. Аналогічні способи демонстрації джерел звуку використовуються на фіг. 36-36C.

[650] Фіг. 37 демонструє представлення описаних тут методів одержання просторово-часової інтерполяції. Блок 50 просторово-часової інтерполяції пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4, може здійснювати просторово-часову інтерполяцію, описану нижче більш детально. Просторово-часова інтерполяція може включати в себе одержання просторових компонентів більш високого розрізнення в просторовому і часовому вимірах. Просторові компоненти можуть базуватися на ортогональному розкладанні багатовимірного сигналу, що складає з коефіцієнтів амбіофонії більш високого порядку (HOA) (або, як також можуть іменуватися коефіцієнти HOA, "коефіцієнтів сферичних гармонік").

[651] У проілюстрованому графіку, вектори V_1 і V_2 представляють відповідні вектори двох різних просторових компонентів багатовимірного сигналу. Просторові компоненти можна одержувати поблоковим розкладанням багатовимірного сигналу. У деяких прикладах, просторові компоненти одержані здійсненням поблокової форми SVD відносно кожного блока (який може означати кадр) аудіоданих амбіофонії більш високого порядку (HOA) (де ці аудіодані амбіофонії включають в себе блоки, вибірки або будь-яку іншу форму багатоканальних аудіоданих). Змінну M можна використовувати для позначення довжини кадру аудіо у вибірках.

[652] Відповідно, V_1 і V_2 можуть представляти відповідні вектори з векторів 51_k переднього плану $V[k]$ і векторів 51_{k-1} переднього плану $V[k-1]$ для послідовних блоків коефіцієнтів 11 HOA. V_1 може, наприклад, представляти перший вектор з векторів 51_{k-1} переднього плану $V[k-1]$ для першого кадру ($k-1$), тоді як V_2 може представляти перший вектор з векторів 51_k переднього плану $V[k]$ для другого і наступних кадрів (k). V_1 і V_2 можуть представляти просторовий компонент для одиничного аудіооб'єкта, включеного в багатовимірний сигнал.

[653] Інтерпольовані вектори V_x для кожного x одержуються шляхом зважування V_1 і V_2 відповідно до кількості відрізків часу або "часових вибірок", x , для часового компонента багатовимірного сигналу, до яких можуть застосовуватися інтерпольовані вектори V_x для згладжування часового (і, отже, у ряді випадків просторового) компонента. Передбачаючи композицію SVD, як описано вище, згладжування сигналів 49 pFG можна одержувати, виконуючи векторне ділення кожного вектора часової вибірки (наприклад, вибірки коефіцієнтів 11 HOA) на відповідний інтерпольований V_x . Таким чином, $US[n]=HOA[n]*V_x[n]^{-1}$, де це представляє вектор-рядок помножений на вектор-стовпець, таким чином, створюючи скалярний елемент для US . $V_x[n]^{-1}$ можна одержувати шляхом псевдоінвертування $V_x[n]$.

[654] Відносно зважування V_1 і V_2 , V_1 зважується пропорційно нижче в часовому вимірі, оскільки V_2 виникає пізніше за часом, ніж V_1 . Таким чином, хоча вектори 51_{k-1} переднього плану $V[k-1]$ є просторовими компонентами розкладання, більш пізні за часом вектори 51_k переднього плану $V[k]$ представляють різні значення просторового компонента протягом часу. Відповідно, ваговий коефіцієнт V_1 зменшується, коли ваговий коефіцієнт V_2 росте при збільшенні x по t . Тут, d_1 і d_2 представляють вагові коефіцієнти.

[655] На фіг. 38 показана блок-схема, що демонструє штучні матриці US , US_1 і US_2 , для послідовних блоків SVD для багатовимірного сигналу відповідно до описаних тут методів. Інтерпольовані V -вектори можуть застосовуватися до векторів-рядків штучних матриць US для

відновлення вихідного багатовимірного сигналу. Зокрема, блок 50 просторово-часової інтерполяції може множити псевдообернені інтерпольованим векторам 53 переднього плану $V[k]$ на результат множення сигналів 49 nFG на вектори 51_k переднього плану $V[k]$ (які можуть бути позначені як коефіцієнти НОА переднього плану) для одержання $K/2$ інтерпольованих вибірок, які можна використовувати замість $K/2$ вибірок сигналів nFG як перші $K/2$ вибірок, як показано в прикладі матриці U_2 , наведеному на фіг. 38.

[656] На фіг. 39 показана блок-схема, що демонструє розкладання наступних кадрів сигналу амбіофонії більш високого порядку (НОА) з використанням розкладання на сингулярні значення і згладжування просторово-часових компонентів відповідно до методів, описаних у цьому винаході. Кадр $n-1$ і кадр n (які також можуть бути позначені як кадр n і кадр $n+1$) представляють наступні кадри в часі, причому кожен кадр містить 1024 відрізків часу і має порядок НОА 4, що дає $(4+1)^2=25$ коефіцієнтів. US-матриці, що є штучно згладженими U-матрицями в кадрі $n-1$ і кадрі n , можна одержувати, застосовуючи інтерпольовані V-вектори, як показано. Кожний сірий вектор-рядок або вектор-стовпець представляє один аудіооб'єкт.

[657] Обчислення представлення НОА активних сигналів на векторній основі

[658] Миттєвий CVECK створюється шляхом узяття кожного із сигналів на векторній основі, представлених у XVECK, і множення його на відповідний йому (деквантований) просторовий вектор, VVECK. Кожен VVECK представлений у MVECK. Таким чином, для сигналу НОА порядку L і M сигналів на векторній основі, буде M сигналів на векторній основі, кожен з яких буде мати розмір, заданий довжиною кадру, P . Ці сигнали, таким чином, можна представити як: $XVECK_{mp}$, $p=0, \dots, P-1$; $m=0, \dots, M-1$. Відповідно, буде M просторових векторів, VVECK, розміром $(L+1)^2$. Їх можна представити як: $MVECK_{ml}$, $l=0, \dots, (L+1)^2-1$; $m=0, \dots, M-1$. Представлення НОА для кожного сигналу на векторній основі, CVECK_m, є множенням матриці на вектор, заданим як:

$$CVECK_m = (XVECK_m (MVECK_m)^T) T,$$
 що створює матрицю $(L+1)^2$ на P . Повне представлення НОА задається підсумовуванням внеску кожного сигналу на векторній основі наступним чином:

$$CVECK = m=0M-1 CVECK[m].$$

[659] Просторово-часова інтерполяція V-векторів

Однак, для підтримки плавної просторово-часової неперервності, вищезгадане обчислення здійснюється тільки для частини довжини кадру, $P-B$. Перші B вибірок матриці НОА, замість цього, здійснюються з використанням інтерпольованого набору $MVECK_{ml}$, $m=0, \dots, M-1$; $l=0, \dots, (L+1)^2$, виведеного з поточного $MVECK_m$ і попередніх значень $MVECK_{m-1}$. Це приводить до просторового вектора більш високої часової щільності, оскільки ми виводимо вектор для кожної часової вибірки, p , наступним чином:

$$MVECK_{mp} = pB-1 MVECK_m + B-1-pB-1 MVECK_{m-1}, p=0, \dots, B-1.$$

Для кожної часової вибірки, p , новий вектор розміром НОА $(L+1)^2$ обчислюється як:

$$CVECK_p = (XVECK_{mp}) MVECK_{mp}, p=0, \dots, B-1.$$

Ці перші B вибірок доповнюються $P-B$ вибірками попередньої секції для одержання повного представлення НОА, CVECK_m, m -го сигналу на векторній основі.

[660] На декодері (наприклад, пристрої 24 аудіодекодування, показаному в прикладі, наведеному на фіг. 5), для визначеного характерного, передньопланового або переважного звуку на векторній основі, V-вектор з попереднього кадру і V-вектор з поточного кадру можна інтерпольувати з використанням лінійної (або нелінійної) інтерполяції для створення інтерпольованого V-вектора більш високого розрізнення (за часом) протягом конкретного відрізка часу. Блок 76 просторово-часової інтерполяції може здійснювати цю інтерполяцію, де блок 76 просторово-часової інтерполяції може потім множити вектор US у поточному кадрі на інтерпольований V-вектор більш високого розрізнення для створення матриці НОА протягом цього конкретного відрізка часу.

[661] Альтернативно, блок 76 просторово-часової інтерполяції може множити вектор US на V-вектор поточного кадру для створення першої матриці НОА. Декодер може додатково множити вектор US на V-вектор з попереднього кадру для створення другої матриці НОА. Блок 76 просторово-часової інтерполяції потім може застосовувати лінійну (або нелінійну) інтерполяцію до першої і другої матриць НОА протягом конкретного відрізка часу. Вихід цієї інтерполяції може збігатися з виходом множення вектора US на інтерпольований V-вектор, за умови загальних вхідних матриць/векторів.

[662] У цьому відношенні, методи можуть передбачати, що пристрій 20 аудіокодування і/або пристрій 24 аудіодекодування виконані з можливістю діяти згідно з наступними пунктами.

[663] Пункт 135054-1С. Пристрій, наприклад пристрій 20 аудіокодування або пристрій 24 аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю одержання множини просторових компонентів більш високого розрізнення в просторі і часі, у

якому просторові компоненти базуються на ортогональному розкладанні багатовимірною сигналу, утвореного коефіцієнтами сферичних гармонік.

[664] Пункт 135054-1D. Пристрій, наприклад пристрій 20 аудіокодування або пристрій 24 аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю згладжування щонайменше одного з просторових компонентів і часових компонентів першої множини коефіцієнтів сферичних гармонік і другої множини коефіцієнтів сферичних гармонік.

[665] Пункт 135054-1E. Пристрій, наприклад пристрій 20 аудіокодування або пристрій 24 аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю одержання множини просторових компонентів більш високого розрізнення в просторі і часі, у якому просторові компоненти базуються на ортогональному розкладанні багатовимірною сигналу, утвореного коефіцієнтами сферичних гармонік.

[666] Пункт 135054-1G. Пристрій, наприклад пристрій 20 аудіокодування або пристрій 24 аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю одержання розкладених коефіцієнтів сферичних гармонік підвищеного розрізнення для відрізка часу шляхом, щонайменше частково, збільшення розрізнення відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік.

[667] Пункт 135054-2G. Пристрій за пунктом 135054-1G, у якому перше розкладання містить першу матрицю V, яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

[668] Пункт 135054-3G. Пристрій за пунктом 135054-1G, у якому друге розкладання містить другу матрицю V, яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[669] Пункт 135054-4G. Пристрій за пунктом 135054-1G, у якому перше розкладання містить першу матрицю V, яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік, і при цьому друге розкладання містить другу матрицю V, яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[670] Пункт 135054-5G. Пристрій за пунктом 135054-1G, у якому відрізок часу містить підкадр кадру аудіо.

[671] Пункт 135054-6G. Пристрій за пунктом 135054-1G, у якому відрізок часу містить часову вибірку кадру аудіо.

[672] Пункт 135054-7G. Пристрій за пунктом 135054-1G, у якому один або більше процесорів виконані з можливістю одержання інтерпольованого розкладання першого розкладання і другого розкладання для коефіцієнта сферичної гармоніки з першої множини коефіцієнтів сферичних гармонік.

[673] Пункт 135054-8G. Пристрій за пунктом 135054-1G, у якому один або більше процесорів виконані з можливістю одержання інтерпольованих розкладань першого розкладання для першої частини першої множини коефіцієнтів сферичних гармонік, включених у перший кадр, і другого розкладання для другої частини другої множини коефіцієнтів сферичних гармонік, включених у другий кадр, у якому один або більше процесорів додатково виконані з можливістю застосування інтерпольованих розкладань до першого часового компонента першої частини першої множини коефіцієнтів сферичних гармонік, включених у перший кадр для генерації першого штучного часового компонента першої множини коефіцієнтів сферичних гармонік, і застосування відповідних інтерпольованих розкладань до другого часового компонента другої частини другої множини коефіцієнтів сферичних гармонік, включених у другий кадр, для генерації другого штучного часового компонента другої множини коефіцієнтів сферичних гармонік.

[674] Пункт 135054-9G. Пристрій за пунктом 135054-8G, у якому перший часовий компонент генерується шляхом здійснення синтезу на векторній основі відносно першої множини коефіцієнтів сферичних гармонік.

[675] Пункт 135054-10G. Пристрій за пунктом 135054-8G, у якому другий часовий компонент генерується шляхом здійснення синтезу на векторній основі відносно другої множини коефіцієнтів сферичних гармонік.

[676] Пункт 135054-11G. Пристрій за пунктом 135054-8G, у якому один або більше процесорів додатково виконані з можливістю прийому першого штучного часового компонента і другого штучного часового компонента, обчислення інтерпольованих розкладань першого розкладання для першої частини першої множини коефіцієнтів сферичних гармонік і другого розкладання для другої частини другої множини коефіцієнтів сферичних гармонік і застосування обернених інтерпольованих розкладань до першого штучного часового компонента для

відновлення першого часового компонента і до другого штучного часового компонента для відновлення другого часового компонента.

5 [677] Пункт 135054-12G. Пристрій за пунктом 135054-1G, у якому один або більше процесорів виконані з можливістю інтерполяції першого просторового компонента першої множини коефіцієнтів сферичних гармонік і другого просторового компонента другої множини коефіцієнтів сферичних гармонік.

[678] Пункт 135054-13G. Пристрій за пунктом 135054-12G, у якому перший просторовий компонент містить першу матрицю U , яка представляє ліві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

10 [679] Пункт 135054-14G. Пристрій за пунктом 135054-12G, у якому другий просторовий компонент містить другу матрицю U , яка представляє ліві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

15 [680] Пункт 135054-15G. Пристрій за пунктом 135054-12G, у якому перший просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для першої множини коефіцієнтів сферичних гармонік і другий просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для другої множини коефіцієнтів сферичних гармонік.

20 [681] Пункт 135054-16G. Пристрій за пунктом 135054-12G, у якому перший просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для першої множини коефіцієнтів сферичних гармонік і другий просторовий компонент представляє M відрізків часу коефіцієнтів сферичних гармонік для другої множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів виконані з можливістю одержання розкладених інтерпольованих коефіцієнтів сферичних гармонік для відрізка часу і інтерполяції останніх N елементів першого просторового компонента і перших N елементів другого просторового компонента.

25 [682] Пункт 135054-17G. Пристрій за пунктом 135054-1G, у якому друга множина коефіцієнтів сферичних гармонік іде за першою множиною коефіцієнтів сферичних гармонік у часовій області.

30 [683] Пункт 135054-18G. Пристрій за пунктом 135054-1G, у якому один або більше процесорів додатково виконані з можливістю розкладання першої множини коефіцієнтів сферичних гармонік для генерації першого розкладання першої множини коефіцієнтів сферичних гармонік.

35 [684] Пункт 135054-19G. Пристрій за пунктом 135054-1G, у якому один або більше процесорів додатково виконані з можливістю розкладання другої множини коефіцієнтів сферичних гармонік для генерації другого розкладання другої множини коефіцієнтів сферичних гармонік.

40 [685] Пункт 135054-20G. Пристрій за пунктом 135054-1G, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно першої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори першої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення першої множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

45 [686] Пункт 135054-21G. Пристрій за пунктом 135054-1G, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно другої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори другої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення другої множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

50 [687] Пункт 135054-22G. Пристрій за пунктом 135054-1G, у якому перша і друга множини коефіцієнтів сферичних гармонік представляють плоскохвильове представлення звукового поля.

55 [688] Пункт 135054-23G. Пристрій за пунктом 135054-1G, у якому перша і друга множини коефіцієнтів сферичних гармонік представляють один або більше змішаних один з одним монофонічних аудіооб'єктів.

[689] Пункт 135054-24G. Пристрій за пунктом 135054-1G, у якому перша і друга множини коефіцієнтів сферичних гармонік містять відповідні перший і другий коефіцієнти сферичних гармонік, що представляють тривимірне звукове поле.

[690] Пункт 135054-25G. Пристрій за пунктом 135054-1G, у якому перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, що має порядок, який більше одиниці.

5 [691] Пункт 135054-26G. Пристрій за пунктом 135054-1G, у якому перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, що має порядок, який дорівнює чотирьом.

10 [692] Пункт 135054-27G. Пристрій за пунктом 135054-1G, у якому інтерполяція є зваженою інтерполяцією першого розкладання і другого розкладання, причому вагові коефіцієнти зваженої інтерполяції, застосовувані до першого розкладання, зворотно пропорційні часу, представленому векторами першого і другого розкладання, і при цьому вагові коефіцієнти зваженої інтерполяції, застосовувані до другого розкладання, пропорційні часу, представленому векторами першого і другого розкладання.

15 [693] Пункт 135054-28G. Пристрій за пунктом 135054-1G, у якому розкладені інтерпольовані коефіцієнти сферичних гармонік згладжують щонайменше один із просторових компонентів і часових компонентів першої множини коефіцієнтів сферичних гармонік і другої множини коефіцієнтів сферичних гармонік.

[694]

20 [695] На фіг. 40A-40J показані блок-схеми, що демонструють приклади пристроїв 510A-510J аудіокодування, які можуть здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. У кожному із прикладів, наведених на фіг. 40A-40J, кожний з пристроїв 510A і 510B аудіокодування, у деяких прикладах, представляє будь-який пристрій, здатний кодувати аудіодані, наприклад настільний комп'ютер, портативний комп'ютер, робочу станцію, планшет або графічний планшет, спеціалізований звукозаписувальний пристрій, стільниковий телефон (включаючи так називані "смартфони"), персональний пристрій медіапрогравача, персональний ігровий пристрій або пристрій будь-якого іншого типу, здатний кодувати аудіодані.

25 [696] Хоча вони показані як єдиний пристрій, тобто пристрої 510A-510J у прикладах, наведених на фіг. 40A-40J, різні компоненти або блоки, згадані нижче як включені в пристрої 510A-510J, можуть фактично формувати окремі пристрої, зовнішні відносно пристроїв 510A-510J. Іншими словами, хоча вони описані в цьому винаході як здійснювані єдиним пристроєм, тобто пристроями 510A-510J у прикладах, наведених на фіг. 40A-40J, методи можуть реалізуватися або інакше здійснюватися системою, що містить множинні пристрої, причому кожний з цих пристроїв може включати в себе один або більше з різних компонентів або блоків, більш докладно описаних нижче. Відповідно, методи не підлягають обмеженню прикладами, наведеними на фіг. 40A-40J.

30 [697] У деяких прикладах, пристрої 510A-510J аудіокодування представляють пристрої аудіокодування, альтернативні описаним вище відносно прикладів, наведених на фіг. 3 і 4. Протягом нижченаведеного розгляду пристроїв 510A-510J аудіокодування різні подібності відносно операції зазначені відносно різних блоків 30-52 пристрою 20 аудіокодування, описаного вище з посиланням на фіг. 4. У багатьох відношеннях, пристрої 510A-510J аудіокодування можуть, як описано нижче, діяти, по суті, аналогічним чином пристрою 20 аудіокодування, хоча і з незначними відхиленнями або модифікаціями.

35 [698] Як показано в прикладі, наведеному на фіг. 40A, пристрій 510A аудіокодування містить блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Блок 512 стиснення аудіосигналу може представляти блок, який стискає коефіцієнти сферичних гармонік (SHC) 511 ("SHC 511"), що також можуть бути позначені як коефіцієнти 511 амбіофонії більш високого порядку (HOA). Блок 512 стиснення аудіосигналу може, в ряді випадків, представляти блок, який може без втрат стискати або здійснювати втратне стиснення відносно SHC 511. SHC 511 може представляти множину SHC, де щонайменше один з множини SHC 40 відповідає сферичній базисній функції, що має порядок, більший одиниці (де SHC цієї множини іменуються амбіофонією більш високого порядку (HOA) для розрізнення від амбіофонії більш низького порядку, одним прикладом якої є так називаний "В-формат"), як описано більш детально вище. Хоча блок 512 стиснення аудіосигналу може без втрат стискати SHC 511, у деяких прикладах, блок 512 стиснення аудіосигналу видаляє ті з SHC 511, які не помітні або не 45 мають відношення до опису звукового поля при відтворенні (у тому розумінні, що може не сприйматися людською звуковою системою). У цьому розумінні, втратний характер цього стиснення може не перебивати вплив на сприйману якість звукового поля при відтворенні зі стисненої версії SHC 511.

50 [699] У прикладі, наведеному на фіг. 40A, блок стиснення аудіосигналу включає в себе блок 518 розкладання і блок 520 витягання компонентів звукового поля. Блок 518 розкладання може 60

бути аналогічним блоку 30 лінійного оборотного перетворення пристрою 20 аудіокодування. Таким чином, блок 518 розкладання може представляти блок, виконаний з можливістю здійснення форми аналізу, що іменується розкладання на сингулярні значення. Хоча вони описані відносно SVD, методи можна здійснювати відносно будь-якого аналогічного перетворення або розкладання, що забезпечує набори лінійно розкорельованих даних. Також посилання на "набори" у цьому винаході має означати "ненульові" набори, якщо прямо не зазначене зворотне, і не має означати класичне математичне визначення множин, що включає в себе так названу "порожню множину".

[700] У будь-якому випадку, блок 518 розкладання здійснює розкладання на сингулярні значення (яке, знову ж, може бути позначене аббревіатурою "SVD") для перетворення коефіцієнтів 511 сферичних гармонік у два або більше наборів перетворених коефіцієнтів сферичних гармонік. У прикладі, наведеному на фіг. 40, блок 518 розкладання може здійснювати SVD відносно SHC 511 для генерації так названої матриці 519 V, матриці 519B S і матриці 519C U. У прикладі, наведеному на фіг. 40, блок 518 розкладання виводить кожну з матриць окремо замість виведення векторів US[k] у комбінованій формі, як розглянуто вище відносно блока 30 лінійного оборотного перетворення.

[701] Як згадано вище, матриця V^* у згаданому вище математичному виразі SVD позначається як спряжена транспонована матриця V для відображення того, що SVD може застосовуватися до матриць, які містять комплексні числа. У випадку застосування до матриць, які містять тільки дійсні числа, комплексно-спряжену матрицю V (або, іншими словами, матрицю V^*) можна розглядати такою, що дорівнює матриці V. Нижче передбачається, для спрощення ілюстрації, що SHC 511 містять дійсні числа, у результаті чого матриця V виводиться через SVD замість матриці V^* . Хоча передбачається, що це матриця V, методи можуть застосовуватися аналогічним чином до SHC 511, що має комплексні коефіцієнти, де виходом SVD є матриця V^* . Відповідно, методи не підлягають обмеженню в цьому відношенні тільки забезпеченням застосування SVD для генерації матриці V, але можуть включати в себе застосування SVD до SHC 511, що має комплексні компоненти, для генерації матриці V^* .

[702] У будь-якому випадку, блок 518 розкладання може здійснювати поблокову форму SVD відносно кожного блока (який може означати кадр) аудіоданих амбіофонії більш високого порядку (НОА) (де ці аудіодані амбіофонії включають в себе блоки або вибірки SHC 511 або будь-яку іншу форму багатоканальних аудіоданих). Змінну M можна використовувати для позначення довжини кадру аудіо у вибірках. Наприклад, коли кадр аудіо включає в себе 1024 вибірки аудіосигналу, M дорівнює 1024. Блок 518 розкладання, таким чином, може здійснювати поблокове SVD відносно блока SHC 511, що має M на $(N+1)^2$ SHC, де N, знову ж, позначає порядок аудіоданих НОА. Блок 518 розкладання може генерувати, за допомогою здійснення цього SVD, матрицю 519A V, матрицю 519B S і матрицю 519C U, де кожна з матриць 519A-519C ("матриці 519") може представляти відповідні матриці V, S і U, описані більш детально вище. Блок 518 розкладання може передавати або виводити ці матриці 519A на блок 520 витягання компонентів звукового поля. Матриця 519A V може мати розмір $(N+1)^2$ на $(N+1)^2$, матриця 519B S може мати розмір $(N+1)^2$ на $(N+1)^2$ і матриця U може мати розмір M на $(N+1)^2$, де M означає кількість вибірок у кадрі аудіо. Типове значення M дорівнює 1024, хоча методи цього винаходу не підлягають обмеженню цим типовим значенням для M.

[703] Блок 520 витягання компонентів звукового поля може представляти блок, виконаний з можливістю визначати і потім витягати характерні компоненти звукового поля і фонові компоненти звукового поля, ефективно відділяючи характерні компоненти звукового поля від фонових компонентів звукового поля. У цьому відношенні, блок 520 витягання компонентів звукового поля може здійснювати багато які з операцій, описаних вище відносно блока 44 аналізу звукового поля, блока 48 вибору фону і блока 36 вибору переднього плану пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. За умови, що характерні компоненти звукового поля, у деяких прикладах, вимагають базисних функцій більш високого порядку (відносно фонових компонентів звукового поля) (і, таким чином, більше SHC) для точного представлення характерного характеру цих компонентів, відділення характерних компонентів від фонових компонентів може надавати можливість виділення більшої кількості бітів характерним компонентам і виділення меншої кількості бітів (відносно мови) фоновим компонентам. Відповідно, шляхом застосування цього перетворення (у формі SVD або будь-якій іншій формі перетворення, що включає в себе PCA), методи, описані в цьому винаході, можуть полегшувати виділення бітів різним SHC і, таким чином, стиснення SHC 511.

[704] Крім того, методи також можуть передбачати, що, як більш докладно описано нижче з посиланням на фіг. 40B, зниження порядку фонових компонентів звукового поля, за умови, що базисні функції більш високого порядку, у деяких прикладах, не вимагається для представлення

цих фонових ділянок звукового поля з урахуванням дифузного або фонового характеру цих компонентів. Методи, таким чином, забезпечують стиснення дифузних або фонових аспектів звукового поля, зберігаючи помітні характерні компоненти або аспекти звукового поля шляхом застосування SVD до SHC 511.

5 [705] Як показано далі в прикладі, наведеному на фіг. 40, блок 520 витягання компонентів звукового поля включає в себе блок 522 транспонування, блок 524 аналізу помітних компонентів і математичний блок 526. Блок 522 транспонування представляє блок, виконаний з
10 можливістю транспонування матриці 519A V для генерації транспонованої матриці 519A V, що позначається як "матриця 523 V^T". Блок 522 транспонування може виводити цю матрицю 523 V^T на математичний блок 526. Матриця 523 V^T може мати розмір (N+1)² на (N+1)².

[706] Блок 524 аналізу помітних компонентів представляє блок, виконаний з можливістю здійснення аналізу помітності відносно матриці 519B S. Блок 524 аналізу помітних компонентів може, у цьому відношенні, здійснювати операції, аналогічні описаним вище відносно блока 44 аналізу звукового поля пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг.
15 4. Блок 524 аналізу помітних компонентів може аналізувати діагональні значення матриці 519B S, вибрати змінну D кількості цих компонентів, що мають найбільше значення. Іншими словами, блок 524 аналізу помітних компонентів може визначати значення D, яке розділяє два підпростори (наприклад, передньоплановий або переважний підпростір і підпростір фону або оточення), шляхом аналізу нахилу кривої, створеної зменшуваними діагональними значеннями
20 S, де великі сингулярні значення представляють звуки переднього плану або характерні звуки і низькі сингулярні значення представляють фонові компоненти звукового поля. У деяких прикладах, блок 524 аналізу помітних компонентів може використовувати першу і другу похідні кривої сингулярних значень. Блок 524 аналізу помітних компонентів також може обмежувати число D величиною між одиницею і п'ятьма. У порядку іншого прикладу, блок 524 аналізу
25 помітних компонентів може обмежувати число D величиною між одиницею і (N+1)². Альтернативно, блок 524 аналізу помітних компонентів може попередньо задавати число D, наприклад значення чотири. У будь-якому випадку, оцінивши число D, блок 524 аналізу помітних компонентів витягає підпростори переднього плану і фону з матриць U, V і S.

[707] У деяких прикладах, блок 524 аналізу помітних компонентів може здійснювати цей
30 аналіз кожні M вибірок, який може відновлюватися на основі кадрів. У цьому відношенні, D може змінюватися від кадру до кадру. В інших прикладах, блок 524 аналізу помітних компонентів може здійснювати цей аналіз більше одного разу за кадр, аналізуючи дві або більше ділянок кадру. Відповідно, методи не підлягають обмеженню в цьому відношенні прикладами, наведеними в цьому винаході.

35 [708] Фактично, блок 524 аналізу помітних компонентів може аналізувати сингулярні значення діагональної матриці, що позначається як матриця 519B S у прикладі, наведеному на фіг. 40, яка ідентифікує значення, що мають відносне значення, яке перевищує інші значення діагональної матриці 519B S. Блок 524 аналізу помітних компонентів може ідентифікувати значення D, витягаючи ці значення для генерації матриці 525A S_{DIST} і матриці 525B S_{BG}.
40 Матриця 525A S_{DIST} може представляти діагональну матрицю, яка містить D стовпців, що мають (N+1)² вихідної матриці 519B S. У ряді випадків, матриця 525B S_{BG} може представляти матрицю, що має (N+1)²-D стовпців, кожний з яких включає в себе (N+1)² перетворених коефіцієнтів сферичних гармонік вихідної матриці 519B S. Хоча описана як матриця S_{DIST}, що представляє матрицю, яка містить D стовпців, що мають (N+1)² значень вихідної матриці 519B S, блок 524
45 аналізу помітних компонентів може зрізати цю матрицю для генерації матриці S_{DIST}, що має D стовпців, які мають значення D вихідної матриці 519B S, за умови, що матриця 519B S є діагональною матрицею, і (N+1)² значень D стовпців після D-го значення в кожному стовпці часто мають значення нуль. Хоча описані відносно повної матриці 525A S_{DIST} і повної матриці 525B S_{BG}, методи можна реалізувати відносно зрізаних версій цих матриць 525A S_{DIST} і зрізаної версії
50 цієї матриці 525B S_{BG}. Відповідно, методи цього винаходу не підлягають обмеженню в цьому відношенні.

[709] Іншими словами, матриця 525A S_{DIST} може мати розмір D-by-(N+1)², тоді як матриця 525B S_{BG} може мати розмір (N+1)²-D-by-(N+1)². Матриця 525A S_{DIST} може включати в себе головні компоненти або, іншими словами, сингулярні значення, що визначаються як помітні в
55 тому розумінні, що вони є характерними (DIST) аудіокомпонентами звукового поля, тоді як матриця 525B S_{BG} може включати в себе ті сингулярні значення, що визначаються як фонові (BG) або, іншими словами, оточуючі або нехарактерні аудіокомпоненти звукового поля. Хоча вони показані як окремі матриці 525A і 525B у прикладі, наведеному на фіг. 40, матриці 525A і 525B можуть указуватися як єдина матриця з використанням змінної D для позначення кількості

стовпців (зліва направо) цієї єдиної матриці, що представляють матрицю $525 S_{DIST}$. У деяких прикладах, змінну D можна задати рівною чотирьом.

[710] Блок 524 аналізу помітних компонентів також може аналізувати матрицю $519C U$ для генерації матриці $525C U_{DIST}$ і матриці $525D U_{BG}$. Часто, блок 524 аналізу помітних компонентів може аналізувати матрицю $519B S$ для ідентифікації змінної D , генерації матриці $525C U_{DIST}$ і матриці $525B U_{BG}$ на основі змінної D . Таким чином, після ідентифікації D стовпців матриці $519B S$, що є помітними, блок 524 аналізу помітних компонентів може розбивати матрицю $519C U$ на основі цієї визначеної змінної D . У цьому випадку, блок 524 аналізу помітних компонентів може генерувати матрицю $525C U_{DIST}$ для включення D стовпців (зліва направо) $(N+1)^2$ перетворених коефіцієнтів сферичних гармонік вихідної матриці $519C U$ і матриці $525D U_{BG}$ для включення $(N+1)^2$ - D стовпців, що залишилися, $(N+1)^2$ перетворених коефіцієнтів сферичних гармонік вихідної матриці $519C U$. Матриця $525C U_{DIST}$ може мати розмір M на D , тоді як матриця $525D U_{BG}$ може мати розмір M на $(N+1)^2$ - D . Хоча вони показані як окремі матриці $525C$ і $525D$ у прикладі, наведеному на фіг. 40, матриці $525C$ і $525D$ можуть указуватися як єдина матриця з використанням змінної D для позначення кількості стовпців (зліва направо) цієї єдиної матриці, що представляють матрицю U_{DIST} $525B$.

[711] Блок 524 аналізу помітних компонентів також може аналізувати матрицю $523 V^T$ для генерації матриці $525E V_{DIST}^T$ і матриці $525F V_{BG}^T$. Часто, блок 524 аналізу помітних компонентів може аналізувати матрицю $519B S$ для ідентифікації змінної D , генерації матриці $525E V_{DIST}^T$ і матриці $525F V_{BG}^T$ на основі змінної D . Таким чином, після ідентифікації D стовпців матриці $519B S$, що є помітними, блок 254 аналізу помітних компонентів може розбивати матрицю $519A V$ на основі цієї визначеної змінної D . У цьому випадку, блок 524 аналізу помітних компонентів може генерувати матрицю $525E V_{DIST}^T$ для включення $(N+1)^2$ рядків (зверху вниз) значень D вихідної матриці $523 V^T$ і матрицю $525F V_{BG}^T$ для включення $(N+1)^2$ рядків, що залишилися, $(N+1)^2$ - D значень вихідної матриці $523 V^T$. Матриця $525E V_{DIST}^T$ може мати розмір $(N+1)^2$ на D , тоді як матриця $525D V_{BG}^T$ може мати розмір $(N+1)^2$ на $(N+1)^2$ - D . Хоча вони показані як окремі матриці $525E$ і $525F$ у прикладі, наведеному на фіг. 40, матриці $525E$ і $525F$ можуть указуватися як єдина матриця з використанням змінної D для позначення кількості стовпців (зліва направо) цієї єдиної матриці, що представляють матрицю $525E V_{DIST}$. Блок 524 аналізу помітних компонентів може виводити матрицю $525 S_{DIST}$, матрицю $525B S_{BG}$, матрицю $525C U_{DIST}$, матрицю $525D U_{BG}$ і матрицю $525F V_{BG}^T$ на математичний блок 526, також виводячи матрицю $525E V_{DIST}^T$ на блок 516 генерації бітового потоку.

[712] Математичний блок 526 може представляти блок, виконаний з можливістю здійснювати матричні множення або будь-яку іншу математичну операцію, яка може здійснюватися відносно однієї або більше матриць (або векторів). Зокрема, як показано в прикладі, наведеному на фіг. 40, математичний блок 526 може представляти блок, виконаний з можливістю здійснення матричного множення для множення матриці $525C U_{DIST}$ на матрицю $525A S_{DIST}$ для генерації векторів $527 U_{DIST}^T S_{DIST}$ розміром M на D . Блок 526 матричного множення також може представляти блок, виконаний з можливістю здійснення матричного множення для множення матриці $525D U_{BG}$ на матрицю $525B S_{BG}$ і потім на матрицю $525F V_{BG}^T$ для генерації матриці $525F U_{BG}^T S_{BG} V_{BG}^T$ для генерації коефіцієнтів 531 сферичних гармонік фону розміром M на $(N+1)^2$ (які можуть представляти коефіцієнти 511 сферичних гармонік, що представляють фонові компоненти звукового поля). Математичний блок 526 може виводити вектори $527 U_{DIST}^T S_{DIST}$ і коефіцієнти 531 сферичних гармонік фону на блок 514 аудіокодування.

[713] Таким чином, пристрій 510 аудіокодування відрізняється від пристрою 20 аудіокодування тим, що пристрій 510 аудіокодування включає в себе цей математичний блок 526, виконаний з можливістю генерації векторів $527 U_{DIST}^T S_{DIST}$ і коефіцієнтів 531 сферичних гармонік фону шляхом матричного множення наприкінці процесу кодування. Блок 30 лінійного оборотного перетворення пристрою 20 аудіокодування здійснює множення матриць U і S для виведення векторів $33 US[k]$ у відносному початку процесу кодування, що може полегшувати подальші операції, наприклад переупорядковування, не показане у прикладі, наведеному на фіг. 40. Крім того, пристрій 20 аудіокодування, замість відновлення фонового SHC 531 наприкінці процесу кодування, вибирає коефіцієнти 47 НОА фону безпосередньо з коефіцієнтів 11 НОА, таким чином, потенційно уникаючи матричних множень для відновлення фонового SHC 531.

[714] Блок 514 аудіокодування може представляти блок, що здійснює форму кодування для додаткового стиснення векторів $527 U_{DIST}^T S_{DIST}$ і коефіцієнтів 531 сферичних гармонік фону. Блок 514 аудіокодування може діяти, по суті, аналогічно блоку 40 психоакустичного аудіокодера пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. У ряді випадків, цей блок 514 аудіокодування може представляти один або більше екземплярів блока кодування

удосконаленого аудіокодування (AAC). Блок 514 аудіокодування може кодувати кожен стовпець або кожен рядок векторів $527 U_{DIST} * S_{DIST}$. Часто, блок 514 аудіокодування може викликати екземпляр блока кодування AAC для кожної з комбінацій порядків/підпорядків, що залишилися в коефіцієнтах 531 сферичних гармонік фону. Додаткову інформацію про те, як можна кодувати

5 коефіцієнти 531 сферичних гармонік фону з використанням блока кодування AAC можна знайти в матеріалах з'їзду Eric Hellerud et al., за назвою "Encoding Higher Order Ambisonics with AAC", представлених на 124-ому з'їзді, 17-20 травня 2008 р., і доступних за адресою: <http://ro.uow.edu.au/cgi/viewcontent.cgi?article=8025&context=engpapers>. Блок 14 аудіокодування може виводити кодовану версію векторів $527 U_{DIST} * S_{DIST}$ (позначених "кодовані вектори 515

10 $U_{DIST} * S_{DIST}$ ") і кодовану версію коефіцієнтів 531 сферичних гармонік фону (позначених "кодовані коефіцієнти 515B сферичних гармонік фону") на блок 516 генерації бітового потоку. У ряді випадків, блок 514 аудіокодування може аудіокодувати коефіцієнти 531 сферичних гармонік фону з використанням більш низької цільової бітової швидкості, ніж використовувані для кодування векторів $527 U_{DIST} * S_{DIST}$, таким чином потенційно стискаючи коефіцієнти 531

15 сферичних гармонік фону більше в порівнянні з векторами $527 U_{DIST} * S_{DIST}$.

[715] Блок 516 генерації бітового потоку представляє блок, що форматує дані для узгодження з відомим форматом (який може означати формат, відомий пристрою декодування), таким чином генеруючи бітовий потік 517. Блок 516 генерації бітового потоку може діяти, по суті, аналогічним чином описаному вище відносно блока 42 генерації бітового потоку пристрою 24

20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. Блок 516 генерації бітового потоку може включати в себе мультиплексор, що мультиплексує кодовані вектори $515 U_{DIST} * S_{DIST}$, кодовані коефіцієнти 515B сферичних гармонік фону і матрицю $525E V_{DIST}^T$.

[716] На фіг. 40B показана блок-схема, що демонструє приклад пристрою 510B аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510B аудіокодування може бути аналогічним пристрою 510 аудіокодування тим, що пристрій 510B аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514

25 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510B аудіокодування може бути аналогічним блоку пристрою 510 аудіокодування тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання. Блок 512 стиснення аудіосигналу пристрою 510B аудіокодування може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510 аудіокодування тим, що блок 520 витягання компонентів звукового поля включає в себе додатковий блок, позначений як блок 528A зниження порядку ("блок 528

30 зниження порядку"). З цієї причини, блок 520 витягання компонентів звукового поля пристрою 510B аудіокодування позначається як "блок 520B витягання компонентів звукового поля".

[717] Блок 528A зниження порядку представляє блок, виконаний з можливістю здійснення додаткового зниження порядку коефіцієнтів 531 сферичних гармонік фону. У ряді випадків, блок 528A зниження порядку може повертати звукове поле, представлене коефіцієнтами 531

40 сферичних гармонік фону, для зниження кількості коефіцієнтів 531 сферичних гармонік фону, необхідної для представлення звукового поля. У ряді випадків, за умови, що коефіцієнти 531 сферичних гармонік фону представляють фонові компоненти звукового поля, блок 528A зниження порядку може видаляти, виключати або інакше усувати (часто шляхом обнулення) коефіцієнти 531 сферичних гармонік фону, що відповідають сферичним базисним функціям більш високого порядку. У цьому відношенні, блок 528A зниження порядку може здійснювати операції, аналогічні операціям блока 48 вибору фону пристрою 20 аудіокодування, показаного в

45 прикладі, наведеному на фіг. 4. Блок 528A зниження порядку може виводити приведену версію коефіцієнтів 531 сферичних гармонік фону (позначених як "приведені коефіцієнти 529 сферичних гармонік фону") на блок 514 аудіокодування, що може здійснювати аудіокодування вищеописаним чином для кодування приведених коефіцієнтів 529 сферичних гармонік фону і,

50 таким чином, генерувати кодовані приведені коефіцієнти 515B сферичних гармонік фону.

[718] Різні перераховані нижче пункти можуть представляти різні аспекти методів, описаних у цьому винаході.

[719] Пункт 132567-1. Пристрій, наприклад, пристрій 510 аудіокодування або пристрій 510B аудіокодування, який містить: один або більше процесорів, виконаних з можливістю здійснювати

55 розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і представляти множину коефіцієнтів

сферичних гармонік як функцію щонайменше частини однієї або більше з матриці U , матриці S і матриці V .

[720] Пункт 132567-2. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю генерації бітового потоку для включення представлення з множини коефіцієнтів сферичних гармонік як одного або більше векторів матриці U , матриці S і матриці V , що включають у себе їх комбінації або їх похідні.

[721] Пункт 132567-3. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, визначати один або більше векторів U_{DIST} , включених у матрицю U , що описують характерні компоненти звукового поля.

[722] Пункт 132567-4. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, визначати один або більше векторів U_{DIST} , включених у матрицю U , що описують характерні компоненти звукового поля, визначати один або більше векторів S_{DIST} , включених у матрицю S , що також описують характерні компоненти звукового поля, і перемножувати один або більше векторів U_{DIST} і один або більше векторів S_{DIST} для генерації векторів $U_{DIST} * S_{DIST}$.

[723] Пункт 132567-5. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, визначати один або більше векторів U_{DIST} , включених у матрицю U , що описують характерні компоненти звукового поля, визначати один або більше векторів S_{DIST} , включених у матрицю S , що також описують характерні компоненти звукового поля, і перемножувати один або більше векторів U_{DIST} і один або більше векторів S_{DIST} для генерації одного або більше векторів $U_{DIST} * S_{DIST}$, і при цьому один або більше процесорів додатково виконані з можливістю аудіокодування одного або більше векторів $U_{DIST} * S_{DIST}$ для генерації аудіокодованої версії одного або більше векторів $U_{DIST} * S_{DIST}$.

[724] Пункт 132567-6. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, визначати один або більше векторів U_{BG} включених у матрицю U .

[725] Пункт 132567-7. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, аналізувати матрицю S для ідентифікації характерних і фонових компонентів звукового поля.

[726] Пункт 132567-8. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, аналізувати матрицю S для ідентифікації характерних і фонових компонентів звукового поля і визначати, на основі аналізу матриці S , один або більше векторів U_{DIST} матриці U , що описують характерні компоненти звукового поля, і один або більше векторів U_{BG} матриці U , що описують фонові компоненти звукового поля.

[727] Пункт 132567-9. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, аналізувати матрицю S для ідентифікації характерних і фонових компонентів звукового поля на основі кадрів аудіо і визначати, на основі аналізу матриці S для кожного кадру аудіо, один або більше векторів U_{DIST} матриці U , що описують характерні компоненти звукового поля, і один або більше векторів U_{BG} матриці U , що описують фонові компоненти звукового поля.

[728] Пункт 132567-10. Пристрій за пунктом 132567-1, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, аналізувати матрицю S для ідентифікації характерних і фонових компонентів звукового поля, визначати, на основі аналізу матриці S , один або більше векторів U_{DIST} матриці U , що описують характерні компоненти звукового поля, і один або більше векторів U_{BG} матриці U , що описують фонові компоненти звукового поля, визначати, на основі аналізу матриці S , один або більше векторів S_{DIST} і один або більше векторів S_{BG} матриці S , відповідних одному або більше векторам U_{DIST} і одному або більше векторам U_{BG} , і визначати, на основі аналізу матриці S , один або більше векторів V_{DIST}^T і один або більше векторів V_{BG}^T транспонованої матриці V , відповідних одному або більше векторам U_{DIST} і одному або більше векторам U_{BG} .

[729] Пункт 132567-11. Пристрій за пунктом 132567-10, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, множити один або більше векторів U_{BG} на один або більше векторів S_{BG} і потім на один або більше векторів V_{BG}^T для генерації одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$, і при цьому один або більше процесорів додатково виконані з можливістю аудіокодування векторів $U_{BG} * S_{BG} * V_{BG}^T$ для генерації аудіокодованої версії векторів $U_{BG} * S_{BG} * V_{BG}^T$.

[730] Пункт 132567-12. Пристрій за пунктом 132567-10, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, множити один або більше векторів U_{BG} на один або більше векторів S_{BG} і потім на один або більше векторів V_{BG}^T для генерації одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$ і здійснювати процес зниження порядку для виключення тих з коефіцієнтів одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$, які пов'язані з одним або більше порядками сферичних гармонічних базисних функцій, і, таким чином, генерувати версію зниженого порядку одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$.

[731] Пункт 132567-13. Пристрій за пунктом 132567-10, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, множити один або більше векторів U_{BG} на один або більше векторів S_{BG} і потім на один або більше векторів V_{BG}^T для генерації одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$ і здійснювати процес зниження порядку для виключення тих з коефіцієнтів одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$, які пов'язані з одним або більше порядками сферичних гармонічних базисних функцій, і, таким чином, генерувати версію зниженого порядку одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$, і при цьому один або більше процесорів додатково виконані з можливістю аудіокодування версії зниженого порядку одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$ для генерації аудіокодованої версії одного або більше векторів зниженого порядку $U_{BG} * S_{BG} * V_{BG}^T$.

[732] Пункт 132567-14. Пристрій за пунктом 132567-10, у якому один або більше процесорів додатково виконані з можливістю, при представленні множини коефіцієнтів сферичних гармонік, множити один або більше векторів U_{BG} на один або більше векторів S_{BG} і потім на один або більше векторів V_{BG}^T для генерації одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$, здійснювати процес зниження порядку для виключення тих з коефіцієнтів одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$, які пов'язані з одним або більше порядками, більшими, ніж для однієї з сферичних гармонічних базисних функцій, і, таким чином, генерувати версію зниженого порядку одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$, і аудіокодувати версію зниженого порядку одного або більше векторів $U_{BG} * S_{BG} * V_{BG}^T$ для генерації аудіокодованої версії одного або більше векторів зниженого порядку $U_{BG} * S_{BG} * V_{BG}^T$.

[733] Пункт 132567-15. Пристрій за пунктом 132567-10, у якому один або більше процесорів додатково виконані з можливістю генерації бітового потоку для включення одного або більше векторів V_{DIST}^T .

[734] Пункт 132567-16. Пристрій за пунктом 132567-10, у якому один або більше процесорів додатково виконані з можливістю генерації бітового потоку для включення одного або більше векторів V_{DIST}^T без аудіокодування одного або більше векторів V_{DIST}^T .

[735] Пункт 132567-1F. Пристрій, наприклад пристрій 510 або 510B аудіокодування, який містить: один або більше процесорів для здійснення розкладання на сингулярні значення відносно багатоканальних аудіоданих, що представляють щонайменше частину звукового поля, для генерації матриці U , яка представляє ліві сингулярні вектори багатоканальних аудіоданих, матриці S , яка представляє сингулярні значення багатоканальних аудіоданих, і матриці V , яка представляє праві сингулярні вектори багатоканальних аудіоданих, і представлення багатоканальних аудіоданих як функції щонайменше частини однієї або більше з матриці U , матриці S і матриці V .

[736] Пункт 132567-2F. Пристрій за пунктом 132567-1F, у якому багатоканальні аудіодані містять множину коефіцієнтів сферичних гармонік.

[737] Пункт 132567-3F. Пристрій за пунктом 132567-2F, у якому один або більше процесорів додатково виконані з можливістю здійснення відповідно до будь-якої комбінації пунктів 132567-2-132567-16.

[738] З кожного з різних описаних вище пунктів випливає, що будь-який з пристроїв 510A-510J аудіокодування може здійснювати спосіб або іншим чином містити засіб для здійснення кожного етапу способу, для якого пристрій 510A-510J аудіокодування виконаний з можливістю здійснення в ряді випадків, цей засіб може містити один або більше процесорів. У ряді випадків, один або більше процесорів можуть представляти процесор спеціального призначення, який виконує інструкції, що зберігаються на довгостроковому зчитуваному комп'ютером носії даних. Іншими словами, різні аспекти методів у кожному з наборів прикладів кодування можуть передбачати довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати спосіб, для здійснення якого призначений пристрій 510A-510J аудіокодування.

[739] Наприклад, пункт 132567-17 може бути виведений з вищенаведеного пункту 132567-1 як спосіб, який включає здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка

представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і представлення множини коефіцієнтів сферичних гармонік як функції щонайменше частини однієї або більше з матриці U , матриці S і матриці V .

[740] У порядку іншого прикладу, пункт 132567-18 може бути виведений з вищенаведеного пункту 132567-1 як пристрій, наприклад пристрій 510В аудіокодування, що містить засіб для здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і засіб для представлення множини коефіцієнтів сферичних гармонік як функції щонайменше частини однієї або більше з матриці U , матриці S і матриці V .

[741] У порядку ще одного прикладу, пункт 132567-18 може бути виведений з вищенаведеного пункту 132567-1 як довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і представляти множину коефіцієнтів сферичних гармонік як функцію щонайменше частини однієї або більше з матриці U , матриці S і матриці V .

[742] Різні пункти аналогічно можуть бути виведені з пунктів 132567-2-132567-16 для різних пристроїв, способів і довгострокових зчитуваних комп'ютером носіїв даних, виведених, як згадано вище. Те ж саме можна здійснювати для різних інших пунктів, перерахованих протягом цього розкриття.

[743] Фіг. 40С показана блок-схема, що демонструє приклад пристрою 510С аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510С аудіокодування може бути аналогічним пристрою 510В аудіокодування тим, що пристрій 510С аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510С аудіокодування може бути аналогічним блоку пристрою 510В аудіокодування тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання.

[744] Блок 512 стиснення аудіосигналу пристрою 510С аудіокодування, однак, може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510В аудіокодування тим, що блок 520 витягання компонентів звукового поля включає в себе додатковий блок, позначений як блок 532 переупорядковування векторів. З цієї причини, блок 520 витягання компонентів звукового поля пристрою 510С аудіокодування позначається як "блок 520С витягання компонентів звукового поля".

[745] Блок 532 переупорядковування векторів може представляти блок, виконаний з можливістю преупорядковування векторів $527 U_{DIST} * S_{DIST}$ для генерації переупорядкованих одного або більше векторів $533 U_{DIST} * S_{DIST}$. У цьому відношенні, блок 532 переупорядковування векторів може діяти аналогічно описаному вище відносно блока 34 переупорядковування пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. Блок 520С витягання компонентів звукового поля може викликати блок 532 переупорядковування векторів для преупорядковування векторів $527 U_{DIST} * S_{DIST}$, оскільки порядок векторів $527 U_{DIST} * S_{DIST}$ (де кожен вектор з векторів $527 U_{DIST} * S_{DIST}$ може представляти один або більше характерних монофонічних аудіооб'єктів, присутніх у звуковому полі) може відрізнятися від ділянок аудіоданих з вищезгаданої причини. Таким чином, за умови, що блок 512 стиснення аудіосигналу, у деяких прикладах, діє на цих ділянках аудіоданих, у загальному випадку, іменованих кадрами аудіо (які можуть мати M вибірок коефіцієнтів 511 сферичних гармонік, де M , у деяких прикладах, задане рівним 1024), позиція векторів, що відповідають цим окремим монофонічним аудіооб'єктам, представлених у матриці 519С U з якої виводяться вектори $527 U_{DIST} * S_{DIST}$, може розрізнятися від кадру аудіо до кадру аудіо.

[746] Передача цих векторів $527 U_{DIST} * S_{DIST}$ безпосередньо на блок 514 аудіокодування без переупорядковування цих векторів $527 U_{DIST} * S_{DIST}$ від кадру аудіо до кадру аудіо може знижувати міру стиснення, що досягається для деяких схем стиснення, наприклад успадкованих

схем стиснення, які діють краще, коли монофонічні аудіооб'єкти корелюють (поканально, що визначається в цьому прикладі порядком векторів $527 U_{DIST} * S_{DIST}$ один відносно одного) по кадрах аудіо. Крім того, за відсутності переупорядковування, кодування векторів $527 U_{DIST} * S_{DIST}$ може знижувати якість аудіоданих при відновленні. Наприклад, кодери AAC, що можуть бути

5 представлені в прикладі, наведеному на фіг. 40C, блоком 514 аудіокодування, можуть більш ефективно стискати переупорядковані один або більше векторів $533 U_{DIST} * S_{DIST}$ від кадру до кадру в порівнянні зі стисненням, що досягається при безпосередньому кодуванні векторів $527 U_{DIST} * S_{DIST}$ від кадру до кадру. Хоча вище наведений опис відносно кодерів AAC, методи можна здійснювати відносно будь-якого кодера, який забезпечує краще стиснення, коли монофонічні

10 аудіооб'єкти задаються по кадрах у конкретному порядку або позиції (поканально).

[747] Як більш докладно описано нижче, методи можуть передбачати, що пристрій 510C аудіокодування преупорядковує один або більше векторів (тобто векторів $527 U_{DIST} * S_{DIST}$) для генерації переупорядкованих одного або більше векторів вектори $533 U_{DIST} * S_{DIST}$ і, таким чином, полегшує стиснення векторів $527 U_{DIST} * S_{DIST}$ успадкованим аудіокодером, наприклад блоком 514

15 аудіокодування. Пристрій 510C аудіокодування може додатково здійснювати методи, описані в цьому винаході, для аудіокодування переупорядкованих одного або більше векторів $533 U_{DIST} * S_{DIST}$ з використанням блока 514 аудіокодування для генерації кодованої версії 515A переупорядкованих одного або більше векторів $533 U_{DIST} * S_{DIST}$.

[748] Наприклад, блок 520C витягання компонентів звукового поля може викликати блок 532 переупорядковування векторів для переупорядковування одного або більше перших векторів $527 U_{DIST} * S_{DIST}$ з першого кадру аудіо, що іде за часом за другим кадром, якому один або більше

20 других векторів $527 U_{DIST} * S_{DIST}$ відповідають. Хоча описано, що перший кадр аудіо іде за часом за другим кадром аудіо, перший кадр аудіо може передувати за часом другому кадру аудіо. Відповідно, методи не підлягають обмеженню прикладом, описаним в цьому винаході.

[749] Блок 532 переупорядковування векторів може спочатку здійснювати аналіз енергії відносно кожного з перших векторів $527 U_{DIST} * S_{DIST}$ і других векторів $527 U_{DIST} * S_{DIST}$, обчислювати середньоквадратичну енергію для щонайменше частини (але часто усього) першого кадру аудіо і частини (але часто усього) другого кадру аудіо і, таким чином, генерувати (передбачаючи D рівним чотирьом) вісім енергій, по одній для кожного з перших векторів $527 U_{DIST} * S_{DIST}$

25 $U_{DIST} * S_{DIST}$ першого кадру аудіо і по одній для кожного з других векторів $527 U_{DIST} * S_{DIST}$ другого кадру аудіо. Потім блок 532 переупорядковування векторів може порівнювати кожну енергію з перших векторів $527 U_{DIST} * S_{DIST}$ по черзі з кожним із других векторів $527 U_{DIST} * S_{DIST}$, як описано вище відносно таблиць 1-4.

[750] Іншими словами, при використанні SVD на кадровій основі (або споріднених способів, наприклад KLT і PCA) розкладання на сигналах НОА, важко гарантувати узгодженість

35 упорядкування векторів від кадру до кадру. Наприклад, якщо в звуковому полі, що лежить в основі, існує два об'єкти, розкладання (яке, коли здійснюється належним чином, може іменуватися "ідеальним розкладанням") може приводити до поділу двох об'єктів таким чином, що один вектор буде представляти один об'єкт у матриці U. Однак, навіть коли розкладання може бути позначене як "ідеальне розкладання", вектори можуть змінювати позицію в матриці U (і, відповідно, у матриці S і V) від кадру до кадру. Крім того, цілком можуть існувати різниці фаз, де блок 532 переупорядковування векторів може інвертувати фазу з використанням інверсії фази (шляхом скалярного множення кожного елемента інвертованого вектора на мінус або негативну одиницю). Для подачі цих векторів покaдрово в ту саму "машину

40 AAC/аудіокодування" може знадобитися ідентифікувати порядок (або, іншими словами, узгодити сигнали), уточнювати фазу і застосовувати акуратну інтерполяцію на границях кадру. Без цього, аудіокодек, що лежить в основі, може створювати надзвичайно різкі артефакти, у тому числі так називані 'часове розмивання' або 'випереджальне ехо'.

[751] Відповідно до різних аспектів методів, описаних у цьому винаході, пристрій 510C аудіокодування може застосовувати множинні методології для ідентифікації/узгодження векторів, з використанням енергії і крос-кореляції на границях кадру векторів. Пристрій 510C аудіокодування також може гарантувати, що зміна фази на 180 градусів, що часто виникає на границях кадру, коректується. Блок 532 переупорядковування векторів може застосовувати форму інтервалу інтерполяції плавного збільшення/зменшення між векторами, щоб гарантувати

45 плавний перехід між кадрами.

[752] Таким чином, пристрій аудіокодування 530C може переупорядковувати один або більше векторів для генерації переупорядкованих одного або більше перших векторів і, таким чином, полегшення кодування успадкованим аудіокодером, причому один або більше векторів описують характерні компоненти звукового поля, і аудіокодування переупорядкованих одного

або більше векторів з використанням успадкованого аудіокодера для генерації кодової версії переупорядкованих одного або більше векторів.

[753] Різні аспекти методів, описаних у цьому винаході, можуть передбачати, що пристрій 510С аудіокодування може діяти згідно з наступними пунктами.

[754] Пункт 133143-1А. Пристрій, наприклад пристрій 510С аудіокодування, який містить: один або більше процесорів, виконаних з можливістю здійснення порівняння енергії між одним або більше першими векторами і одним або більше другими векторами для визначення переупорядкованих одного або більше перших векторів і полегшення виділення одного або обох з одного або більше перших векторів і одного або більше других векторів, причому один або більше перших векторів описують характерні компоненти звукового поля в першій частині аудіоданих і один або більше других векторів описують характерні компоненти звукового поля в другій частині аудіоданих.

[755] Пункт 133143-2А. Пристрій за пунктом 133143-1А, у якому один або більше перших векторів не представляють фонові компоненти звукового поля в першій частині аудіоданих і при цьому один або більше других векторів не представляють фонові компоненти звукового поля в другій частині аудіоданих.

[756] Пункт 133143-3А. Пристрій за пунктом 133143-1А, у якому один або більше процесорів додатково виконані з можливістю, після здійснення порівняння енергії, здійснення крос-кореляції між одним або більше першими векторами й одним або більше другими векторами для ідентифікації одного або більше перших векторів, що корелюють з одним або більше другими векторами.

[757] Пункт 133143-4А. Пристрій за пунктом 133143-1А, у якому один або більше процесорів додатково виконані з можливістю відкидати один або більше із других векторів на основі порівняння енергії для генерації приведених одного або більше других векторів, що мають менше векторів, ніж один або більше других векторів, здійснення крос-кореляції між щонайменше одним з одного або більше перших векторів і приведеними одним або більше другими векторами для ідентифікації одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з одного або більше перших векторів, і переупорядковувати щонайменше один з одного або більше перших векторів на основі крос-кореляції для генерації переупорядкованих одного або більше перших векторів.

[758] Пункт 133143-5А. Пристрій за пунктом 133143-1А, у якому один або більше процесорів додатково виконані з можливістю відкидати один або більше із других векторів на основі порівняння енергії для генерації приведених одного або більше других векторів, що мають менше векторів, ніж один або більше других векторів, здійснення крос-кореляції між щонайменше одним з одного або більше перших векторів і приведеними одним або більше другими векторами для ідентифікації одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з одного або більше перших векторів, переупорядковувати щонайменше один з одного або більше перших векторів на основі крос-кореляції для генерації переупорядкованих одного або більше перших векторів і кодувати переупорядковані один або більше перших векторів для генерації аудіокодової версії переупорядкованих одного або більше перших векторів.

[759] Пункт 133143-6А. Пристрій за пунктом 133143-1А, у якому один або більше процесорів додатково виконані з можливістю відкидати один або більше із других векторів на основі порівняння енергії для генерації приведених одного або більше других векторів, що мають менше векторів, ніж один або більше других векторів, здійснення крос-кореляції між щонайменше одним з одного або більше перших векторів і приведеними одним або більше другими векторами для ідентифікації одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з одного або більше перших векторів, переупорядковувати щонайменше один з одного або більше перших векторів на основі крос-кореляції для генерації переупорядкованих одного або більше перших векторів, кодувати переупорядковані один або більше перших векторів для генерації аудіокодової версії переупорядкованих одного або більше перших векторів і генерації бітового потоку для включення кодової версії переупорядкованих одного або більше перших векторів.

[760] Пункт 133143-7А. Пристрій за пунктами 133143-3А-133143-6А, у якому перша частина аудіоданих містить перший кадр аудіо, що має М вибірок, причому друга частина аудіоданих містить другий кадр аудіо, що має ту ж кількість, М, вибірок, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, здійснювати крос-кореляцію відносно останніх значень М-З щонайменше одного з одного або більше перших векторів і перших значень М-З кожного з приведених одного або більше других векторів для ідентифікації

одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з одного або більше перших векторів, і при цьому Z менше, ніж M .

5 [761] Пункт 133143-8A. Пристрій за пунктами 133143-3A-133143-6A, у якому перша частина аудіоданих містить перший кадр аудіо, що має M вибірок, причому друга частина аудіоданих містить другий кадр аудіо, що має ту ж кількість, M , вибірок, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, здійснювати крос-кореляцію відносно останніх значень $M-Y$ щонайменше одного з одного або більше перших векторів і перших значень $M-Z$ кожного з приведених одного або більше других векторів для ідентифікації одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з
10 одного або більше перших векторів, і при цьому обидва Z і Y менше M .

[762] Пункт 133143-9A. Пристрій за пунктами 133143-3A-133143-6A, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, інвертувати щонайменше один з одного або більше перших векторів і одного або більше других векторів.

15 [763] Пункт 133143-10A. Пристрій за пунктом 133143-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації одного або більше перших векторів і одного або більше других векторів.

[764] Пункт 133143-11A. Пристрій за пунктом 133143-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і генерації одного або більше перших векторів і одного або більше других векторів як функції однієї або більше з матриці U , матриці S і матриці V .
20 25

[765] Пункт 133143-12A. Пристрій за пунктом 133143-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, здійснення аналізу помітності відносно матриці S для ідентифікації одного або більше векторів U_{DIST} матриці U і одного або більше векторів S_{DIST} матриці S і визначення одного або більше перших векторів і одного або більше других векторів шляхом щонайменше часткового множення одного або більше векторів U_{DIST} на один або більше векторів S_{DIST} .
30 35

[766] Пункт 133143-13A. Пристрій за пунктом 133143-1A, у якому перша частина аудіоданих передує за часом другій частині аудіоданих.

[767] Пункт 133143-14A. Пристрій за пунктом 133143-1A, у якому перша частина аудіоданих виникає через час після другої частини аудіоданих.
40

[768] Пункт 133143-15A. Пристрій за пунктом 133143-1A, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні порівняння енергії, обчислювати середньоквадратичну енергію для кожного з одного або більше перших векторів і одного або більше других векторів і порівнювати середньоквадратичну енергію, обчислену для щонайменше одного з одного або більше перших векторів, з середньоквадратичною енергією, обчисленою для кожного з одного або більше других векторів.
45

[769] Пункт 133143-16A. Пристрій за пунктом 133143-1A, у якому один або більше процесорів додатково виконані з можливістю преупорядковування щонайменше одного з одного або більше перших векторів на основі порівняння енергії для генерації переупорядкованих одного або більше перших векторів, і при цьому один або більше процесорів додатково виконані з можливістю, при переупорядкуванні перших векторів, застосовувати інтервал інтерполяції плавного збільшення/зменшення між одним або більше першими векторами, щоб гарантувати плавний перехід при генерації переупорядкованих одного або більше перших векторів.
50

[770] Пункт 133143-17A. Пристрій за пунктом 133143-1A, у якому один або більше процесорів додатково виконані з можливістю преупорядковування одного або більше перших векторів на основі щонайменше порівняння енергії для генерації переупорядкованих одного або більше перших векторів, генерації бітового потоку для включення переупорядкованих одного або більше перших векторів або кодової версії переупорядкованих одного або більше перших векторів і указання в бітовому потоці інформації переупорядковування, яка описує, як були переупорядковані один або більше перших векторів.
55

[771] Пункт 133143-18А. Пристрій за пунктом 133143-1А, у якому порівняння енергії полегшує витягання одного або обох з одного або більше перших векторів і одного або більше других векторів для сприяння аудіокодуванню одного або обох з одного або більше перших векторів і одного або більше других векторів.

5 [772] Пункт 133143-1В. Пристрій, наприклад пристрій 510С аудіокодування, який містить: один або більше процесорів, виконаних з можливістю здійснення крос-кореляції відносно одного або більше перших векторів і одного або більше других векторів для визначення переупорядкованих одного або більше перших векторів і полегшення виділення одного або обох з одного або більше перших векторів і одного або більше других векторів, причому один
10 або більше перших векторів описують характерні компоненти звукового поля в першій частині аудіоданих і один або більше других векторів описують характерні компоненти звукового поля в другій частині аудіоданих.

[773] Пункт 133143-2В. Пристрій за пунктом 133143-1В, у якому один або більше перших векторів не представляють фонові компоненти звукового поля в першій частині аудіоданих і при
15 цьому один або більше других векторів не представляють фонові компоненти звукового поля в другій частині аудіоданих.

[774] Пункт 133143-3В. Пристрій за пунктом 133143-1В, у якому один або більше процесорів додатково виконані з можливістю, до здійснення крос-кореляції, здійснення порівняння енергії між одним або більше першими векторами й одним або більше другими векторами для
20 генерації приведених одного або більше других векторів, що мають менше векторів, ніж один або більше других векторів, і при цьому один або більше процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, здійснення крос-кореляції між одним або більше першими векторами і приведеними одним або більше другими векторами для полегшення аудіокодування одного або обох з одного або більше перших векторів і одного або більше
25 других векторів.

[775] Пункт 133143-4В. Пристрій за пунктом 133143-3В, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні порівняння енергії, обчислювати середньоквадратичну енергію для кожного з одного або більше перших векторів і одного або
30 більше других векторів і порівнювати середньоквадратичну енергію, обчислену для щонайменше одного з одного або більше перших векторів, з середньоквадратичною енергією, обчисленою для кожного з одного або більше других векторів.

[776] Пункт 133143-5В. Пристрій за пунктом 133143-3В, у якому один або більше процесорів додатково виконані з можливістю відкидати один або більше із других векторів на основі порівняння енергії для генерації приведених одного або більше других векторів, що мають
35 менше векторів, ніж один або більше других векторів, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, здійснення крос-кореляції між щонайменше одним з одного або більше перших векторів і приведеними одним або більше другими векторами для ідентифікації одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з одного або більше перших векторів, і при цьому один або
40 більше процесорів додатково виконані з можливістю преупорядковування щонайменше одного з одного або більше перших векторів на основі крос-кореляції для генерації переупорядкованих одного або більше перших векторів.

[777] Пункт 133143-6В. Пристрій за пунктом 133143-3В, у якому один або більше процесорів додатково виконані з можливістю відкидати один або більше із других векторів на основі порівняння енергії для генерації приведених одного або більше других векторів, що мають
45 менше векторів, ніж один або більше других векторів, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, здійснення крос-кореляції між щонайменше одним з одного або більше перших векторів і приведеними одним або більше другими векторами для ідентифікації одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з одного або більше перших векторів, і при цьому один або
50 більше процесорів додатково виконані з можливістю преупорядковувати щонайменше один з одного або більше перших векторів на основі крос-кореляції для генерації переупорядкованих одного або більше перших векторів і кодувати переупорядковані один або більше перших векторів для генерації аудіокодуваної версії переупорядкованих одного або більше перших
55 векторів.

[778] Пункт 133143-7В. Пристрій за пунктом 133143-3В, у якому один або більше процесорів додатково виконані з можливістю відкидати один або більше із других векторів на основі порівняння енергії для генерації приведених одного або більше других векторів, що мають
60 менше векторів, ніж один або більше других векторів, у якому один або більше процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, здійснення крос-кореляції між

щонайменше одним з одного або більше перших векторів і приведеними одним або більше
 другими векторами для ідентифікації одного з приведених одного або більше других векторів,
 які корелюють з щонайменше одним з одного або більше перших векторів, і при цьому один або
 більше процесорів додатково виконані з можливістю переупорядковувати щонайменше один з
 5 одного або більше перших векторів на основі крос-кореляції для генерації переупорядкованих
 одного або більше перших векторів, кодувати переупорядковані один або більше перших
 векторів для генерації аудіокодованої версії переупорядкованих одного або більше перших
 векторів і генерувати бітовий потік для включення кодованої версії переупорядкованих одного
 або більше перших векторів.

[779] Пункт 133143-8В. Пристрій за пунктами 133143-3В-133143-7В, у якому перша частина
 аудіоданих містить перший кадр аудіо, що має M вибірок, причому друга частина аудіоданих
 містить другий кадр аудіо, що має ту ж кількість, M , вибірок, у якому один або більше процесорів
 додатково виконані з можливістю, при здійсненні крос-кореляції, здійснювати крос-кореляцію
 відносно останніх значень $M-Z$ щонайменше одного з одного або більше перших векторів і
 15 перших значень $M-Z$ кожного з приведених одного або більше других векторів для ідентифікації
 одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з
 одного або більше перших векторів, і при цьому Z менше, ніж M .

[780] Пункт 133143-9В. Пристрій за пунктами 133143-3В-133143-7В, у якому перша частина
 аудіоданих містить перший кадр аудіо, що має M вибірок, причому друга частина аудіоданих
 20 містить другий кадр аудіо, що має ту ж кількість, M , вибірок, у якому один або більше процесорів
 додатково виконані з можливістю, при здійсненні крос-кореляції, здійснювати крос-кореляцію
 відносно останніх значень $M-Y$ щонайменше одного з одного або більше перших векторів і
 перших значень $M-Z$ кожного з приведених одного або більше других векторів для ідентифікації
 одного з приведених одного або більше других векторів, які корелюють з щонайменше одним з
 25 одного або більше перших векторів, і при цьому обидва Z і Y менше M .

[781] Пункт 133143-10В. Пристрій за пунктом 133143-1В, у якому один або більше
 процесорів додатково виконані з можливістю, при здійсненні крос-кореляції, інвертувати
 щонайменше один з одного або більше перших векторів і одного або більше других векторів.

[782] Пункт 133143-11В. Пристрій за пунктом 133143-1В, у якому один або більше
 30 процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення
 відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для
 генерації одного або більше перших векторів і одного або більше других векторів.

[783] Пункт 133143-12В. Пристрій за пунктом 133143-1В, у якому один або більше
 процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення
 35 відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для
 генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних
 гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних
 гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів
 сферичних гармонік, і генерації одного або більше перших векторів і одного або більше других
 40 векторів як функції однієї або більше з матриці U , матриці S і матриці V .

[784] Пункт 133143-13В. Пристрій за пунктом 133143-1В, у якому один або більше
 процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення
 відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для
 генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних
 45 гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних
 гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів
 сферичних гармонік, здійснення аналізу помітності відносно матриці S для ідентифікації одного
 або більше векторів U_{DIST} матриці U і одного або більше векторів S_{DIST} матриці S і визначення
 одного або більше перших векторів і одного або більше других векторів шляхом щонайменше
 50 часткового множення одного або більше векторів U_{DIST} на один або більше векторів S_{DIST} .

[785] Пункт 133143-14В. Пристрій за пунктом 133143-1В, у якому один або більше
 процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення
 відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для
 генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних
 55 гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних
 гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів
 сферичних гармонік, і, при визначенні одного або більше перших векторів і одного або більше
 других векторів, здійснення аналізу помітності відносно матриці S для ідентифікації одного або
 більше векторів V_{DIST} матриці V як щонайменше одного з одного або більше перших векторів і
 60 одного або більше других векторів.

[786] Пункт 133143-15В. Пристрій за пунктом 133143-1В, у якому перша частина аудіоданих передує за часом другій частині аудіоданих.

[787] Пункт 133143-16В. Пристрій за пунктом 133143-1В, у якому перша частина аудіоданих виникає через час після другої частини аудіоданих.

5 [788] Пункт 133143-17В. Пристрій за пунктом 133143-1В, у якому один або більше процесорів додатково виконані з можливістю преупорядковувати щонайменше один з одного або більше перших векторів на основі крос-кореляції для генерації переупорядкованих одного або більше перших векторів і, при переупорядкуванні перших векторів, застосовувати інтервал інтерполяції плавного збільшення/зменшення між одним або більше першими векторами, щоб
10 гарантувати плавний перехід при генерації переупорядкованих одного або більше перших векторів.

[789] Пункт 133143-18В. Пристрій за пунктом 133143-1В, у якому один або більше процесорів додатково виконані з можливістю преупорядкування одного або більше перших векторів на основі щонайменше крос-кореляції для генерації переупорядкованих одного або
15 більше перших векторів, генерації бітового потоку для включення переупорядкованих одного або більше перших векторів або кодованої версії переупорядкованих одного або більше перших векторів і указання в бітовому потоці, як були переупорядковані один або більше перших векторів.

[790] Пункт 133143-19В. Пристрій за пунктом 133143-1В, у якому крос-кореляція полегшує
20 витягання одного або обох з одного або більше перших векторів і одного або більше других векторів для сприяння аудіокодуванню одного або обох з одного або більше перших векторів і одного або більше других векторів.

[791] На фіг. 40D показана блок-схема, що демонструє приклад пристрою 510D аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для
25 стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510D аудіокодування може бути аналогічним пристрою 510С аудіокодування тим, що пристрій 510D аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510D аудіокодування може бути аналогічним блоку пристрою 510С аудіокодування
30 тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання.

[792] Блок 512 стиснення аудіосигналу пристрою 510D аудіокодування, однак, може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510С аудіокодування тим, що блок
35 520 витягання компонентів звукового поля включає в себе додатковий блок, позначений як блок 534 квантування ("блок 534 квантування"). З цієї причини, блок 520 витягання компонентів звукового поля пристрою 510D аудіокодування позначається як "блок 520D витягання компонентів звукового поля".

[793] Блок 534 квантування представляє блок, виконаний з можливістю квантування одного або більше векторів $525E \ V_{DIST}^T$ і/або одного або більше векторів $525F \ V_{BG}^T$ для генерації
40 відповідних одного або більше векторів $525G \ V_{Q_DIST}^T$ і/або одного або більше вектори $525H \ V_{Q_BG}^T$. Блок 534 квантування може квантувати (що є членом обробки сигналу для математичного округлення за допомогою усунення бітів, використовуваних для представлення значення) один або більше векторів $525E \ V_{DIST}^T$ для зниження кількості бітів, використовуваних для представлення одного або більше векторів $525E \ V_{DIST}^T$ у бітовому потоці 517. У деяких прикладах, блок 534 квантування може квантувати 32-бітові значення одного або більше
45 векторів $525E \ V_{DIST}^T$, замінюючи ці 32-бітові значення округленими 16-бітовими значеннями для генерації одного або більше векторів $525G \ V_{Q_DIST}^T$. У цьому відношенні, блок 534 квантування може діяти аналогічно описаному вище відносно блока 52 квантування пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4.

[794] Квантування такого роду може вносити помилку в представлення звукового поля, що
50 змінюється відповідно до грубості квантування. Іншими словами, чим більше бітів використовується для представлення одного або більше векторів $525E \ V_{DIST}^T$, тим менше помилка квантування. Помилку квантування внаслідок квантування векторів $525E \ V_{DIST}^T$ (яка може бути позначена " E_{DIST} ") можна визначити відніманням одного або більше векторів $525E \ V_{DIST}^T$ з одного або більше векторів $525G \ V_{Q_DIST}^T$.

[795] Відповідно до методів, описаних у цьому винаході, пристрій 510D аудіокодування може
55 компенсувати одну або більше з помилок квантування E_{DIST} шляхом проектування помилки E_{DIST} на або іншої зміни одного або більше з векторів $527 \ U_{DIST} \ S_{DIST}$ або коефіцієнтів 531 сферичних гармонік фону, генерованих шляхом множення одного або більше векторів $525D \ U_{BG}$ на один або більше векторів $525B \ S_{BG}$ і потім на один або більше векторів $525F \ V_{BG}^T$. У деяких
60 прикладах, пристрій 510D аудіокодування може компенсувати тільки помилку E_{DIST} у векторах

527 $U_{DIST} * S_{DIST}$. В інших прикладах, пристрій 510D аудіокодування може компенсувати тільки помилку E_{BG} у коефіцієнтах сферичних гармонік фону. В інших прикладах, пристрій 510D аудіокодування може компенсувати помилку E_{DIST} як у векторах 527 $U_{DIST} * S_{DIST}$, так і в коефіцієнтах сферичних гармонік фону.

5 [796] У ході експлуатації, блок 524 аналізу помітних компонентів може бути виконаний з можливістю виведення одного або більше векторів 525A S_{DIST} , одного або більше векторів 525B S_{BG} , одного або більше векторів 525C U_{DIST} , одного або більше векторів 525D U_{BG} , одного або більше векторів 525E V_{DIST}^T і одного або більше векторів 525F V_{BG}^T на математичний блок 526. Блок 524 аналізу помітних компонентів також може виводити один або більше векторів 525E V_{DIST}^T на блок 534 квантування. Блок 534 квантування може квантувати один або більше векторів 525E V_{DIST}^T для генерації одного або більше векторів 525G $V_{Q_DIST}^T$. Блок 534 квантування може забезпечувати один або більше векторів 525G $V_{Q_DIST}^T$ на математичний блок 526, також забезпечуючи один або більше векторів 525G $V_{Q_DIST}^T$ на блок 532 переупорядковування векторів (як описано вище). Блок 532 переупорядковування векторів може діяти відносно одного або більше векторів 525G $V_{Q_DIST}^T$ аналогічно описаному вище відносно векторів 525E V_{DIST}^T .

20 [797] Після прийому цих векторів 525A-525G ("векторів 525"), математичний блок 526 може спочатку визначати характерні коефіцієнти сферичних гармонік, що описують характерні компоненти звукового поля, і коефіцієнти сферичних гармонік фону, що описують фонові компоненти звукового поля. Блок 526 матричної математики може бути виконаний з можливістю визначення характерних коефіцієнтів сферичних гармонік шляхом множення одного або більше векторів 525C U_{DIST} на один або більше векторів 525A S_{DIST} і потім на один або більше векторів 525E V_{DIST}^T . Математичний блок 526 може бути виконаний з можливістю визначення коефіцієнтів сферичних гармонік фону шляхом множення одного або більше векторів 525D U_{BG} на один або більше векторів 525A S_{BG} і потім на один або більше векторів 525E V_{BG}^T .

30 [798] Потім математичний блок 526 може визначати один або більше скомпенсованих векторів 527' $U_{DIST} * S_{DIST}$ (які можуть бути аналогічні векторам 527 $U_{DIST} * S_{DIST}$ за винятком того, що ці вектори включають в себе значення для компенсації помилки E_{DIST}) шляхом здійснення операції псевдообернення відносно одного або більше векторів 525G $V_{Q_DIST}^T$ і потім множення характерних сферичних гармонік на псевдообернене одного або більше векторів 525G $V_{Q_DIST}^T$. Блок 532 переупорядковування векторів може діяти вищеописаним чином для генерації переупорядкованих векторів 527', які потім аудіокодується блоком 515A аудіокодування для генерації аудіокодованих переупорядкованих векторів 515', знову ж, як описано вище.

35 [799] Потім математичний блок 526 може проектувати помилку E_{DIST} у коефіцієнти сферичних гармонік фону. Математичний блок 526 може, для здійснення цієї проєкції, визначати або інакше виявляти вихідні коефіцієнти 511 сферичних гармонік шляхом додавання характерних коефіцієнтів сферичних гармонік до коефіцієнтів сферичних гармонік фону. Потім математичний блок 526 може віднімати квантованні характерні коефіцієнти сферичних гармонік (які можуть генеруватися шляхом множення векторів 525C U_{DIST} на вектори 525A S_{DIST} і потім на вектори 525G $V_{Q_DIST}^T$) і коефіцієнти сферичних гармонік фону з коефіцієнтів 511 сферичних гармонік для визначення помилки, що залишилася, внаслідок квантування векторів 519 V_{DIST}^T . Потім математичний блок 526 може додавати цю помилку до квантованих коефіцієнтів сферичних гармонік фону для генерації скомпенсованих квантованих коефіцієнтів 531' сферичних гармонік фону.

45 [800] У будь-якому випадку, блок 528A зниження порядку може здійснювати, як описано вище, скорочення скомпенсованих квантованих коефіцієнтів 531' сферичних гармонік фону до приведених коефіцієнтів 529' сферичних гармонік фону, які можуть аудіокодуватися блоком 514 аудіокодування вищеописаним чином, для генерації аудіокодованих приведених коефіцієнтів 515B' сферичних гармонік фону.

50 [801] Таким чином, методи можуть передбачати, що пристрій 510D аудіокодування квантує один або більше перших векторів, наприклад вектори 525E V_{DIST}^T , що представляють один або більше компонентів звукового поля, і компенсує помилку, внесену внаслідок квантування одного або більше перших векторів в один або більше других векторів, наприклад векторів 527 $U_{DIST} * S_{DIST}$ і/або векторів коефіцієнтів 531 сферичних гармонік фону, які також представляють ті ж один або більше компонентів звукового поля.

55 [802] Крім того, методи можуть забезпечувати компенсацію цієї помилки квантування згідно з наступними пунктами.

60 [803] Пункт 133146-1B. Пристрій, наприклад пристрій 510D аудіокодування, який містить: один або більше процесорів, виконаних з можливістю квантування одного або більше перших векторів, що представляють один або більше характерних компонентів звукового поля, і

компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, які також представляють одні і ті ж один або більше характерних компонентів звукового поля.

5 [804] Пункт 133146-2В. Пристрій за пунктом 133146-1В, у якому один або більше процесорів виконані з можливістю квантування одного або більше векторів із транспонованої матриці V , генерованої, щонайменше частково, шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що описують звукове поле.

10 [805] Пункт 133146-3В. Пристрій за пунктом 133146-1В, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів виконані з можливістю квантування одного
15 або більше векторів із транспонованої матриці V .

[806] Пункт 133146-4В. Пристрій за пунктом 133146-1В, у якому один або більше процесорів виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка
20 представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, у якому один або більше процесорів виконані з можливістю квантування одного або більше векторів із транспонованої матриці V , і при цьому один або більше процесорів виконані з можливістю компенсації помилки, що вноситься внаслідок квантування в один або більше векторів U^*S ,
25 обчислених шляхом множення одного або більше векторів U матриці U на один або більше векторів S матриці S .

[807] Пункт 133146-5В. Пристрій за пунктом 133146-1В, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів U_{DIST} матриці U , кожний з яких відповідає
30 одному з характерних компонентів звукового поля, визначення одного або більше векторів S_{DIST} матриці S , кожний з яких відповідає одному і тому ж одному з характерних компонентів звукового поля, і визначення одного або більше векторів V_{DIST}^T транспонованої матриці V , кожний з яких відповідає одному і тому ж одному з характерних компонентів звукового поля,
35 [808] у якому один або більше процесорів виконані з можливістю квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{\text{Q_DIST}}^T$, і при цьому один або більше процесорів виконані з можливістю компенсації помилки, що вноситься внаслідок квантування в один або більше векторів $U_{\text{DIST}}^*S_{\text{DIST}}$, обчислених шляхом множення одного або більше векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S , для генерації одного або більше векторів $U_{\text{DIST}}^*S_{\text{DIST}}$ з компенсацією помилки.

[809] Пункт 133146-6В. Пристрій за пунктом 133146-5В, у якому один або більше процесорів виконані з можливістю визначення характерних коефіцієнтів сферичних гармонік на основі
45 одного або більше векторів U_{DIST} , одного або більше векторів S_{DIST} і одного або більше векторів V_{DIST}^T і здійснення псевдообернення відносно векторів $V_{\text{Q_DIST}}^T$ для поділу характерних коефіцієнтів сферичних гармонік на один або більше векторів $V_{\text{Q_DIST}}^T$ і, таким чином, генерації одного або більше векторів $U_{\text{C_DIST}}^*S_{\text{C_DIST}}$ з компенсацією помилки, які компенсують,
50 щонайменше частково, помилку, що вноситься за допомогою квантування векторів V_{DIST}^T .

[810] Пункт 133146-7В. Пристрій за пунктом 133146-5В, у якому один або більше процесорів додатково виконані з можливістю аудіокодування одного або більше векторів $U_{\text{DIST}}^*S_{\text{DIST}}$ з компенсацією помилки.

55 [811] Пункт 133146-8В. Пристрій за пунктом 133146-1В, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів U_{BG} матриці U , що описують один або більше
60

фонових компонентів звукового поля, і одного або більше векторів U_{DIST} матриці U , що описують один або більше характерних компонентів звукового поля, визначення одного або більше векторів S_{BG} матриці S , що описують один або більше фонових компонентів звукового поля, і одного або більше векторів S_{DIST} матриці S , що описують один або більше характерних компонентів звукового поля, і визначення одного або більше векторів V_{DIST}^T і одного або більше векторів V_{BG}^T транспонованої матриці V , причому вектори V_{DIST}^T описують один або більше характерних компонентів звукового поля і V_{BG}^T описують один або більше фонових компонентів звукового поля, у якому один або більше процесорів виконані з можливістю квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{Q_DIST}^T$, і у якому один або більше процесорів додатково виконані з можливістю компенсації щонайменше частини помилки, що вноситься внаслідок квантування в коефіцієнти сферичних гармонік фону, сформовані шляхом множення одного або більше векторів U_{BG} на один або більше векторів S_{BG} і потім на один або більше векторів V_{BG}^T , для генерації коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[812] Пункт 133146-9B. Пристрій за пунктом 133146-8B, у якому один або більше процесорів виконані з можливістю визначення помилки на основі векторів V_{DIST}^T і одного або більше векторів $U_{DIST} \cdot S_{DIST}$, сформованих шляхом множення векторів U_{DIST} на вектори S_{DIST} , і додавання визначеної помилки до коефіцієнтів сферичних гармонік фону для генерації коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[813] Пункт 133146-10B. Пристрій за пунктом 133146-8B, у якому один або більше процесорів додатково виконані з можливістю аудіокодування коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[814] Пункт 133146-11B. Пристрій за пунктом 133146-1B,

[815] у якому один або більше процесорів виконані з можливістю компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля, для генерації одного або більше других векторів з компенсацією помилки, і при цьому один або більше процесорів додатково виконані з можливістю генерації бітового потоку для включення одного або більше других векторів з компенсацією помилки і квантованих одного або більше перших векторів.

[816] Пункт 133146-12B. Пристрій за пунктом 133146-1B, у якому один або більше процесорів виконані з можливістю компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля, для генерації одного або більше других векторів з компенсацією помилки, і при цьому один або більше процесорів додатково виконані з можливістю аудіокодування одного або більше других векторів з компенсацією помилки і генерації бітового потоку для включення аудіокодованих одного або більше других векторів з компенсацією помилки і квантованих одного або більше перших векторів.

[817] Пункт 133146-1C. Пристрій, наприклад пристрій 510D аудіокодування, який містить: один або більше процесорів, виконаних з можливістю квантування одного або більше перших векторів, що представляють один або більше характерних компонентів звукового поля, і компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що представляють один або більше фонових компонентів звукового поля.

[818] Пункт 133146-2C. Пристрій за пунктом 133146-1C, у якому один або більше процесорів виконані з можливістю квантування одного або більше векторів із транспонованої матриці V , генерованої, щонайменше частково, шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[819] Пункт 133146-3C. Пристрій за пунктом 133146-1C, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів виконані з можливістю квантування одного або більше векторів із транспонованої матриці V .

[820] Пункт 133146-4C. Пристрій за пунктом 133146-1C, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік,

матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів U_{DIST} матриці U , кожний з яких відповідає одному з характерних компонентів звукового поля, визначення одного або більше векторів S_{DIST} матриці S , кожний з яких відповідає одному і тому ж одному з характерних компонентів звукового поля, і визначення одного або більше векторів V_{DIST}^T транспонованої матриці V , кожний з яких відповідає одному і тому ж одному з характерних компонентів звукового поля, у якому один або більше процесорів виконані з можливістю квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{Q_DIST}^T$ і компенсації щонайменше частини помилки, що вноситься внаслідок квантування в один або більше векторів $U_{DIST}^T S_{DIST}$, обчислених шляхом множення одного або більше векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S , для генерації одного або більше векторів $U_{DIST}^T S_{DIST}$ з компенсацією помилки.

[821] Пункт 133146-5C. Пристрій за пунктом 133146-4C, у якому один або більше процесорів виконані з можливістю визначення характерних коефіцієнтів сферичних гармонік на основі одного або більше векторів U_{DIST} , одного або більше векторів S_{DIST} і одного або більше векторів V_{DIST}^T і здійснення псевдообернення відносно векторів $V_{Q_DIST}^T$ для поділу характерних коефіцієнтів сферичних гармонік на один або більше векторів $V_{Q_DIST}^T$ і, таким чином, генерування одного або більше векторів $U_{C_DIST}^T S_{C_DIST}$, які компенсують, щонайменше частково, помилку, внесену за допомогою квантування векторів V_{DIST}^T .

[822] Пункт 133146-6C. Пристрій за пунктом 133146-4C, у якому один або більше процесорів додатково виконані з можливістю аудіокодування одного або більше векторів $U_{DIST}^T S_{DIST}$ з компенсацією помилки.

[823] Пункт 133146-7C. Пристрій за пунктом 133146-1C, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів U_{BG} матриці U , що описують один або більше фонових компонентів звукового поля, і одного або більше векторів U_{DIST} матриці U , що описують один або більше характерних компонентів звукового поля, визначення одного або більше векторів S_{BG} матриці S , що описують один або більше фонових компонентів звукового поля, і одного або більше векторів S_{DIST} матриці S , що описують один або більше характерних компонентів звукового поля, і визначення одного або більше векторів V_{DIST}^T і одного або більше векторів V_{BG}^T транспонованої матриці V , причому вектори V_{DIST}^T описують один або більше характерних компонентів звукового поля і V_{BG}^T описують один або більше фонових компонентів звукового поля, у якому один або більше процесорів виконані з можливістю квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{Q_DIST}^T$, і при цьому один або більше процесорів виконані з можливістю компенсації помилки, що вноситься внаслідок квантування в коефіцієнти сферичних гармонік фону, сформовані шляхом множення одного або більше векторів U_{BG} на один або більше векторів S_{BG} і потім на один або більше векторів V_{BG}^T , для генерації коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[824] Пункт 133146-8C. Пристрій за пунктом 133146-7C, у якому один або більше процесорів виконані з можливістю визначення помилки на основі векторів V_{DIST}^T і одного або більше векторів $U_{DIST}^T S_{DIST}$, сформованих шляхом множення векторів U_{DIST} на вектори S_{DIST} , і додавання визначеної помилки до коефіцієнтів сферичних гармонік фону для генерації коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[825] Пункт 133146-9C. Пристрій за пунктом 133146-7C, у якому один або більше процесорів додатково виконані з можливістю аудіокодування коефіцієнтів сферичних гармонік фону з компенсацією помилки.

[826] Пункт 133146-10C. Пристрій за пунктом 133146-1C, у якому один або більше процесорів додатково виконані з можливістю компенсації помилки, що вноситься внаслідок квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля, для генерації одного або більше других векторів з компенсацією помилки і генерації бітового потоку для включення одного або більше других векторів з компенсацією помилки і квантованих одного або більше перших векторів.

[827] Пункт 133146-11C. Пристрій за пунктом 133146-1C, у якому один або більше процесорів додатково виконані з можливістю компенсації помилки, що вноситься внаслідок

квантування одного або більше перших векторів в один або більше других векторів, що також представляють ті ж один або більше компонентів звукового поля, для генерації одного або більше других векторів з компенсацією помилки, аудіокодування одного або більше других векторів з компенсацією помилки і генерації бітового потоку для включення аудіокодованих одного або більше других векторів з компенсацією помилки і квантованих одного або більше перших векторів.

[828] Іншими словами, при використанні SVD на кадровій основі (або споріднених способів, наприклад KLT і PCA) розкладання на сигналах НОА з метою зменшення ширини смуги, методи, описані в цьому винаході, можуть передбачати, що пристрій 510D аудіокодування квантує перші декілька векторів матриці U (помножені на відповідні сингулярні значення матриці S), а також відповідні вектори вектора V . Вони будуть містити 'фонові' або 'характерні' компоненти звукового поля. Потім методи можуть передбачати, що пристрій 510D аудіокодування кодує вектори $U \cdot S$ з використанням машини аудіокодування типу "чорний ящик", наприклад кодера AAC. Вектор V може піддаватися скалярному або векторному квантуванню.

[829] Крім того, деякі з векторів, що залишилися, у матриці U можна множити на відповідні сингулярні значення матриці S і матриці V і також кодувати з використанням машини аудіокодування типу "чорний ящик". Вони будуть містити 'фонові' компоненти звукового поля. Просте 16-бітове скалярне квантування векторів V може приводити до службового навантаження приблизно 80 кбіт/с для 4-го порядку (25 коефіцієнтів) і 160 кбіт/с для 6-го порядку (49 коефіцієнтів). Більш грубе квантування може приводити до більш значних помилок квантування. Методи, описані в цьому винаході, можуть компенсувати помилку квантування векторів V шляхом 'проекування' помилки квантування вектора V на компоненти переднього плану і фону.

[830] Методи в цьому винаході можуть включати в себе обчислення квантованої версії фактичного вектора V . Цей квантований вектор V може іменуватися V' (де $V' = V + e$). Сигнал НОА, що лежить в основі, для компонентів переднього плану, які методи намагаються відтворити, задається як $H_f = USV$, де U , S і V містять тільки елементи переднього плану. З метою цього розгляду, US буде замінена одиничним набором векторів U . Таким чином, $H_f = UV$. За умови, що ми маємо помилковий V' , методи намагаються відтворити H_f якомога точніше. Таким чином, методи можуть передбачати, що пристрій 510D аудіокодування знаходить U' , таким чином, що $H_f = U'V'$. Пристрій 510D аудіокодування може використовувати методологію псевдообернення, що дозволяє $U' = H_f [V']^{\wedge}(-1)$. З використанням так названої машини аудіокодування типу "чорний ящик" для кодування U' , методи можуть мінімізувати помилку в H , обумовлену так названим помилковим вектором V' .

[831] Аналогічним чином, методи також можуть передбачати, що пристрій аудіокодування проектує помилку внаслідок квантування V у фонові елементи. Пристрій 510D аудіокодування може бути виконаний з можливістю відтворення повного сигналу НОА, що є комбінацією сигналів НОА переднього плану і фону, тобто $H = H_f + H_b$. Це, знову ж, можна моделювати як $H = H_f + e + H_b$, внаслідок помилки квантування в V' . Таким чином, замість подачі H_b на 'аудіокодер типу чорного ящика', ми подаємо $(e + H_b)$ на аудіокодер, що фактично компенсує помилку в V' . На практиці, це дозволяє компенсувати помилку тільки до порядку, визначеного пристроєм 510D аудіокодування для відправлення для фонових елементів.

[832] На фіг. 40Е показана блок-схема, що демонструє приклад пристрою 510Е аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510Е аудіокодування може бути аналогічним пристрою 510D аудіокодування тим, що пристрій 510Е аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510Е аудіокодування може бути аналогічним блоку пристрою 510D аудіокодування тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання.

[833] Блок 512 стиснення аудіосигналу пристрою 510Е аудіокодування, однак, може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510D аудіокодування тим, що математичний блок 526 блока 520 витягання компонентів звукового поля здійснює додаткові аспекти методів, описаних у цьому винаході, для додаткового приведення матриця 519А V до включення приведеної версії транспонованої матриці 519А V у бітовий потік 517. З цієї причини, блок 520 витягання компонентів звукового поля пристрою 510Е аудіокодування позначається як "блок 520Е витягання компонентів звукового поля".

[834] У прикладі, наведеному на фіг. 40Е, блок 528 зниження порядку, замість пересилання приведених коефіцієнтів 529' сферичних гармонік фону на блок 514 аудіокодування, повертає приведені коефіцієнти 529' сферичних гармонік фону на математичний блок 526. Як згадано

вище, ці приведені коефіцієнти 529' сферичних гармонік фону можуть бути скорочені шляхом видалення коефіцієнтів, які відповідають сферичним базисним функціям, що мають один або більше ідентифікованих порядків і/або підпорядків. Знижений порядок приведених коефіцієнтів 529' сферичних гармонік фону може бути позначений змінною N_{BG} .

5 [835] За умови, що блок 520E витягання компонентів звукового поля може не здійснювати зниження порядку відносно переупорядкованих одного або більше векторів 533' $U_{DIST}^T S_{DIST}$, порядок цього розкладання коефіцієнтів сферичних гармонік, що описують характерні компоненти звукового поля (який може бути позначений змінною N_{DIST}) може бути більше порядку фону, N_{BG} . Іншими словами, N_{BG} звичайно може бути менше N_{DIST} . Одна можлива

10 причина того, що N_{BG} може бути менше N_{DIST} , полягає в припущенні, що фонові компоненти не мають великої направленості, завдяки чому сферичні базисні функції більш високого порядку не вимагаються, що дозволяє знижувати порядок і приводить до того, що N_{BG} менше N_{DIST} .

15 [836] За умови, що переупорядковані один або більше векторів 539 $V_{Q_DIST}^T$ раніше відправлялися відкрито, без аудіокодування цих векторів 539 у бітовому потоці 517, як показано в прикладах, наведених на фіг. 40A-40D, переупорядковані один або більше векторів 539 $V_{Q_DIST}^T$ можуть споживати значну ширину смуги. У порядку одного прикладу, кожний з переупорядкованих одного або більше векторів 539 $V_{Q_DIST}^T$, при квантуванні в 16-бітові скалярні значення, може споживати приблизно 20 кбіт/с для аудіоданих амбіофонії четвертого порядку (де кожен вектор має 25 коефіцієнтів) і 40 кбіт/с для аудіоданих амбіофонії шостого порядку (де кожен вектор має 49 коефіцієнтів).

20 [837] Відповідно до різних аспектів методів, описаних у цьому винаході, блок 520E витягання компонентів звукового поля може знижувати кількість бітів, які потрібно вказувати для коефіцієнтів сферичних гармонік або їх розкладань, наприклад переупорядкованих одного або більше векторів 539 $V_{Q_DIST}^T$. У деяких прикладах, математичний блок 526 може визначати, на

25 основі коефіцієнтів 529' сферичних гармонік зниженого порядку, ті з переупорядкованих векторів 539 $V_{Q_DIST}^T$, які підлягають видаленню і рекомендації з коефіцієнтами 529' сферичних гармонік зниженого порядку, і ті з переупорядкованих векторів 539 $V_{Q_DIST}^T$, які служать для формування векторів 521 V_{SMALL}^T . Таким чином, математичний блок 526 може визначати порядок коефіцієнтів 529' сферичних гармонік зниженого порядку, де цей порядок може бути

30 позначений N_{BG} . Переупорядковані вектори 539 $V_{Q_DIST}^T$ можуть мати порядок, позначений змінною N_{DIST} , де N_{DIST} більше, ніж порядок N_{BG} .

[838] Потім математичний блок 526 може виділяти перші N_{BG} порядків переупорядкованих векторів 539 $V_{Q_DIST}^T$, видаляти вектори, які вказують розкладені коефіцієнти сферичних гармонік, що відповідають сферичним базисним функціям, які мають порядок, що менше або

35 дорівнює N_{BG} . Ці видалені переупорядковані вектори 539 $V_{Q_DIST}^T$ потім можна використовувати для формування проміжних коефіцієнтів сферичних гармонік шляхом множення переупорядкованих векторів 533' $U_{DIST}^T S_{DIST}$, які представляють розкладені версії коефіцієнтів 511 сферичних гармонік, що відповідають сферичним базисним функціям, які мають порядок, що менше або дорівнює N_{BG} , на видалені переупорядковані вектори 539 $V_{Q_DIST}^T$ для

40 формування проміжних характерних коефіцієнтів сферичних гармонік. Потім математичний блок 526 може генерувати модифіковані коефіцієнти 537 сферичних гармонік фону шляхом додавання проміжних характерних коефіцієнтів сферичних гармонік до коефіцієнтів 529' сферичних гармонік зниженого порядку. Потім математичний блок 526 може передавати ці модифіковані коефіцієнти 537 сферичних гармонік фону на блок 514 аудіокодування, який

45 аудіокодує ці коефіцієнти 537 для формування аудіокодованих модифікованих коефіцієнтів 515B' сферичних гармонік фону.

[839] Потім математичний блок 526 може передавати один або більше векторів 521 V_{SMALL}^T , що можуть представляти вектори 539, які представляють розкладену форму коефіцієнтів 511 сферичних гармонік, що відповідають сферичним базисним функціям, які мають порядок, що

50 більше N_{BG} і менше або дорівнює N_{DIST} . У цьому відношенні, математичний блок 526 може здійснювати операції, аналогічні операціям блока 46 приведення коефіцієнтів пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4. Математичний блок 526 може передавати один або більше векторів 521 V_{SMALL}^T на блок 516 генерації бітового потоку, який може генерувати бітовий потік 517 для включення векторів 521 V_{SMALL}^T часто у свій вихідний неаудіокодованій формі. За умови, що вектори 521 V_{SMALL}^T включає в себе менше векторів, ніж переупорядковані вектори 539 $V_{Q_DIST}^T$, методи можуть полегшувати виділення меншої кількості бітів переупорядкованим векторам 539 $V_{Q_DIST}^T$ за рахунок укавання тільки векторів 521 V_{SMALL}^T у бітовому потоці 517.

60 [840] Хоча вони не показані як квантовані, у ряді випадків, пристрій 510E аудіокодування може квантувати вектори 525F V_{BG}^T . У ряді випадків, наприклад, коли блок 514 аудіокодування

не використовується для стиснення коефіцієнтів сферичних гармонік фону, пристрій 510E аудіокодування може квантувати вектори $525F V_{BG}^T$.

[841] Таким чином, методи можуть передбачати, що пристрій 510E аудіокодування визначає щонайменше один з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, для рекомбінації з коефіцієнтами сферичних гармонік фону для зниження кількості бітів, які необхідно виділяти одному або більше векторам у бітовому потоці, причому коефіцієнти сферичних гармонік описують звукове поле і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[842] Таким чином, методи можуть передбачати, що пристрій 510E аудіокодування сконфігурований згідно з наступними пунктами.

[843] Пункт 133149-1A. Пристрій, наприклад пристрій 510E аудіокодування, який містить: один або більше процесорів, виконаних з можливістю визначення щонайменше одного з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, для рекомбінації з коефіцієнтами сферичних гармонік фону для зниження кількості бітів, які необхідно виділяти одному або більш векторам у бітовому потоці, причому коефіцієнти сферичних гармонік описують звукове поле, і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[844] Пункт 133149-2A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю генерації скороченого набору з одного або більше векторів шляхом видалення визначених щонайменше одного з одного або більше векторів.

[845] Пункт 133149-3A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю генерації скороченого набору з одного або більше векторів шляхом видалення визначених щонайменше одного з одного або більше векторів, рекомбінації видалених щонайменше одного з одного або більше векторів з коефіцієнтами сферичних гармонік фону для генерації модифікованих коефіцієнтів сферичних гармонік фону і генерації бітового потоку для включення скороченого набору з одного або більше векторів і модифікованих коефіцієнтів сферичних гармонік фону.

[846] Пункт 133149-4A. Пристрій за пунктом 133149-3A, у якому скорочений набір з одного або більше векторів включений у бітовий потік без початкового аудіокодування.

[847] Пункт 133149-5A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю генерації скороченого набору з одного або більше векторів шляхом видалення визначених щонайменше одного з одного або більше векторів, рекомбінації видалених щонайменше одного з одного або більше векторів з коефіцієнтами сферичних гармонік фону для генерації модифікованих коефіцієнтів сферичних гармонік фону, аудіокодування модифікованих коефіцієнтів сферичних гармонік фону і генерації бітового потоку для включення скороченого набору з одного або більше векторів і аудіокодованих модифікованих коефіцієнтів сферичних гармонік фону.

[848] Пункт 133149-6A. Пристрій за пунктом 133149-1A, у якому один або більше векторів містять вектори, що представляють щонайменше деякий аспект одного або більше характерних компонентів звукового поля.

[849] Пункт 133149-7A. Пристрій за пунктом 133149-1A, у якому один або більше векторів містять один або більше векторів із транспонованої матриці V , генерованої, щонайменше частково, шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[850] Пункт 133149-8A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і при цьому один або більше векторів містять один або більше векторів з транспонованої матриці V .

[851] Пункт 133149-9A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення зниження порядку відносно коефіцієнтів сферичних гармонік фону для видалення тих з коефіцієнтів сферичних гармонік фону, що відповідають сферичним базисним функціям, які мають ідентифікований порядок i /або підпорядок, причому коефіцієнти сферичних гармонік фону відповідають порядку N_{BG} .

[852] Пункт 133149-10A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення зниження порядку відносно коефіцієнтів сферичних гармонік фону для видалення тих з коефіцієнтів сферичних гармонік фону, що відповідають сферичним базисним функціям, які мають ідентифікований порядок i /або

підпорядок, причому коефіцієнти сферичних гармонік фону відповідають порядку N_{BG} , який менше, ніж порядок характерних коефіцієнтів сферичних гармонік, N_{DIST} , і при цьому характерні коефіцієнти сферичних гармонік представляють характерні компоненти звукового поля.

[853] Пункт 133149-11A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення зниження порядку відносно коефіцієнтів сферичних гармонік фону для видалення тих з коефіцієнтів сферичних гармонік фону, що відповідають сферичним базисним функціям, які мають ідентифікований порядок і/або підпорядок, причому коефіцієнти сферичних гармонік фону відповідають порядку N_{BG} , який менше, ніж порядок характерних коефіцієнтів сферичних гармонік, N_{DIST} , і при цьому характерні коефіцієнти сферичних гармонік представляють характерні компоненти звукового поля і не піддаються зниженню порядку.

[854] Пункт 133149-12A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і визначення одного або більше векторів V_{DIST}^T і одного або більше векторів V_{BG}^T транспонованої матриці V , причому один або більше векторів V_{DIST}^T описують один або більше характерних компонентів звукового поля і один або більше векторів V_{BG}^T описують один або більше фонових компонентів звукового поля, і при цьому один або більше векторів включають в себе один або більше векторів V_{DIST}^T .

[855] Пункт 133149-13A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів V_{DIST}^T і одного або більше векторів V_{BG}^T транспонованої матриці V , причому один або більше векторів V_{DIST}^T описують один або більше характерних компонентів звукового поля і один або більше векторів V_{BG}^T описують один або більше фонових компонентів звукового поля, і квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{Q_DIST}^T$, і при цьому один або більше векторів включають в себе один або більше векторів $V_{Q_DIST}^T$.

[856] Пункт 133149-14A. Пристрій за будь-яким з пунктів 133149-12A або 133149-13A, у якому один або більше процесорів додатково виконані з можливістю визначення одного або більше векторів U_{DIST} і одного або більше векторів U_{BG} матриці U , причому один або більше векторів U_{DIST} описують один або більше характерних компонентів звукового поля і один або більше векторів U_{BG} описують один або більше фонових компонентів звукового поля, і визначення одного або більше векторів S_{DIST} і одного або більше векторів S_{BG} матриці S , причому один або більше векторів S_{DIST} описують один або більше характерних компонентів звукового поля і один або більше векторів S_{BG} описують один або більше фонових компонентів звукового поля.

[857] Пункт 133149-15A. Пристрій за пунктом 133149-14A, у якому один або більше процесорів додатково виконані з можливістю визначення коефіцієнтів сферичних гармонік фону як функції одного або більше векторів U_{BG} , одного або більше векторів S_{BG} і одного або більше V_{BG}^T , здійснення зниження порядку відносно коефіцієнтів сферичних гармонік фону для генерації приведених коефіцієнтів сферичних гармонік фону, що мають порядок, який дорівнює N_{BG} , множення одного або більше векторів U_{DIST} на один або більше векторів S_{DIST} для генерації одного або більше векторів $U_{DIST} * S_{DIST}$, видалення визначеного щонайменше одного з одного або більше векторів з одного або більше векторів для генерації скороченого набору з одного або більше векторів, множення одного або більше векторів $U_{DIST} * S_{DIST}$ на видалений щонайменше один з одного або більше векторів V_{DIST}^T або одного або більше векторів $V_{Q_DIST}^T$ для генерації проміжних характерних коефіцієнтів сферичних гармонік і додавання проміжних характерних коефіцієнтів сферичних гармонік до коефіцієнта сферичної гармоніки фону для рекомбінації видалених щонайменше одного з одного або більше векторів V_{DIST}^T або одного або більше векторів $V_{Q_DIST}^T$ з коефіцієнтами сферичних гармонік фону.

[858] Пункт 133149-16A. Пристрій за пунктом 133149-14A, у якому один або більше процесорів додатково виконані з можливістю визначення коефіцієнтів сферичних гармонік фону як функції одного або більше векторів U_{BG} , одного або більше векторів S_{BG} і одного або більше V_{BG}^T , здійснення зниження порядку відносно коефіцієнтів сферичних гармонік фону для

генерації приведених коефіцієнтів сферичних гармонік фону, що мають порядок, який дорівнює N_{BG} , множення одного або більше векторів U_{DIST} на один або більше векторів S_{DIST} для генерації одного або більше векторів $U_{DIST} * S_{DIST}$, переупорядкування одного або більше векторів $U_{DIST} * S_{DIST}$ для генерації переупорядкованих одного або більше векторів $U_{DIST} * S_{DIST}$, видалення визначеного щонайменше одного з одного або більше векторів з одного або більше векторів для генерації скороченого набору з одного або більше векторів, множення переупорядкованих одного або більше векторів $U_{DIST} * S_{DIST}$ на видалений щонайменше один з одного або більше векторів V_{DIST}^T або одного або більше векторів $V_{Q_DIST}^T$ для генерації проміжних характерних коефіцієнтів сферичних гармонік і додавання проміжних характерних коефіцієнтів сферичних гармонік до коефіцієнта сферичної гармоніки фону для рекомбінації видалених щонайменше одного з одного або більше векторів V_{DIST}^T або одного або більше векторів $V_{Q_DIST}^T$ з коефіцієнтами сферичних гармонік фону.

[859] Пункт 133149-17A. Пристрій за будь-яким з пунктів 133149-15A або 133149-16A, у якому один або більше процесорів додатково виконані з можливістю аудіокодування коефіцієнтів сферичних гармонік фону після додавання проміжних характерних коефіцієнтів сферичних гармонік до коефіцієнтів сферичних гармонік фону і генерації бітового потоку для включення аудіокодованих коефіцієнтів сферичних гармонік фону.

[860] Пункт 133149-18A. Пристрій за пунктом 133149-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, визначення одного або більше векторів V_{DIST}^T і одного або більше векторів V_{BG}^T транспонованої матриці V , причому один або більше векторів V_{DIST}^T описують один або більше характерних компонентів звукового поля і один або більше векторів V_{BG}^T описують один або більше фонових компонентів звукового поля, квантування одного або більше векторів V_{DIST}^T для генерації одного або більше векторів $V_{Q_DIST}^T$ і переупорядкування одного або більше векторів $V_{Q_DIST}^T$ для генерації переупорядкованих одного або більше векторів $V_{Q_DIST}^T$, і при цьому один або більше векторів включають в себе переупорядковані один або більше векторів $V_{Q_DIST}^T$.

[861] На фіг. 40F показана блок-схема, що демонструє ілюстративний пристрій 510F аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510F аудіокодування може бути аналогічним пристрою 510C аудіокодування тим, що пристрій 510F аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510F аудіокодування може бути аналогічним блоку пристрою 510C аудіокодування тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання і блок 532 переупорядкування векторів, які можуть діяти аналогічно відповідним блокам пристрою 510C аудіокодування. У деяких прикладах, пристрій 510F аудіокодування може включати в себе блок 534 квантування, як описано з посиланням на фіг. 40D і 40E, для квантування одного або більше векторів з будь-яких векторів 525C U_{DIST} , векторів 525D U_{BG} , векторів 525E V_{DIST}^T і векторів 525J V_{BG}^T .

[862] Блок 512 стиснення аудіосигналу пристрою 510F аудіокодування, однак, може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510C аудіокодування тим, що блок 524 аналізу помітних компонентів блока 520 витягання компонентів звукового поля може здійснювати аналіз контенту для вибору кількості компонентів переднього плану, позначений як D у контексті фіг. 40A-40J. Іншими словами, блок 524 аналізу помітних компонентів може діяти відносно матриць 519 U , S і V вищеописаним чином для ідентифікації, чи згенеровані розкладені версії коефіцієнтів сферичних гармонік із синтетичних аудіооб'єктів або з природного запису за допомогою мікрофона. Потім блок 524 аналізу помітних компонентів може визначати D на основі цього синтетичного визначення.

[863] Крім того, блок 512 стиснення аудіосигналу пристрою 510F аудіокодування може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510C аудіокодування тим, що блок 520 витягання компонентів звукового поля може включати в себе додатковий блок, блок 528F зниження порядку і збереження енергії (проілюстрований як "блок 528F зниження порядку і збереження енергії"). З цих причин, блок 520 витягання компонентів звукового поля пристрою 510F аудіокодування позначається як "блок 520F витягання компонентів звукового поля".

[864] Блок 528F зниження порядку і збереження енергії представляє блок, виконаний з можливістю здійснення зниження порядку фонових компонентів матриці 525H V_{BG} , що

представляють праві сингулярні вектори з множини коефіцієнтів 511 сферичних гармонік, і при цьому збереження повної енергії (і забруднюючого звукового тиску) звукового поля, описаного частково повною матрицею 525H V_{BG} . У цьому відношенні, блок 528F зниження порядку і збереження енергії може здійснювати операції, аналогічні описаним вище відносно блока 48 вибору фону і блока 38 компенсації енергії пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4.

[865] Повна матриця 525H V_{BG} має розмірність $(N+1)^2 \times (N+1)^2 - D$, де D представляє кількість головних компонентів або, іншими словами, сингулярних значень, що визначаються як помітні в тому розумінні, що вони є характерними аудіокомпонентами звукового поля. Таким чином, повна матриця 525H V_{BG} включає в себе ті сингулярні значення, що визначаються як фонові (BG) або, іншими словами, оточуючі або нехарактерні аудіокомпоненти звукового поля.

[866] Як описано вище відносно, наприклад, блока 524 зниження порядку на фіг. 40B-40E, блок 528F зниження порядку і збереження енергії може видаляти, виключати або інакше усувати (часто шляхом обнулення) фонових сингулярних значень матриці 525H V_{BG} , що відповідають сферичним базисним функціям більш високого порядку. Блок 528F зниження порядку і збереження енергії може виводити приведену версію матриці 525H V_{BG} (позначеної як "матриця 525I V_{BG}' " і іменованої далі "приведеною матрицею 525I V_{BG}' ") на блок 522 транспонування. Приведена матриця 525I V_{BG}' може мати розмірність $(\eta+1)^2 \times (N+1)^2 - D$, де $\eta < N$. Блок 522 транспонування застосовує операцію транспонування до приведеної матриці 525I V_{BG}' для генерації і виведення транспонованої приведеної матриці 525J $V_{BG}'^T$ на математичний блок 526, що може діяти для реконструкції фонових звукових компонентів звукового поля шляхом обчислення $U_{BG} * S_{BG} * V_{BG}'^T$ з використанням матриці 525D U_{BG} , матриці 525B S_{BG} і транспонованої приведеної матриці 525J $V_{BG}'^T$.

[867] Відповідно до описаних тут методів, блок 528F зниження порядку і збереження енергії додатково виконаний з можливістю компенсації можливого зниження повної енергії фонових звукових компонентів звукового поля, обумовленого зниженням порядку повної матриці 525H V_{BG} для генерації приведеної матриці 525I V_{BG}' . У деяких прикладах, блок 528F зниження порядку і збереження енергії компенсує шляхом визначення коефіцієнта підсилення компенсації у формі значень посилення для застосування до кожного з $(N+1)^2 - D$ стовпців приведеної матриці 525I V_{BG}' для збільшення середньоквадратичної (RMS) енергії приведеної матриці 525I V_{BG}' , щоб вони дорівнювали або щонайменше були більш наближені до RMS повної матриці 525H V_{BG} , до виведення приведеної матриці 525I V_{BG}' на блок 522 транспонування.

[868] У ряді випадків, блок 528F зниження порядку і збереження енергії може визначати енергію RMS кожного стовпця повної матриці 525H V_{BG} і енергію RMS кожного стовпця приведеної матриці 525I V_{BG}' , потім визначати значення посилення для стовпця як відношення першої до останнього, як зазначено в наступному рівнянні:

$$\alpha = v_{BG} / v_{BG}',$$

[869] де α - значення посилення для стовпця, v_{BG} представляє одиничний стовпець матриці 525H V_{BG} , і v_{BG}' представляє відповідний одиничний стовпець матриці 525I V_{BG}' . Це може бути представлено в матричній формі як:

$$A = V_{BG}^{RMS} / V_{BG}'^{RMS},$$

$$A = [\alpha_1 \dots \alpha_{(N+1)^2 - D}],$$

де V_{BG}^{RMS} - вектор RMS, елементи якого позначають RMS кожного стовпця матриці 525H

V_{BG} , $V_{BG}'^{RMS}$ - вектор RMS, елементи якого позначають RMS кожного стовпця приведеної матриці 525I V_{BG}' , і A - вектор значень посилення, що має елементи для кожного стовпця матриці 525H V_{BG} . Блок 528F зниження порядку і збереження енергії застосовує скалярне множення до кожного стовпця приведеної матриці 525I V_{BG}' з використанням відповідного значення посилення, α , або у векторній формі:

$$V_{BG}'' = V_{BG}' A^T,$$

[870] де V_{BG}'' представляє приведену матрицю 525I V_{BG}'' , що включає в себе компенсацію енергії. Блок 528F зниження порядку і збереження енергії може виводити приведену матрицю 525I V_{BG}'' , що включає в себе компенсацію енергії, на блок 522 транспонування для вирівнювання (або приблизного вирівнювання) RMS приведеної матриці 525I V_{BG}'' з RMS повної

матриці 525H V_{BG} . Вихідна розмірність приведеної матриці 525I V_{BG}' , що включає в себе компенсацію енергії, може дорівнювати $(\hat{n}+1)^2 \times (N+1)^2 - D$.

[871] У деяких прикладах, для визначення кожного RMS відповідних стовпців приведеної матриці 525I V_{BG}' і повної матриці 525H V_{BG} , блок 528F зниження порядку і збереження енергії може спочатку застосовувати рендеризатор опорних коефіцієнтів сферичних гармонік (SHC) до стовпців. Застосування рендеризатора опорних SHC блоком 528F зниження порядку і збереження енергії дозволяє визначати RMS в області SHC для визначення енергії загального звукового поля, описаного кожним стовпцем кадру, представленого приведеною матрицею 525I V_{BG}' і повною матрицею 525H V_{BG} . Таким чином, у таких прикладах, блок 528F зниження порядку і збереження енергії може застосовувати рендеризатор опорних SHC до кожного стовпця повної матриці 525H V_{BG} і до кожного скороченого стовпця приведеної матриці 525I V_{BG}' , визначати відповідні значення RMS для стовпця і скороченого стовпця і визначати значення посилення для стовпця як відношення значення RMS для стовпця до значення RMS для скороченого стовпця. У деяких прикладах, зниження порядку до приведеної матриці 525I V_{BG}' забезпечує

```
R=ReferenceRenderer;
for m=numDist+1: numChannels
fullV=V(:,m); //бере один стовпець V=>fullV
reducedV =[fullV(1:numBG); zeros(numChannels-numBG, 1)];
alpha=sqrt(sum((fullV'*R).^2)/ sum ((reducedV'*R).^2));
if isnan(alpha)||isinf(alpha), alpha=1; end;
V_out(:,m)=reducedV*alpha;
end
```

[872] У вищенаведеному псевдокоді, numChannels може представляти $(N+1)^2 - D$, numBG може представляти $(\hat{n}+1)^2$, V може представляти матрицю 525H V_{BG} , і V_out може представляти приведену матрицю 525I V_{BG}' , і R може представляти рендеризатор опорних SHC блока 528F зниження порядку і збереження енергії. Розмірність V може дорівнювати $(N+1)^2 \times (N+1)^2 - D$, і розмірність V_out може дорівнювати $(\hat{n}+1)^2 \times (N+1)^2 - D$.

[873] У результаті, пристрій 510F аудіокодування може, при представленні множини коефіцієнтів 511 сферичних гармонік, реконструювати фонові звукові компоненти з використанням матриці 525I V_{BG}' зниженого порядку, що включає в себе компенсацію можливої втрати енергії в результаті процесу зниження порядку.

[874] На фіг. 40G показана блок-схема, що демонструє ілюстративний пристрій 510G аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. У прикладі, наведеному на фіг. 40G, пристрій 510G аудіокодування включає в себе блок 520F витягання компонентів звукового поля. У свою чергу, блок 520F витягання компонентів звукового поля включає в себе блок 524G аналізу помітних компонентів.

[875] Блок 512 стиснення аудіосигналу пристрою 510G аудіокодування, однак, може відрізнятися від блока 512 стиснення аудіосигналу пристрою аудіокодування 510F тим, що блок 512 стиснення аудіосигналу пристрою 510G аудіокодування включає в себе блок 524G аналізу помітних компонентів. Блок 524G аналізу помітних компонентів може представляти блок, виконаний з можливістю визначення помітності або характерності аудіоданих, що представляють звукове поле, з використанням інформації на основі направленості, пов'язаної з аудіоданими.

[876] Хоча визначення на основі енергії можуть поліпшувати рендеризацію звукового поля, розкладеного за допомогою SVD для ідентифікації характерних аудіокомпонентів звукового поля, визначення на основі енергії також можуть приводити до того, що пристрій помилково ідентифікує фонові аудіокомпоненти як характерні аудіокомпоненти, у випадках, коли фонові аудіокомпоненти демонструють високий рівень енергії. Таким чином, поділ характерних і фонових аудіокомпонентів тільки на основі енергії може бути ненадійним, оскільки енергетичні (наприклад, більш голосні) фонові аудіокомпоненти можуть неправильно ідентифікуватися як характерні аудіокомпоненти. Щоб більш надійно розрізняти між характерними і фоновими аудіокомпонентами звукового поля, різні аспекти методів, описаних у цьому винаході, можуть передбачати блок 524G аналізу помітних компонентів для здійснення аналізу на основі направленості SHC 511 для виділення характерних і фонових аудіокомпонентів з розкладених версій SHC 511.

[877] Блок 524G аналізу помітних компонентів може, у прикладі, наведеному на фіг. 40H, представляти блок, сконфігурований або інакше виконаний з можливістю виділення характерних (або передньопланових) елементів з фонових елементів, включених в одну або

більше з матриці 519A V, матриці 519B S і матриці 519C U, аналогічно блокам 524 аналізу помітних компонентів раніше описаних пристроїв 510-510F аудіокодування. Відповідно до деяких методів на основі SVD, найбільш енергетичні компоненти (наприклад, перші декілька векторів однієї або більше з матриць 519A-519C V, S і U або матриці, виведеної з них) можна розглядати як характерні компоненти. Однак, найбільш енергетичні компоненти (які представлені векторами) однієї або більше з матриць 519A-519C можуть не у всіх сценаріях представляти найбільш направлені компоненти/сигнали.

[878] На відміну від раніше описаних блоків 524 аналізу помітних компонентів, блок 524G аналізу помітних компонентів може реалізувати один або більше аспектів методів, описаних тут, для ідентифікації елементів переднього плану на основі направленості векторів однієї або більше з матриць 519A-519C або матриці, виведеної з них. У деяких прикладах, блок 524G аналізу помітних компонентів може ідентифікувати або вибрати як характерні аудіокомпоненти (де компоненти також можуть іменуватися "об'єктами") один або більше векторів на основі енергії і направленості векторів. Наприклад, блок 524G аналізу помітних компонентів може ідентифікувати ті вектори однієї або більше з матриць 519A-519C (або матриці, виведеної з них), які демонструють високу енергію і високу направленість (наприклад, представлену як показник направленості) як характерні аудіокомпоненти. В результаті, якщо блок 524G аналізу помітних компонентів визначає, що конкретний вектор є відносно менш направленим у порівнянні з іншими векторами однієї або більше з матриць 519A-519C (або матриці, виведеної з них), то, незалежно від рівня енергії, пов'язаного з конкретним вектором, блок 524G аналізу помітних компонентів може визначити, що конкретний вектор представляє аудіокомпоненти фону (або оточення) звукового поля, представленого SHC 511. У цьому відношенні, блок 524G аналізу помітних компонентів може здійснювати операції, аналогічні описаним вище відносно блока 44 аналізу звукового поля пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4.

[879] У деяких реалізаціях, блок 524G аналізу помітних компонентів може ідентифікувати характерні аудіооб'єкти (які, як згадано вище, також можуть іменуватися "компонентами") на основі направленості, шляхом здійснення наступних операцій. Блок 524G аналізу помітних компонентів може множити (наприклад, з використанням одного або більше процесів матричного множення) матрицю 519A V на матрицю 519B S. Шляхом перемножування матриці 519A V і матриці 519B S, блок 524G аналізу помітних компонентів може одержувати матрицю VS. Додатково, блок 524G аналізу помітних компонентів може підносити до квадрата (тобто підносити до степеня два) щонайменше деякі з елементів кожного з векторів (який може бути рядком) матриці VS. У ряді випадків, блок 524G аналізу помітних компонентів може підсумовувати квадрати елементів кожного вектора, які пов'язані з порядком, що дорівнює 1. У порядку одного прикладу, якщо кожен вектор матриці включає в себе 25 елементів, блок 524G аналізу помітних компонентів може, відносно кожного вектора, підносити до квадрата елементи кожного вектора, починаючи з п'ятого елемента і закінчуючи двадцять п'ятим елементом, підсумовуючи квадрати елементів для визначення показника направленості (або індикатора направленості). Кожна операція підсумовування може давати показник направленості для відповідного вектора. У цьому прикладі, блок 524G аналізу помітних компонентів може визначити, що ті елементи кожного рядка, що пов'язані з порядком, який менше або дорівнює 1, а саме елементи з першого по четвертий, більшою мірою стосуються величини енергії й меншою мірою направленості цих елементів. Таким чином, амбіофонія більш низького порядку, пов'язана з порядком нуль або один, відповідають сферичним базисним функціям, які, як показано на фіг. 1 і фіг. 2, не забезпечують багато відносно напрямку хвилі тиску, але проте забезпечують деяку голосність (яка представляє енергію).

[880] Операції, описані у вищенаведеному прикладі, також можуть бути виражені відповідно до наступного псевдокоду. Нижченаведений псевдокод включає в себе анотації, у формі операторів коментарів, що поміщені між послідовними екземплярами рядків символів "/*" і "*/" (без лапок).

```
[U, S, V]=svd(audioframe, 'ecom');
VS=V*S;
```

/* Наступний рядок стосується незалежного аналізу кожного рядка і підсумовування значень у першому (у порядку одного прикладу) рядку від п'ятого елемента до двадцять п'ятого елемента для визначення показника направленості або метрики направленості для відповідного вектора. Піднести до квадрата елементи до підсумовування. Елементи в кожному рядку, які пов'язані з порядком, що більше 1, пов'язані з амбіофонією більш високого порядку і, таким чином, з більшою імовірністю є направленими. */

```
sumVS=sum(VS(5:end, :).^2,1);
```

/* Наступний рядок стосується сортування суми квадратів для генерованої матриці VS і вибору набору найбільших значень (наприклад, трьох або чотирьох найбільших значень). */

```
[~,idxVS]=sort(sumVS, 'descend');
```

```
U=U(:,idxVS);
```

```
5 V=V(:,idxVS);
```

```
S=S(idxVS, idxVS);
```

[881] Іншими словами, відповідно до вищенаведеного псевдокоду, блок 524G аналізу помітних компонентів може вибирати елементи кожного вектора матриці VS, розкладеної з SHC 511, що відповідає сферичній базисній функції, яка має порядок, що більше одиниці. Потім блок 524G аналізу помітних компонентів може підносити до квадрата ці елементи для кожного вектора матриці VS, підсумовувати квадрати елементів для ідентифікації, обчислювати або інакше визначати метрику або показник направленості для кожного вектора матриці VS. Потім, блок 524G аналізу помітних компонентів може сортувати вектори матриці VS на основі відповідних метрик направленості кожного з векторів. Блок 524G аналізу помітних компонентів може сортувати ці вектори в порядку убутання метрик направленості, таким чином, що вектори з найвищою відповідною направленістю є першими і вектори з найнижчою відповідною направленістю є останніми. Блок 524G аналізу помітних компонентів може потім вибирати ненульовий піднабір векторів, що мають найвищу відносну метрику направленості.

[882] Відповідно до деяких аспектів описаних тут методів, пристрій 510G аудіокодування, або один або більше його компонентів, може ідентифікувати або інакше використовувати попередньо визначену кількість векторів матриці VS як характерні аудіокомпоненти. Наприклад, після вибору елементів з 5 по 25 кожного рядка матриці VS і піднесення до квадрата і підсумовування вибраних елементів для визначення відносної метрики направленості для кожного відповідного вектора, блок 524G аналізу помітних компонентів може реалізувати додатковий вибір серед векторів для ідентифікації векторів, що представляють характерні аудіокомпоненти. У деяких прикладах, блок 524G аналізу помітних компонентів може вибирати попередньо визначену кількість векторів матриці VS, порівнюючи показники направленості векторів. У порядку одного прикладу, блок 524G аналізу помітних компонентів може вибирати чотири вектори, представлені в матриці VS, які мають чотири найвищих показники направленості (і які є першими чотирма векторами сортованої матриці VS). У свою чергу, блок 524G аналізу помітних компонентів може визначити, що чотири вибрані вектори представляють чотири найбільш характерних аудіооб'єкта, пов'язаних з відповідним представленням SHC звукового поля.

[883] У деяких прикладах, блок 524G аналізу помітних компонентів може переупорядковувати вектори, виведені з матриці VS, для відображення характерності чотирьох вибраних векторів, як описано вище. В одному прикладі, блок 524G аналізу помітних компонентів може переупорядковувати вектори таким чином, що чотири вибраних елементи переміщуються у верхню частину матриці VS. Наприклад, блок 524G аналізу помітних компонентів може змінювати матрицю VS таким чином, що всі чотири вибраних елементи розташовуються в першому (або найбільш верхньому) рядку результуючої переупорядкованої матриці VS. Хоча воно описано тут відносно блока 524G аналізу помітних компонентів, у різних реалізаціях, переупорядкування можуть здійснювати інші компоненти пристрою 510G аудіокодування, наприклад блок 532 переупорядкування векторів.

[884] Блок 524G аналізу помітних компонентів може передавати результуючу матрицю (тобто матрицю VS, переупорядковану або ні, що може бути) на блок 516 генерації бітового потоку. У свою чергу, блок 516 генерації бітового потоку може використовувати матрицю 525K VS для генерації бітового потоку 517. Наприклад, якщо блок 524G аналізу помітних компонентів переупорядкував матрицю 525K VS, блок 516 генерації бітового потоку може використовувати верхній рядок переупорядкованої версії матриці 525K VS як характерні аудіооб'єкти, наприклад, за допомогою квантування або відкидання векторів, що залишилися, переупорядкованої версії матриці 525K VS. За допомогою квантування векторів, що залишилися, переупорядкованої версії матриці 525K VS, блок 16 генерації бітового потоку може розглядати вектори, що залишилися, як оточуючі або фонові аудіодані.

[885] У прикладах, де блок 524G аналізу помітних компонентів не переупорядковував матрицю 525K VS, блок 516 генерації бітового потоку може відрізнати характерні аудіодані від фонових аудіоданих, на основі конкретних елементів (наприклад, елементів з 5-го по 25-й) кожного рядка матриці 525K VS, як вибраних блоком 524G аналізу помітних компонентів. Наприклад, блок 516 генерації бітового потоку може генерувати бітовий потік 517 за допомогою квантування або відкидання перших чотирьох елементів кожного рядка матриці 525K VS.

[886] Таким чином, пристрій 510G аудіокодування і/або його компоненти, наприклад блок 524G аналізу помітних компонентів, можуть реалізувати методи цього винаходу для визначення або іншого використання відношень енергій більш високих і більш низьких коефіцієнтів аудіоданих, для розрізнення між характерними аудіооб'єктами і фоновими аудіоданими, що представляють звукове поле. Наприклад, як описано, блок 524G аналізу помітних компонентів може використовувати відношення енергії на основі значень різних елементів матриці 525K VS, генерованої блоком 524H аналізу помітних компонентів. Шляхом об'єднання даних, забезпечених матрицею 519A V і матрицею 519B S, блок 524G аналізу помітних компонентів може генерувати матрицю 525K VS для забезпечення інформації про направленість і повну енергію різних компонентів аудіоданих, у формі векторів і відповідних даних (наприклад, показників направленості). Зокрема, матриця 519A V може забезпечувати інформацію, що стосується визначень направленості, тоді як матриця 519B S може забезпечувати інформацію, що стосується визначень повної енергії для компонентів аудіоданих.

[887] В інших прикладах, блок 524G аналізу помітних компонентів може генерувати матрицю 525K VS з використанням переупорядкованих векторів $539 V_{\text{DIST}}^T$. У цих прикладах, блок 524G аналізу помітних компонентів може визначати характерність на основі матриці 519A V, до будь-якої зміни на основі матриці 519B S. Іншими словами, згідно з цими прикладами, блок 524G аналізу помітних компонентів може визначати направленість з використанням тільки матриці 519A V, не здійснюючи етап генерації матриці 525K VS. Зокрема, матриця 519A V може забезпечувати інформацію про те, яким чином змішуються компоненти (наприклад, вектори матриці 519A V) аудіоданих, і, потенційно, інформацію про різні синергетичні ефекти даних, що переносяться векторами. Наприклад, матриця 519A V може забезпечувати інформацію про "напрямок надходження" різних аудіокомпонентів, представлених векторами, наприклад напрямком надходження кожного аудіокомпонента, що пересилається на пристрій 510G аудіокодування за допомогою EigenMike®. Використовуваний тут термін "компонент аудіоданих" можна використовувати взаємозамінно з "елементом" будь-якої з матриць 519 або будь-яких виведених з них матриць.

[888] Відповідно до деяких реалізацій методів цього винаходу, блок 524G аналізу помітних компонентів може поповнювати або доповнювати представлення SHC сторонньою інформацією для виробництва різних описаних тут визначень. У порядку одного прикладу, блок 524G аналізу помітних компонентів може доповнювати SHC сторонньою інформацією для визначення помітності різних аудіокомпонентів, представлених у матрицях 519A-519C. У порядку іншого прикладу, блок 524G аналізу помітних компонентів і/або блок 532 переупорядковування векторів може доповнювати НОА сторонніми даними для розрізнення між характерними аудіооб'єктами і фоновими аудіоданими.

[889] У деяких прикладах, блок 524G аналізу помітних компонентів може виявляти, що ділянки (наприклад, характерні аудіооб'єкти) аудіоданих відображають кейнсіанську енергію. Приклад таких характерних об'єктів може бути пов'язаний з людським голосом, який модулюють. У випадку аудіоданих на основі голосу, які модулюють, блок 524G аналізу помітних компонентів може визначити, що енергія модулюючих даних, як відношення до енергій компонентів, що залишилися, залишається приблизно постійною (наприклад, постійна в пороговому діапазоні) або приблизно нерухомою протягом часу. Традиційно, якщо енергетичні характеристики характерних аудіокомпонентів з кейнсіанською енергією (наприклад пов'язаних з модулюючим голосом) змінюються від одного кадру аудіо до іншого, пристрій може бути не здатний ідентифікувати послідовність аудіокомпонентів як єдиний сигнал. Однак блок 524G аналізу помітних компонентів може реалізувати методи цього винаходу для визначення направленості або апертури об'єкта відстані, представленої як вектор у різних матрицях.

[890] Зокрема, блок 524G аналізу помітних компонентів може визначити, що характеристики, наприклад направленість і/або апертура, навряд чи суттєво змінюються по кадрах аудіо. Використовуваний тут термін апертура представляє відношення коефіцієнтів більш високого порядку до коефіцієнтів більш низького порядку в аудіоданих. Кожен рядок матриці 519A V може включати в себе вектори, які відповідають конкретному SHC. Блок 524G аналізу помітних компонентів може визначити, що SHC більш низького порядку (наприклад, пов'язаному з порядком, який менше або дорівнює 1) властиво представляти дані оточення, тоді як елементам більш високого порядку властиво представляти характерні дані. Додатково, блок 524G аналізу помітних компонентів може визначити, що, у багатьох випадках, SHC більш високого порядку (наприклад, пов'язані з порядком, який більше 1) відображають більш високі енергії і що відношення енергії SHC більш високого порядку до SHC більш низького порядку залишається, по суті, аналогічним (або приблизно постійним) від кадру аудіо до кадру аудіо.

[891] Один або більше компонентів блока 524G аналізу помітних компонентів можуть визначати характеристики аудіоданих, наприклад направленість і апертуру, з використанням матриці 519A V. Таким чином, компоненти пристрою 510G аудіокодування, наприклад блок 524G аналізу помітних компонентів, можуть реалізувати описані тут методи для визначення помітності і/або розрізнення характерних аудіооб'єктів від фонового аудіо, з використанням інформації на основі направленості. З використанням направленості для визначення помітності і/або характерності, блок 524G аналізу помітних компонентів може досягати більш надійних визначень, ніж у випадках пристрою, виконаного з можливістю визначення помітності і/або характерності з використанням тільки даних на основі енергії. Хоча він описаний вище відносно визначень помітності і/або характерності на основі направленості, блок 524G аналізу помітних компонентів може реалізувати методи цього винаходу для використання направленості крім інших характеристик, наприклад енергії, для визначення помітності і/або характерності конкретних компонентів аудіоданих, що представлені векторами однієї або більше з матриць 519A-519C (або будь-якої виведеної з них матриці).

[892] У деяких прикладах, спосіб включає в себе ідентифікацію одного або більше характерних аудіооб'єктів з одного або більше коефіцієнтів сферичних гармонік (SHC), пов'язаних з аудіооб'єктами, на основі направленості, визначеної для одного або більше з аудіооб'єктів. В одному прикладі, спосіб додатково включає в себе визначення направленості одного або більше аудіооб'єктів на основі коефіцієнтів сферичних гармонік, пов'язаних з аудіооб'єктами. У деяких прикладах, спосіб додатково включає в себе здійснення розкладання на сингулярні значення відносно коефіцієнтів сферичних гармонік для генерації матриці U, яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S, яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V, яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік; і представлення множини коефіцієнтів сферичних гармонік як функції щонайменше частини однієї або більше з матриці U, матриці S і матриці V, причому визначення відповідної направленості одного або більше аудіооб'єктів базується, щонайменше частково, на матриці V.

[893] В одному прикладі, спосіб додатково включає в себе переупорядкування одного або більше векторів матриці V таким чином, що вектори, що мають більш високий показник направленості, розташовуються над векторами, що мають більш низький показник направленості, у переупорядкованій матриці V. В одному прикладі, спосіб додатково включає в себе визначення, що вектори, які мають більш високий показник направленості, включають в себе більше інформації направленості, ніж вектори, які мають більш низький показник направленості. В одному прикладі, спосіб додатково включає в себе множення матриці V на матрицю S для генерації матриці VS, причому матриця VS включає в себе один або більше векторів. В одному прикладі, спосіб додатково включає в себе вибір елементів кожного рядка матриці VS, які пов'язані з порядком, що більше 1, піднесення до квадрата кожного з вибраних елементів для формування відповідних квадратів елементів і, для кожного рядка матриці VS, підсумовування квадратів всіх елементів для визначення показника направленості для відповідного вектора.

[894] У деяких прикладах, кожен рядок матриці VS включає в себе 25 елементів. В одному прикладі, вибір елементів кожного рядка матриці VS, пов'язаної з порядком, що більше 1, включає в себе вибір всіх елементів, починаючи з 5-го елемента кожного рядка матриці VS і закінчуючи 25-им елементом кожного рядка матриці VS. В одному прикладі, спосіб додатково включає в себе вибір піднабору векторів матриці VS для представлення характерних аудіооб'єктів. У деяких прикладах, вибір піднабору включає в себе вибір чотирьох векторів матриці VS, і вибрані чотири вектори мають чотири найбільших показники направленості з усіх векторів матриці VS. В одному прикладі, визначення, що вибраний піднабір векторів представляє характерні аудіооб'єкти, здійснюється на основі направленості й енергії кожного вектора.

[895] У деяких прикладах, спосіб включає в себе ідентифікацію одного або більше характерних аудіооб'єктів з одного або більше коефіцієнтів сферичних гармонік, пов'язаних з аудіооб'єктами, на основі направленості й енергії, визначеної для одного або більше з аудіооб'єктів. В одному прикладі, спосіб додатково включає в себе визначення одного або обох з направленості й енергії одного або більше аудіооб'єктів на основі коефіцієнтів сферичних гармонік, пов'язаних з аудіооб'єктами. У деяких прикладах, спосіб додатково включає в себе здійснення розкладання на сингулярні значення відносно коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U, яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S, яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V, яка представляє праві сингулярні

вектори з множини коефіцієнтів сферичних гармонік, і представлення множини коефіцієнтів сферичних гармонік як функції щонайменше частини однієї або більше з матриці U, матриці S і матриці V, причому визначення відповідної направленості одного або більше аудіооб'єктів базується, щонайменше частково, на матриці V, і при цьому визначення відповідної енергії

5 одного або більше аудіооб'єктів базується, щонайменше частково, матриці S.

[896] В одному прикладі, спосіб додатково включає в себе множення матриці V на матрицю S для генерації матриці VS, причому матриця VS включає в себе один або більше векторів. У деяких прикладах, спосіб додатково включає в себе вибір елементів кожного рядка матриці VS, які пов'язані з порядком, що більше 1, піднесення до квадрата кожного з вибраних елементів

10 для формування відповідних квадратів елементів i, для кожного рядка матриці VS, підсумовування квадратів всіх елементів для генерації показника направленості для відповідного вектора матриці VS. У деяких прикладах, кожен рядок матриці VS включає в себе 25 елементів. В одному прикладі, вибір елементів кожного рядка матриці VS, пов'язаної з порядком, що більше 1, включає вибір всіх елементів, починаючи з 5-го елемента кожного рядка

15 матриці VS і закінчуючи 25-им елементом кожного рядка матриці VS. У деяких прикладах, спосіб додатково включає в себе вибір піднабору векторів для представлення характерних аудіооб'єктів. В одному прикладі, вибір піднабору включає вибір чотирьох векторів матриці VS, і вибрані чотири вектори мають чотири найбільших показники направленості з усіх векторів матриці VS. У деяких прикладах, визначення, що вибраний піднабір векторів представляє

20 характерні аудіооб'єкти, здійснюється на основі направленості й енергії кожного вектора.

[897] У деяких прикладах, спосіб включає в себе визначення, з використанням інформації на основі направленості, одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються

25 щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів сферичних гармонік. В одному прикладі, перетворення включає розкладання на сингулярні значення, що генерує матрицю U, яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матрицю S, яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матрицю V, яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік. В одному прикладі, перетворення включає аналіз головних компонентів для ідентифікації характерних компонентів звукового поля і фонових компонентів звукового поля.

[898] У деяких прикладах, пристрій сконфігурований або інакше виконаний з можливістю здійснення будь-яких описаних тут методів або будь-якої комбінації методів. У деяких

35 прикладах, на зчитуваному комп'ютером носії даних зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати будь-які описані тут методи або будь-яку комбінацію методів. У деяких прикладах, пристрій включає в себе засіб для здійснення будь-яких описаних тут методів або будь-якої комбінації методів.

[899] Таким чином, вищенаведені аспекти методів можуть передбачати, що пристрій 510G аудіокодування виконаний з можливістю діяти згідно з наступними пунктами.

[900] Пункт 134954-1B. Пристрій, наприклад пристрій 510G аудіокодування, який містить: один або більше процесорів, виконаних з можливістю ідентифікації одного або більше характерних аудіооб'єктів з одного або більше коефіцієнтів сферичних гармонік, пов'язаних з аудіооб'єктами, на основі направленості й енергії, визначеної для одного або більше з

45 аудіооб'єктів.

[901] Пункт 134954-2B. Пристрій за пунктом 134954-1B, у якому один або більше процесорів додатково виконані з можливістю визначення одного або обох з направленості й енергії одного або більше аудіооб'єктів на основі коефіцієнтів сферичних гармонік, пов'язаних з аудіооб'єктами.

[902] Пункт 134954-3B. Пристрій за будь-яким з пунктів 134954-1B або 134954-2B або їх комбінації, у якому один або більше процесорів додатково виконані з можливістю здійснювати розкладання на сингулярні значення відносно коефіцієнтів сферичних гармонік, що представляють звукове поле, для генерації матриці U, яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S, яка представляє сингулярні значення з

50 множини коефіцієнтів сферичних гармонік, і матриці V, яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік, і представляти множину коефіцієнтів сферичних гармонік як функцію щонайменше частини однієї або більше з матриці U, матриці S і матриці V, у якому один або більше процесорів виконані з можливістю визначення відповідної направленості одного або більше аудіооб'єктів на основі, щонайменше частково, матриці V, і

55

при цьому один або більше процесорів виконані з можливістю визначення відповідної енергії одного або більше аудіооб'єктів на основі, щонайменше частково, матриці S .

[903] Пункт 134954-4B. Пристрій за пунктом 134954-3B, у якому один або більше процесорів додатково виконані з можливістю множення матриці V на матрицю S для генерації матриці VS , причому матриця VS включає в себе один або більше векторів.

[904] Пункт 134954-5B. Пристрій за пунктом 134954-4B, у якому один або більше процесорів додатково виконані з можливістю вибору елементів кожного рядка матриці VS , які пов'язані з порядком, що більше 1, піднесення до квадрата кожного з вибраних елементів для формування відповідних квадратів елементів i , для кожного рядка матриці VS , підсумовування квадратів всіх елементів для генерації показника направленості для відповідного вектора матриці VS .

[905] Пункт 134954-6B. Пристрій за будь-яким з пунктів 134954-4B і 134954-5B або їх комбінації, у якому кожен рядок матриці VS включає в себе 25 елементів.

[906] Пункт 134954-7B. Пристрій за пунктом 134954-6B, у якому один або більше процесорів виконані з можливістю вибору всіх елементів, починаючи з 5-го елемента кожного рядка матриці VS і закінчуючи 25-им елементом кожного рядка матриці VS .

[907] Пункт 134954-8B. Пристрій за будь-яким з пункту 134954-6B і пункту 134954-7B або їх комбінації, у якому один або більше процесорів додатково виконані з можливістю вибору піднабору векторів для представлення характерних аудіооб'єктів.

[908] Пункт 134954-9B. Пристрій за пунктом 134954-8B, у якому один або більше процесорів виконані з можливістю вибору чотирьох векторів матриці VS , і при цьому вибрані чотири вектори мають чотири найбільших показники направленості з усіх векторів матриці VS .

[909] Пункт 134954-10B. Пристрій за будь-яким з пункту 134954-8B і пункту 134954-9B або їх комбінації, у якому один або більше процесорів додатково виконані з можливістю визначення, що вибраний піднабір векторів представляє характерні аудіооб'єкти, на основі направленості й енергії кожного вектора.

[910] Пункт 134954-1C. Пристрій, наприклад пристрій 510G аудіокодування, який містить: один або більше процесорів, виконаних з можливістю визначення, з використанням інформації на основі направленості, одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення перетворення відносно множини коефіцієнтів сферичних гармонік.

[911] Пункт 134954-2C. Спосіб за пунктом 134954-1C, у якому перетворення включає розкладання на сингулярні значення, що генерує матрицю U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матрицю S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матрицю V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік.

[912] Пункт 134954-3C. Спосіб за пунктом 134954-2C, який додатково включає операції відповідно до будь-якої комбінації пунктів 134954-1A-134954-12A і пунктів 134954-1B-134954-9B.

[913] Пункт 134954-4C. Спосіб за пунктом 134954-1C, у якому перетворення включає аналіз головних компонентів для ідентифікації характерних компонентів звукового поля і фонових компонентів звукового поля.

[914] На фіг. 40H показана блок-схема, що демонструє ілюстративний пристрій 510H аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510H аудіокодування може бути аналогічним пристрою 510G аудіокодування тим, що пристрій 510H аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510H аудіокодування може бути аналогічним блоку пристрою 510G аудіокодування тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання і блок 520G витягання компонентів звукового поля, що можуть діяти аналогічно відповідним блокам пристрою 510G аудіокодування. У деяких прикладах, пристрій 510H аудіокодування може включати в себе блок 534 квантування, як описано з посиланням на фіг. 40D-40E, для квантування одного або більше векторів з будь-яких векторів 525C U_{DIST} , векторів 525D U_{BG} , векторів 525E V_{DIST}^T і векторів 525J V_{BG}^T .

[915] Блок 512 стиснення аудіосигналу пристрою 510H аудіокодування, однак, може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510G аудіокодування тим, що блок 512 стиснення аудіосигналу пристрою 510H аудіокодування включає в себе додатковий блок, позначений як блок 550 інтерполяції. Блок 550 інтерполяції може представляти блок, що інтерполює підкадри першого кадру аудіо з підкадрів першого кадру аудіо і другого у часі

наступного або попереднього кадру аудіо, як більш докладно описано нижче з посиланням на фіг. 45 і 45В. Блок 550 інтерполяції може, при здійсненні цієї інтерполяції, знижувати обчислювальну складність (відносно циклів обробки і/або споживання пам'яті) шляхом потенційного зниження міри, у якій блок 518 розкладання потрібний для розкладання SHC 511.

5 У цьому відношенні, блок 550 інтерполяції може здійснювати операції, аналогічні описаним вище відносно блока 50 просторово-часової інтерполяції пристрою 24 аудіокодування, показаного в прикладі, наведеному на фіг. 4.

[916] Таким чином, розкладання на сингулярні значення, здійснюване блоком 518 розкладання, потенційно вимагає великого обсягу ресурсів процесора і/або пам'яті і, також, у деяких прикладах, займає багато часу для розкладання SHC 511, особливо при збільшенні порядку SHC 511. Для скорочення часу і підвищення ефективності стиснення SHC 511 (відносно циклів обробки і/або споживання пам'яті), методи, описані в цьому винаході, можуть передбачати інтерполяцію одного або більше підкадрів першого кадру аудіо, де кожний з підкадрів може представляти розкладені версії SHC 511. Замість здійснення SVD відносно

10 всього кадру, методи можуть передбачати, що блок 518 розкладання розкладає перший підкадр першого кадру аудіо, генеруючи матрицю 519A' V.

[917] Блок 518 розкладання також може розкласти другий підкадр другого кадру аудіо, де цей другий кадр аудіо може бути наступним у часі або попереднім у часі першому кадру аудіо. Блок 518 розкладання може виводити матрицю 519A' V для цього підкадру другого кадру аудіо.

20 Потім блок 550 інтерполяції може інтерполювати підкадри, що залишилися, першого кадру аудіо на основі матриць 519AA' V, розкладених з першого і другого підкадрів, виводити матрицю 519A V, матрицю 519B S і матрицю 519C U, де розкладання для підкадрів, що залишилися, можна обчислювати на основі SHC 511, матриці 519A V для першого кадру аудіо й інтерпольованих матриць 519A V для підкадрів, що залишилися, першого кадру аудіо. Таким чином, інтерполяція дозволяє уникнути обчислення розкладань для підкадрів, що залишилися, першого кадру аудіо.

25 [918] Крім того, як згадано вище, матриця 519C U може не бути неперервною від кадру до кадру, де характерні компоненти матриці 519C U, розкладеної з першого кадру аудіо SHC 511, можуть вказуватися в інших рядках і/або стовпцях, ніж у матриці 519C U, розкладеній з другого кадру аудіо SHC 511. Завдяки здійсненню цієї інтерполяції, порушення неперервності може

знижуватися, за умови, що лінійна інтерполяція може мати згладжуючу дію, що може знижувати будь-які артефакти, які вносяться внаслідок границь кадру (або, іншими словами, сегментування SHC 511 на кадри). Використання матриці 519A' V для здійснення цієї інтерполяції і потім відновлення матриці 519C U на основі інтерпольованої матриці 519A' V з SHC 511 дозволяє згладжувати будь-які ефекти від переупорядковування матриці 519C U.

30 [919] У ході експлуатації, блок 550 інтерполяції може інтерполювати один або більше підкадрів першого кадру аудіо з першого розкладання, наприклад матриці 519A' V, частини першої множини коефіцієнтів 511 сферичних гармонік, включених у перший кадр, і другого розкладання, наприклад матриці 519A' V, частини другої множини коефіцієнтів 511 сферичних гармонік, включених у другий кадр, для генерації розкладених інтерпольованих коефіцієнтів сферичних гармонік для одного або більше підкадрів.

35 [920] У деяких прикладах, перше розкладання містить першу матрицю 519A' V, яка представляє праві сингулярні вектори частини першої множини коефіцієнтів 511 сферичних гармонік. Аналогічно, у деяких прикладах, друге розкладання містить другу матрицю 519A' V, яка представляє праві сингулярні вектори частини другої множини коефіцієнтів сферичних гармонік.

40 [921] Блок 550 інтерполяції може здійснювати часову інтерполяцію відносно одного або більше підкадрів на основі першої матриці 519A' V і другої матриці 519A' V. Таким чином, блок 550 інтерполяції може інтерполювати в часі, наприклад, другий, третій і четвертий підкадри з чотирьох повних підкадрів для першого кадру аудіо на основі матриці 519A' V, розкладеної з

першого підкадру першого кадру аудіо, і матриці 519A' V, розкладеної з першого підкадру другого кадру аудіо. У деяких прикладах, ця часова інтерполяція є лінійною часовою інтерполяцією, де матриця 519A' V, розкладена з першого підкадру першого кадру аудіо, одержує більшу вагу при інтерполяції другого підкадру першого кадру аудіо, ніж при інтерполяції четвертого підкадру першого кадру аудіо. При інтерполяції третього підкадру, матриці 519A' V можуть одержувати однакову вагу. При інтерполяції четвертого підкадру, матриця 519A' V, розкладена з першого підкадру другого кадру аудіо, може одержувати більшу вагу, ніж матриця 519A' V, розкладена з першого підкадру першого кадру аудіо.

45 [922] Іншими словами, лінійна часова інтерполяція може зважувати матриці 519A' V з урахуванням близькості одного з підкадрів першого кадру аудіо, що підлягає інтерполяції. Для

другого підкадру, що підлягає інтерполяції, матриця 519A' V, розкладена з першого підкадру

першого кадру аудіо, одержує більшу вагу з урахуванням його близькості до другого підкадру, що підлягає інтерполяції, ніж матриця $519A' V$, розкладена з першого підкадру другого кадру аудіо. Вагові коефіцієнти, з цієї причини, можуть бути еквівалентні при інтерполяції третього підкадру на основі матриць $519A' V$. Ваговий коефіцієнт, застосовуваний до матриці $519A' V$, розкладеної з першого підкадру другого кадру аудіо, може бути більше, ніж застосовуваний до матриці $519A' V$, розкладеної з першого підкадру першого кадру аудіо, за умови, що четвертий підкадр, що підлягає інтерполяції, ближче до першого підкадру другого кадру аудіо, ніж перший підкадр першого кадру аудіо.

[923] Хоча, у деяких прикладах, тільки перший підкадр кожного кадру аудіо використовується для здійснення інтерполяції, частина першої множини коефіцієнтів сферичних гармонік може містити два з чотирьох підкадрів першої множини коефіцієнтів 511 сферичних гармонік. У цих і інших прикладах, частина другої множини коефіцієнтів 511 сферичних гармонік містить два з чотирьох підкадрів другої множини коефіцієнтів 511 сферичних гармонік.

[924] Як згадано вище, єдиний пристрій, наприклад пристрій 510H аудіокодування, може здійснювати інтерполяцію, при цьому також розкладаючи частину першої множини коефіцієнтів сферичних гармонік для генерації перших розкладань частини першої множини коефіцієнтів сферичних гармонік. У цих і інших прикладах, блок 518 розкладання може розкладати частину другої множини коефіцієнтів сферичних гармонік для генерації других розкладань частини другої множини коефіцієнтів сферичних гармонік. Хоча вони описані відносно єдиного пристрою, два або більше пристроїв можуть здійснювати методи, описані в цьому винаході, де один із двох пристроїв здійснює розкладання й інший з пристроїв здійснює інтерполяцію відповідно до методів, описаних в цьому винаході.

[925] Іншими словами, 3D аудіо на основі сферичних гармонік може бути параметричним представленням 3D поля тиску відносно ортогональних базисних функцій на сфері. Чим вище порядок N представлення, тим потенційно вище просторове розрізнення, і, часто, тим більше кількість коефіцієнтів сферичних гармонік (SH) (для усього $(N+1)^2$ коефіцієнтів). Для багатьох застосувань, може знадобитися, щоб стиснення смуги коефіцієнтів дозволяло ефективно передавати і зберігати коефіцієнти. Ці методи, передбачені в цьому винаході, можуть забезпечувати процес зниження розмірності на кадровій основі з використанням розкладання на сингулярні значення (SVD). Аналіз SVD може розкладати кожен кадр коефіцієнтів на три матриці U , S і V . У деяких прикладах, методи можуть обробляти деякі з векторів у U як направлені компоненти звукового поля, що лежить в основі. Однак, при такій обробці, ці вектори (у U) є розривними від кадру до кадру, хоча вони представляють один і той же характерний аудіокомпонент. Ці порушення неперервності можуть приводити до значних артефактів, коли компоненти надходять через перетворюючі аудіокодері.

[926] Методи, описані в цьому винаході, дозволяють подолати це порушення неперервності. Таким чином, методи можуть ґрунтуватися на тому спостереженні, що матрицю V можна інтерпретувати як ортогональні просторові осі в області сферичних гармонік. Матриця U може представляти проекцію даних сферичних гармонік (HOA) відносно цих базисних функцій, де порушення неперервності може приписуватися базисним функціям (V), що змінюють кожен кадр і, таким чином, самі є розривними. У цьому полягає відмінність від аналогічного розкладання, наприклад перетворення Фур'є, де базисні функції, у деяких прикладах, постійні від кадру до кадру. У цьому відношенні, SVD можна розглядати як алгоритм переслідування збігу. Методи, описані в цьому винаході, можуть передбачати, що блок 550 інтерполяції підтримує неперервність між базисними функціями (V) від кадру до кадру шляхом інтерполяції між ними.

[927] У деяких прикладах, методи передбачають, що блок 550 інтерполяції розділяє кадр даних SH на чотири підкадри, як описано вище і додатково описано нижче з посиланням на фіг. 45 і 45B. Потім блок 550 інтерполяції може обчислювати SVD для першого підкадру. Аналогічно обчислюється SVD для першого підкадру другого кадру. Для кожного з першого кадру і другого кадру, блок 550 інтерполяції може перетворювати вектори в V у просторову карту шляхом проектування векторів на сферу (з використанням проекційної матриці, наприклад Т-подібної матриці). Потім блок 550 інтерполяції може інтерпретувати вектори в V як форми на сфері. Для інтерполяції матриць V для трьох підкадрів між першим підкадром першого кадру і першим підкадром наступного кадру, блок 550 інтерполяції потім може інтерполювати ці просторові форми і потім перетворювати їх зворотно у SH вектори за допомогою оберненої проекційної матриці. Методи цього винаходу, таким чином, можуть забезпечувати плавний перехід між матрицями V .

[928] Таким чином, пристрій 510H аудіокодування може бути виконаний з можливістю здійснення різних аспектів методів, викладених нижче відносно наступних пунктів.

[929] Пункт 135054-1А. Пристрій, наприклад пристрій 510Н аудіокодування, який містить: один або більше процесорів, виконаних з можливістю інтерполяції одного або більше підкадрів першого кадру з першого розкладання частини першої множини коефіцієнтів сферичних гармонік, включених у перший кадр, і другого розкладання частини другої множини коефіцієнтів сферичних гармонік, включених у другий кадр, для генерації розкладених інтерпольованих коефіцієнтів сферичних гармонік для одного або більше підкадрів.

[930] Пункт 135054-2А. Пристрій за пунктом 135054-1А, у якому перше розкладання містить першу матрицю V , яка представляє праві сингулярні вектори частини першої множини коефіцієнтів сферичних гармонік.

[931] Пункт 135054-3А. Пристрій за пунктом 135054-1А, у якому друге розкладання містить другу матрицю V , яка представляє праві сингулярні вектори частини другої множини коефіцієнтів сферичних гармонік.

[932] Пункт 135054-4А. Пристрій за пунктом 135054-1А, у якому перше розкладання містить першу матрицю V , яка представляє праві сингулярні вектори частини першої множини коефіцієнтів сферичних гармонік, і при цьому друге розкладання містить другу матрицю V , яка представляє праві сингулярні вектори частини другої множини коефіцієнтів сферичних гармонік.

[933] Пункт 135054-5А. Пристрій за пунктом 135054-1А, у якому один або більше процесорів додатково виконані з можливістю, при інтерполяції одного або більше підкадрів, інтерпольовати в часі один або більше підкадрів на основі першого розкладання і другого розкладання.

[934] Пункт 135054-6А. Пристрій за пунктом 135054-1А, у якому один або більше процесорів додатково виконані з можливістю, при інтерполяції одного або більше підкадрів, проектувати перше розкладання в просторову область для генерації перших проекційних розкладань, проектувати друге розкладання в просторову область для генерації других проекційних розкладань, просторово інтерпольовати перші проекційні розкладання і другі проекційні розкладання для генерації першого просторово інтерпольованого проекційного розкладання і другого просторово інтерпольованого проекційного розкладання і інтерпольовати в часі один або більше підкадрів на основі першого просторово інтерпольованого проекційного розкладання і другого просторово інтерпольованого проекційного розкладання.

[935] Пункт 135054-7А. Пристрій за пунктом 135054-6А, у якому один або більше процесорів додатково виконані з можливістю проектувати інтерпольовані в часі коефіцієнти сферичних гармонік, обумовлені інтерполяцією одного або більше підкадрів, зворотно в область сферичних гармонік.

[936] Пункт 135054-8А. Пристрій за пунктом 135054-1А, у якому частина першої множини коефіцієнтів сферичних гармонік містить єдиний підкадр першої множини коефіцієнтів сферичних гармонік.

[937] Пункт 135054-9А. Пристрій за пунктом 135054-1А, у якому частина другої множини коефіцієнтів сферичних гармонік містить єдиний підкадр другої множини коефіцієнтів сферичних гармонік.

[938] Пункт 135054-10А. Пристрій за пунктом 135054-1А,

[939] у якому перший кадр поділяється на чотири підкадри, і

[940] у якому частина першої множини коефіцієнтів сферичних гармонік містить тільки перший підкадр першої множини коефіцієнтів сферичних гармонік.

[941] Пункт 135054-11А. Пристрій за пунктом 135054-1А,

[942] у якому другий кадр поділяється на чотири підкадри, і

[943] у якому частина другої множини коефіцієнтів сферичних гармонік містить тільки перший підкадр другої множини коефіцієнтів сферичних гармонік.

[944] Пункт 135054-12А. Пристрій за пунктом 135054-1А, у якому частина першої множини коефіцієнтів сферичних гармонік містить два з чотирьох підкадрів першої множини коефіцієнтів сферичних гармонік.

[945] Пункт 135054-13А. Пристрій за пунктом 135054-1А, у якому частина другої множини коефіцієнтів сферичних гармонік містить два з чотирьох підкадрів другої множини коефіцієнтів сферичних гармонік.

[946] Пункт 135054-14А. Пристрій за пунктом 135054-1А, у якому один або більше процесорів додатково виконані з можливістю розкладання частини першої множини коефіцієнтів сферичних гармонік для генерації перших розкладань частини першої множини коефіцієнтів сферичних гармонік.

[947] Пункт 135054-15А. Пристрій за пунктом 135054-1А, у якому один або більше процесорів додатково виконані з можливістю розкладання частини другої множини коефіцієнтів сферичних гармонік для генерації других розкладань частини другої множини коефіцієнтів сферичних гармонік.

[948] Пункт 135054-16A. Пристрій за пунктом 135054-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно частини першої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори першої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення першої множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори першої множини коефіцієнтів сферичних гармонік.

[949] Пункт 135054-17A. Пристрій за пунктом 135054-1A, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно частини другої множини коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори другої множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення другої множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[950] Пункт 135054-18A. Пристрій за пунктом 135054-1A, у якому перша і друга множини коефіцієнтів сферичних гармонік представляють плоскохвильове представлення звукового поля.

[951] Пункт 135054-19A. Пристрій за пунктом 135054-1A, у якому перша і друга множини коефіцієнтів сферичних гармонік представляють один або більше змішаних один з одним монофонічних аудіооб'єктів.

[952] Пункт 135054-20A. Пристрій за пунктом 135054-1A, у якому перша і друга множини коефіцієнтів сферичних гармонік містять відповідні перший і другий коефіцієнти сферичних гармонік, що представляють тривимірне звукове поле.

[953] Пункт 135054-21A. Пристрій за пунктом 135054-1A, у якому перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, яка має порядок, що більше одиниці.

[954] Пункт 135054-22A. Пристрій за пунктом 135054-1A, у якому перша і друга множини коефіцієнтів сферичних гармонік пов'язані з щонайменше однією сферичною базисною функцією, яка має порядок, що дорівнює чотирьом.

[955] Хоча вони описані вище як здійснювані пристроєм 510H аудіокодування, різні пристрої аудіодекодування 24 і 540 також можуть здійснювати будь-який з різних аспектів методів, викладених вище відносно пунктів 135054-1A-135054-22A.

[956] На фіг. 40I показана блок-схема, що демонструє ілюстративний пристрій 510I аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510I аудіокодування може бути аналогічним пристрою 510H аудіокодування тим, що пристрій 510I аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510I аудіокодування може бути аналогічним блоку пристрою 510H аудіокодування тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання і блок 520 витягання компонентів звукового поля, що можуть діяти аналогічно відповідним блокам пристрою 510H аудіокодування. У деяких прикладах, пристрій аудіокодування 10I може включати в себе блок квантування 34, як описано з посиланням на фіг. 3D-3E, для квантування одного або більше векторів будь-якої з $25C U_{DIST}$, $25D U_{BG}$, $25E V_{DIST}^T$ і $25J V_{BG}^T$.

[957] Однак, хоча і блок 512 стиснення аудіосигналу пристрою 510I аудіокодування, і блок 512 стиснення аудіосигналу пристрою аудіокодування 510H включають в себе блок витягання компонентів звукового поля, блок 520I витягання компонентів звукового поля пристрою 510I аудіокодування може включати в себе додатковий модуль, іменований блоком 552 стиснення V . Блок 552 стиснення V може представляти блок, виконаний з можливістю стиснення просторового компонента звукового поля, тобто одного або більше з векторів $539 V_{DIST}^T$ у цьому прикладі. Таким чином, розкладання на сингулярні значення, здійснюване відносно SHC може розкласти SHC (що представляє звукове поле) на енергетичні компоненти, представлені векторами матриці S , часові компоненти, представлені матрицею U , і просторові компоненти, представлені матрицею V . Блок 552 стиснення V може здійснювати операції, аналогічні описаним вище відносно блока 52 квантування.

[958] Для прикладу, передбачається, що вектори $539 V_{DIST}^T$ містять два вектори-рядки, що мають 25 елементів кожний (що передбачає представлення НОА звукового поля четвертого порядку). Хоч опис наведений відносно двох векторів-рядків, будь-яка кількість векторів може бути включена у вектори $539 V_{DIST}^T$ аж до $(n+1)^2$, де n позначає порядок представлення НОА звукового поля.

[959] Блок 552 стиснення V може приймати вектори $539 V_{\text{DIST}}^T$ і здійснювати схему стиснення для генерації стисненого вектора V_{DIST}^T представлення 539'. Ця схема стиснення може передбачати, у загальному випадку, будь-яку припустиму схему стиснення для стиснення елементів вектора або даних, і не підлягає обмеженню прикладом, більш детально описаним

5

[960] Блок 552 стиснення V може здійснювати, у порядку прикладу, схему стиснення, яка включає в себе одне або більше з перетворення представлень з плаваючою комою кожного елемента векторів $539 V_{\text{DIST}}^T$ у цілочислові представлення кожного елемента векторів $539 V_{\text{DIST}}^T$, однорідного квантування цілочислових представлень векторів $539 V_{\text{DIST}}^T$ і категоризації і кодування квантованих цілочислових представлень векторів $539 V_{\text{DIST}}^T$. Різні один або більше процесів цієї схеми стиснення можуть динамічно керуватися параметрами для досягнення або приблизного досягнення, у порядку одного прикладу, цільової бітової швидкості для результуючого бітового потоку 517.

10

[961] За умови, що кожний з векторів $539 V_{\text{DIST}}^T$ ортонормальні один одному, кожний з векторів $539 V_{\text{DIST}}^T$ можна кодувати незалежно. У деяких прикладах, як більш докладно описано нижче, кожен елемент кожного вектора $539 V_{\text{DIST}}^T$ можна кодувати з використанням одного і того ж режиму кодування (заданого різними підрежимами).

15

[962] У будь-якому випадку, як згадано вище, ця схема кодування може спочатку використовувати перетворення представлень з плаваючою комою кожного елемента (яке, у деяких прикладах, є 32-бітовим числом з плаваючою комою) кожного з векторів $539 V_{\text{DIST}}^T$ у 16-бітове цілочислове представлення. Блок 552 стиснення V може здійснювати це перетворення плаваючої коми у ціле число шляхом множення кожного елемента даного одного з векторів $539 V_{\text{DIST}}^T$ на 2^{15} , що, у деяких прикладах, здійснюється за допомогою зсуву вправо на 15.

20

[963] Блок 552 стиснення V потім може здійснювати однорідне квантування відносно всіх елементів даного одного з векторів $539 V_{\text{DIST}}^T$. Блок 552 стиснення V може ідентифікувати розмір кроку квантування на основі значення, яке може бути позначене як параметр nbits . Блок 552 стиснення V може динамічно визначати цей параметр nbits на основі цільової бітової швидкості. Блок 552 стиснення V може визначати розмір кроку квантування як функцію цього параметра nbits . У порядку одного прикладу, блок 552 стиснення V може визначати розмір кроку квантування (позначений як "дельта" або " Δ " у цьому винаході) рівним $2^{16-\text{nbits}}$. У цьому прикладі, якщо nbits дорівнює шести, дельта дорівнює 2^{10} і існує 2^6 рівнів квантування. У цьому відношенні, для елемента v вектора, квантований елемент v_q вектора дорівнює $\lceil v/\Delta \rceil$ і $-2^{\text{nbits}-1} < v_q \leq 2^{\text{nbits}-1}$.

25

30

[964] Блок 552 стиснення V потім може здійснювати категоризацію і залишкове кодування квантованих елементів вектора. У порядку одного прикладу, блок 552 стиснення V може, для даного квантованого елемента v_q вектора, ідентифікувати категорію (визначаючи ідентифікатор категорії cid), якій відповідає цей елемент, з використанням наступного рівняння:

35

$$\text{cid} = \begin{cases} 0, & \text{if } v_q = 0 \\ \lfloor \log_2 |v_q| \rfloor + 1, & \text{if } v_q \neq 0 \end{cases}.$$

Потім блок 552 стиснення V може кодувати по Хаффману цей індекс категорії cid , ідентифікуючи також знаковий біт, який вказує, чи має v_q позитивне значення або негативне значення. Потім блок 552 стиснення V може ідентифікувати залишок у цій категорії. У порядку одного прикладу, блок 552 стиснення V може визначати цей залишок згідно з наступним рівнянням:

40

$$\text{residual} = |v_q| - 2^{\text{cid}-1}.$$

45

Потім блок 552 стиснення V може блоково кодувати цей залишок $\text{cid}-1$ бітами.

[965] Наступний приклад демонструє спрощений приклад цього процесу категоризації і залишкового кодування. Спочатку припустимо, що nbits дорівнює шести, так що $v_q \in [-31, 31]$.

Потім припустимо наступне:

cid	v_q	Код Хаффмана для cid
0	0	'1'
1	-1, 1	'01'
2	-3,-2, 2,3	'000'
3	-7,-6,-5,-4, 4,5,6,7	'0010'
4	-15,-14,...,-8, 8,..., 14,15	'00110'
5	-31,-30,..., -16, 16,..., 30,31	'00111'

Також припустимо наступне:

cid	Блоковий код для залишку
0	N/A
1	0, 1
2	01,00, 10,11
3	011,010,001,000, 100,101,110,111
4	0111, 0110,..., 0000, 1000,..., 1110,1111
5	01111,..., 00000, 10000,..., 11111

Таким чином, для $v_q=[6, -17, 0, 0, 3]$, можна визначити наступне:

» cid=3,5,0,0,2

» знак=1,0,x, x,1

» залишок=2,1,x, x,1

» біти для 6='0010'+1'+10'

» біти для -17='00111'+0'+0001'

» біти для 0='0'

» біти для 0='0'

» біти для 3='000'+1'+1'

» Усього бітів=7+10+1+1+5=24

» Середня кількість бітів=24/5=4.8

[966] Хоча це не показано у вищенаведеному спрощеному прикладі, блок 552 стиснення V може вибирати різні кодові книги Хаффмана для різних значень nbits при кодуванні cid. У деяких прикладах, блок 552 стиснення V може забезпечувати різні таблиці хаффманівського кодування для значень nbits 6,..., 15... Крім того, блок 552 стиснення V може включати в себе п'ять різних кодових книг Хаффмана для кожного з різних значень nbits у межах 6,..., 15 для всього 50 кодових книг Хаффмана. У цьому відношенні, блок 552 стиснення V може включати в себе множину різних кодових книг Хаффмана для розміщення кодування cid у декількох різних статистичних контекстах.

[967] Для ілюстрації, блок 552 стиснення V може, для кожного зі значень nbits, включати в себе першу кодову книгу Хаффмана для кодування елементів вектора з першого по четвертий, другу кодову книгу Хаффмана для кодування елементів вектора з п'ятого по дев'ятий, третю кодову книгу Хаффмана для кодування елементів вектора з дев'ятого і вище. Ці перші три кодові книги Хаффмана можна використовувати, коли один з векторів $539 V_{DIST}^T$, підлягаючих стисненню, не прогнозується з наступного в часі відповідного одного з векторів $539 V_{DIST}^T$ і не представляє просторову інформацію синтетичного аудіооб'єкта (заданого, наприклад, спочатку аудіооб'єктом, підданим імпульсно-кодовій модуляції (IKM)). Блок 552 стиснення V може додатково включати в себе, для кожного зі значень nbits, четверту кодову книгу Хаффмана для кодування одного з векторів $539 V_{DIST}^T$, коли цей один з векторів $539 V_{DIST}^T$ прогнозується з наступного в часі відповідного одного з векторів $539 V_{DIST}^T$. Блок 552 стиснення V також може включати в себе, для кожного зі значень nbits, п'яту кодову книгу Хаффмана для кодування одного з векторів $539 V_{DIST}^T$, коли цей один з векторів $539 V_{DIST}^T$ представляє синтетичний аудіооб'єкт. Різні кодові книги Хаффмана можна розробляти для кожного з цих різних статистичних контекстів, тобто непрогнозованого і несинтетичного контексту, прогнозованого контексту і синтетичного контексту в цьому прикладі.

[968] Нижченаведена таблиця демонструє вибір таблиці Хаффмана і біти, що підлягають указанню в бітовому потоці, щоб блок зняття стиснення міг вибирати належну таблицю Хаффмана.

Режим прогнозування	Інформація НТ	Таблиця НТ
0	0	НТ5
0	1	НТ{1,2,3}
1	0	НТ4
1	1	НТ5

У вищенаведеній таблиці, режим прогнозування ("режим прогнозування") указує, чи здійснювалося прогнозування для поточного вектора, тоді як таблиця Хаффмана ("інформація НТ") указує додаткову інформацію кодової книги (або таблиці) Хаффмана, використовувану для вибору однієї з таблиць Хаффмана з першої по п'яту.

[969] Нижченаведена таблиця додатково демонструє цей процес вибору таблиці Хаффмана в різних статистичних контекстах або сценаріях.

	Запис	Синтетичний
Без прогнозування	HT{1,2,3}	HT5
З прогнозуванням	HT4	HT5

У вищенаведеній таблиці, стовпець "запис" указує контекст кодування, коли вектор представляє аудіооб'єкт, що був записаний, тоді як стовпець "синтетичний" указує контекст кодування, коли вектор представляє синтетичний аудіооб'єкт. Рядок "без прогнозування" указує контекст кодування, коли прогнозування не здійснюється відносно елементів вектора, тоді як рядок "з прогнозуванням" указує контекст кодування, коли прогнозування здійснюється відносно елементів вектора. Як показано в цій таблиці, блок 552 стиснення V вибирає HT{1,2,3}, коли вектор представляє записаний аудіооб'єкт, і прогнозування не здійснюється відносно елементів вектора. Блок 552 стиснення V вибирає HT5, коли аудіооб'єкт представляє синтетичний аудіооб'єкт, і прогнозування не здійснюється відносно елементів вектора. Блок 552 стиснення V вибирає HT4, коли вектор представляє записаний аудіооб'єкт, і прогнозування здійснюється відносно елементів вектора. Блок 552 стиснення V вибирає HT5, коли аудіооб'єкт представляє синтетичний аудіооб'єкт, і прогнозування здійснюється відносно елементів вектора.

[970] Таким чином, методи можуть передбачати, що пристрій стиснення аудіосигналу стискає просторовий компонент звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[971] На фіг. 43 показана схема, що більш детально демонструє блок 552 стиснення V, показаний на фіг. 40I. У прикладі, наведеному на фіг. 43, блок 552 стиснення V включає в себе блок 600 однорідного квантування, блок 602 nbits, блок 604 прогнозування, блок 606 режиму прогнозування ("блок 606 режиму прогнозування"), блок 608 кодування категорії і залишку і блок 610 вибору таблиці Хаффмана. Блок 600 однорідного квантування представляє блок, виконаний з можливістю здійснення однорідного квантування, описаного вище відносно одного з просторових компонентів, позначеного як v у прикладі, наведеному на фіг. 43 (який може представляти будь-який з векторів 539 V_{DIST}^T). Блок 602 nbits представляє блок, виконаний з можливістю визначення параметра або значення nbits.

[972] Блок 604 прогнозування представляє блок, виконаний з можливістю здійснення прогнозування відносно квантованого просторового компонента, позначеного як v_q у прикладі, наведеному на фіг. 43. Блок 604 прогнозування може здійснювати прогнозування шляхом здійснення поелементного віднімання поточного одного з векторів 539 V_{DIST}^T з наступного в часі відповідного одного з векторів 539 V_{DIST}^T . Результат цього прогнозування може іменуватися прогнозованим просторовим компонентом.

[973] Блок 606 режиму прогнозування може представляти блок, виконаний з можливістю вибору режиму прогнозування. Блок 610 вибору таблиці Хаффмана може представляти блок, виконаний з можливістю вибору належної таблиці Хаффмана для кодування cid. Блок 606 режиму прогнозування і блок 610 вибору таблиці Хаффмана можуть діяти, у порядку одного прикладу, згідно з наступним псевдокодом:

```

для даного nbits, витягнути всі таблиці Хаффмана, що мають nbits
B00=0; B01=0; B10=0; B11=0; // ініціалізувати для обчислення очікуваних бітів для кожного
режиму кодування
for m=1: (кількість елементів у векторі)
    // обчислювати очікувану кількість бітів для елемента вектора v(m)
    // без прогнозування і з використанням таблиці Хаффмана 5
    B00=B00+calculate_bits(v(m), HT5);
    // без прогнозування і з використанням таблиці Хаффмана {1,2,3}
    B01=B01+calculate_bits(v(m), HTq); q у {1,2,3}
    // обчислювати очікувану кількість бітів для залишку прогнозування e(m)
    e(m)=v(m)-vp(m); // vp(m): елемент вектора попереднього кадру
    // з прогнозуванням і з використанням таблиці Хаффмана 4
    B10=B10+calculate_bits(e(m), HT4);
    // з прогнозуванням і з використанням таблиці Хаффмана 5
    B11=B11+calculate_bits(e(m), HT5);
end
// знайти найкращий режим прогнозування і таблицю Хаффмана, що дають мінімальну
кількість бітів
// найкращі режим прогнозування і таблиця Хаффмана позначені прапорами pflag і Htflag,
відповідно

```

```

[Be, id]=min([B00 B01 B10 B11]);
Switch id
case 1: pflag=0; HTflag=0;
case 2: pflag=0; HTflag=1;
5 case 3: pflag=1; HTflag=0;
case 4: pflag=1; HTflag=1;
end

```

[974] Блок 608 кодування категорії і залишку може представляти блок, виконаний з можливістю здійснення категоризації і залишкового кодування прогнозованого просторового компонента або квантованого просторового компонента (коли прогнозування заборонене), як більш детально описано вище.

[975] Як показано в прикладі, наведеному на фіг. 43, блок 552 стиснення V може виводити різні параметри або значення для включення або в бітовий потік 517, або в побічну інформацію (яка сама може бути бітовим потоком, окремим від бітового потоку 517). Передбачаючи, що інформація вказується в бітовому потоці 517, блок 552 стиснення V може виводити значення nbits, режим прогнозування й інформацію таблиці Хаффмана на блок 516 генерації бітового потоку спільно зі стисненою версією просторового компонента (показаною як стиснений просторовий компонент 539' у прикладі, наведеному на фіг. 40I), що у цьому прикладі може означати код Хаффмана, вибраний для кодування sid, знакового біта і блоково кодованого залишку. Значення nbits може вказуватися один раз у бітовому потоці 517 для усіх векторів 539 V_{DIST}^T , тоді як режим прогнозування й інформація таблиці Хаффмана можуть вказуватися для кожного з векторів 539 V_{DIST}^T . Частина бітового потоку, що вказує стиснену версію просторового компонента, показана в прикладі, наведеному на фіг. 10B і 10C.

[976] Таким чином, пристрій 510H аудіокодування може здійснювати різні аспекти методів, викладених нижче відносно наступних пунктів.

[977] Пункт 141541-1A. Пристрій, наприклад пристрій 510H аудіокодування, який містить: один або більше процесорів, виконаних з можливістю одержання бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[978] Пункт 141541-2A. Пристрій за пунктом 141541-1A, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента.

[979] Пункт 141541-3A. Пристрій за будь-якою комбінацією пунктів 141541-1A і 141541-2A, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.

[980] Пункт 141541-4A. Пристрій за будь-якою комбінацією пунктів 141541-1A-141541-3A, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.

[981] Пункт 141541-5A. Пристрій за пунктом 141541-4A, у якому значення містить значення nbits.

[982] Пункт 141541-6A. Пристрій за будь-якою комбінацією пунктів 141541-4A і 141541-5A, у якому бітовий потік містить стиснену версію з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і при цьому значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.

[983] Пункт 141541-7A. Пристрій за будь-якою комбінацією пунктів 141541-1A-141541-6A, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.

[984] Пункт 141541-8A. Пристрій за будь-якою комбінацією пунктів 141541-1A-141541-7A, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

[985] Пункт 141541-9A. Пристрій за будь-якою комбінацією пунктів 141541-1A-141541-8A, у якому стиснена версія просторового компонента представлена в бітовому потоці з

використанням, щонайменше частково, коду Хаффмана для представлення залишкового значення просторового компонента.

[986] Пункт 141541-10A. Пристрій за будь-якою комбінацією пунктів 141541-1A-141541-9A, у якому пристрій містить пристрій аудіокодування і пристрій генерації бітового потоку.

5 [987] Пункт 141541-12A. Пристрій за будь-якою комбінацією пунктів 141541-1A-141541-11A, у якому векторний синтез включає розкладання на сингулярні значення.

[988] Хоча вони описані як здійснювані пристроєм 510H аудіокодування, методи також можуть здійснюватися будь-яким із пристроїв 24 і/або 540 аудіодекодування.

10 [989] Таким чином, пристрій 510H аудіокодування може додатково здійснювати різні аспекти методів, викладених нижче відносно наступних пунктів.

[990] Пункт 141541-1D. Пристрій, наприклад пристрій 510H аудіокодування, який містить: один або більше процесорів, виконаних з можливістю генерації бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

15 [991] Пункт 141541-2D. Пристрій за пунктом 141541-1D, у якому один або більше процесорів додатково виконані з можливістю, при генерації бітового потоку, генерувати бітовий потік для включення поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента.

20 [992] Пункт 141541-3D. Пристрій за будь-якою комбінацією пунктів 141541-1D і 141541-2D, у якому один або більше процесорів додатково виконані з можливістю, при генерації бітового потоку, генерувати бітовий потік для включення інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.

25 [993] Пункт 141541-4D. Пристрій за будь-якою комбінацією пунктів 141541-1D-141541-3D, у якому один або більше процесорів додатково виконані з можливістю, при генерації бітового потоку, генерувати бітовий потік для включення поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.

[994] Пункт 141541-5D. Пристрій за пунктом 141541-4D, у якому значення містить значення nbits.

30 [995] Пункт 141541-6D. Пристрій за будь-якою комбінацією пунктів 141541-4D і 141541-5D, у якому один або більше процесорів додатково виконані з можливістю, при генерації бітового потоку, генерувати бітовий потік для включення стисненої версії з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і при цьому значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.

35 [996] Пункт 141541-7D. Пристрій за будь-якою комбінацією пунктів 141541-1D-141541-6D, у якому один або більше процесорів додатково виконані з можливістю, при генерації бітового потоку, генерувати бітовий потік для включення коду Хаффмана для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.

40 [997] Пункт 141541-8D. Пристрій за будь-якою комбінацією пунктів 141541-1D-141541-7D, у якому один або більше процесорів додатково виконані з можливістю, при генерації бітового потоку, генерувати бітовий потік для включення знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

45 [998] Пункт 141541-9D. Пристрій за будь-якою комбінацією пунктів 141541-1D-141541-8D, у якому один або більше процесорів додатково виконані з можливістю, при генерації бітового потоку, генерувати бітовий потік для включення коду Хаффмана для представлення залишкового значення просторового компонента.

50 [999] Пункт 141541-10D. Пристрій за будь-якою комбінацією пунктів 141541-1D-141541-10D, у якому векторний синтез включає розкладання на сингулярні значення.

[1000] Пристрій 510H аудіокодування може бути додатково виконаний з можливістю реалізації різних аспектів методів, викладених у наступних пунктах.

55 [1001] Пункт 141541-1E. Пристрій, наприклад пристрій 510H аудіокодування, який містить: один або більше процесорів, виконаних з можливістю стиснення просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

60 [1002] Пункт 141541-2E. Пристрій за пунктом 141541-1E, у якому один або більше процесорів додатково виконані з можливістю, при стисненні просторового компонента, перетворювати просторовий компонент із представлення з плаваючою комою у цілочислове представлення.

[1003] Пункт 141541-3E. Пристрій за будь-якою комбінацією пунктів 141541-1E і 141541-2E, у якому один або більше процесорів додатково виконані з можливістю, при стисненні просторового компонента, динамічно визначати значення, що вказує розмір кроку квантування, і квантувати просторовий компонент на основі значення для генерації квантованого просторового компонента.

[1004] Пункт 141541-4E. Пристрій за будь-якою комбінацією пунктів 141541-1E-141541-3E, у якому один або більше процесорів додатково виконані з можливістю, при стисненні просторового компонента, ідентифікувати категорію, якій відповідає просторовий компонент.

[1005] Пункт 141541-5E. Пристрій за будь-якою комбінацією пунктів 141541-1E-141541-4E, у якому один або більше процесорів додатково виконані з можливістю, при стисненні просторового компонента, ідентифікувати залишкове значення для просторового компонента.

[1006] Пункт 141541-6E. Пристрій за будь-якою комбінацією пунктів 141541-1E-141541-5E, у якому один або більше процесорів додатково виконані з можливістю, при стисненні просторового компонента, здійснювати прогнозування відносно просторового компонента і наступного просторового компонента для генерації прогнозованого просторового компонента.

[1007] Пункт 141541-7E. Пристрій за будь-якою комбінацією пунктів 141541-1E, у якому один або більше процесорів додатково виконані з можливістю, при стисненні просторового компонента, перетворювати просторовий компонент із представлення з плаваючою комою у цілочислове представлення, динамічно визначати значення, що вказує розмір кроку квантування, квантувати цілочислове представлення просторового компонента на основі значення для генерації квантованого просторового компонента, ідентифікувати категорію, якій відповідає просторовий компонент, на основі квантованого просторового компонента для генерації ідентифікатора категорії, визначати знак просторового компонента, ідентифікувати залишкове значення для просторового компонента на основі квантованого просторового компонента й ідентифікатора категорії і генерувати стиснену версію просторового компонента на основі ідентифікатора категорії, знака і залишкового значення.

[1008] Пункт 141541-8E. Пристрій за будь-якою комбінацією пунктів 141541-1E, у якому один або більше процесорів додатково виконані з можливістю, при стисненні просторового компонента, перетворювати просторовий компонент із представлення з плаваючою комою у цілочислове представлення, динамічно визначати значення, що вказує розмір кроку квантування, квантувати цілочислове представлення просторового компонента на основі значення для генерації квантованого просторового компонента, здійснювати прогнозування відносно просторового компонента і наступного просторового компонента для генерації прогнозованого просторового компонента, ідентифікувати категорію, якій відповідає прогнозований просторовий компонент, на основі квантованого просторового компонента для генерації ідентифікатора категорії, визначати знак просторового компонента, ідентифікувати залишкове значення для просторового компонента на основі квантованого просторового компонента й ідентифікатора категорії і генерувати стиснену версію просторового компонента на основі ідентифікатора категорії, знака і залишкового значення.

[1009] Пункт 141541-9E. Пристрій за будь-якою комбінацією пунктів 141541-1E-141541-8E, у якому векторний синтез включає розкладання на сингулярні значення.

[1010] Різні аспекти методів можуть, крім того, передбачати, що пристрій 510H аудіокодування виконаний з можливістю діяти відповідно до наступних пунктів.

[1011] Пункт 141541-1F. Пристрій, наприклад пристрій 510H аудіокодування, який містить: один або більше процесорів, виконаних з можливістю ідентифікації кодової книги Хаффмана для використання при стисненні поточного просторового компонента з множини просторових компонентів на основі порядку поточного просторового компонента відносно просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[1012] Пункт 141541-2F. Пристрій за пунктом 141541-3F, у якому один або більше процесорів додатково виконані з можливістю здійснення будь-якої комбінації етапів, згаданих у пунктах 141541-1A-141541-12A, пунктах 141541-1B-141541-10B і пунктах 141541-1C-141541-9C.

[1013] Різні аспекти методів можуть, крім того, передбачати, що пристрій 510H аудіокодування виконаний з можливістю діяти відповідно до наступних пунктів.

[1014] Пункт 141541-1H. Пристрій, наприклад пристрій 510H аудіокодування, який містить: один або більше процесорів, виконаних з можливістю визначення розміру кроку квантування, що підлягає використанню при стисненні просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[1015] Пункт 141541-2Н. Пристрій за пунктом 141541-1Н, у якому один або більше процесорів додатково виконані з можливістю, при визначенні розміру кроку квантування, визначати розмір кроку квантування на основі цільової бітової швидкості.

5 [1016] Пункт 141541-3Н. Пристрій за пунктом 141541-1Н, у якому один або більше процесорів додатково виконані з можливістю, при виборі одного з множини розмірів кроку квантування, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, і визначення розміру кроку квантування на основі різниці між оцінкою і цільовою бітовою швидкістю.

10 [1017] Пункт 141541-4Н. Пристрій за пунктом 141541-1Н, у якому один або більше процесорів додатково виконані з можливістю, при виборі одного з множини розмірів кроку квантування, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента, визначення різниці між оцінкою і цільовою бітовою швидкістю і визначення розміру кроку квантування шляхом додавання різниці до цільової бітової швидкості.

15 [1018] Пункт 141541-5Н. Пристрій за пунктом 141541-3Н або пунктом 141541-4Н, у якому один або більше процесорів додатково виконані з можливістю, при визначенні оцінки кількості бітів, обчислення оціненої кількості бітів, що підлягають генерації для просторового компонента, за умови, що кодова книга відповідає цільовій бітій швидкості.

20 [1019] Пункт 141541-6Н. Пристрій за пунктом 141541-3Н або пунктом 141541-4Н, у якому один або більше процесорів додатково виконані з можливістю, при визначенні оцінки кількості бітів, обчислення оціненої кількості бітів, що підлягають генерації для просторового компонента, при даному режимі кодування, використовуваному при стисненні просторового компонента.

25 [1020] Пункт 141541-7Н. Пристрій за пунктом 141541-3Н або пунктом 141541-4Н, у якому один або більше процесорів додатково виконані з можливістю, при визначенні оцінки кількості бітів, обчислення першої оцінки кількості бітів, що підлягають генерації для просторового компонента, за умови, що перший режим кодування підлягає використанню при стисненні просторового компонента, обчислення другої оцінки кількості бітів, що підлягають генерації для просторового компонента, за умови, що другий режим кодування підлягає використанню при стисненні просторового компонента, вибору однієї з першої оцінки і другої оцінки, що має найменшу кількість бітів, які підлягають використанню, як визначеної оцінки кількості бітів.

30 [1021] Пункт 141541-8Н. Пристрій за пунктом 141541-3Н або пунктом 141541-4Н, у якому один або більше процесорів додатково виконані з можливістю, при визначенні оцінки кількості бітів, ідентифікації ідентифікатора категорії, що ідентифікує категорію, якій відповідає просторовий компонент, ідентифікації бітової довжини залишкового значення для просторового компонента, яке було б одержане при стисненні просторового компонента, що відповідає категорії, і визначення оцінки кількості бітів шляхом, щонайменше частково, додавання кількості бітів, використовуваних для представлення ідентифікатора категорії, до бітової довжини залишкового значення.

40 [1022] Пункт 141541-9Н. Пристрій за будь-якою комбінацією пунктів 141541-1Н-141541-8Н, у якому векторний синтез включає розкладання на сингулярні значення.

[1023] Хоча вони описані як здійснювані пристроєм 510Н аудіокодування, методи, викладені вище в пунктах 141541-1Н-141541-9Н, також можуть здійснюватися пристроєм 540D аудіодекодування.

[1024] Додатково, різні аспекти методів можуть передбачати, що пристрій 510Н аудіокодування виконаний з можливістю діяти, як викладено у наступних пунктах.

45 [1025] Пункт 141541-1J. Пристрій, наприклад пристрій 510J аудіокодування, який містить: один або більше процесорів, виконаних з можливістю вибору однієї з множини кодових книг, що підлягають використанню при стисненні просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

50 [1026] Пункт 141541-2J. Пристрій за пунктом 141541-1J, у якому один або більше процесорів додатково виконані з можливістю, при виборі однієї з множини кодових книг, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента з використанням кожної з множини кодових книг, і вибору однієї з множини кодових книг, яка привела до визначеної оцінки, що має найменшу кількість бітів.

55 [1027] Пункт 141541-3J. Пристрій за пунктом 141541-1J, у якому один або більше процесорів додатково виконані з можливістю, при виборі однієї з множини кодових книг, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента з використанням однієї або більше з множини кодових книг, причому одна або більше з множини кодових книг вибрані на основі порядку елементів просторового компонента, що підлягають стисненню, відносно інших елементів просторового компонента.

60

[1028] Пункт 141541-4J. Пристрій за пунктом 141541-1J, у якому один або більше процесорів додатково виконані з можливістю, при виборі однієї з множини кодових книг, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент не прогнозується з наступного просторового компонента.

[1029] Пункт 141541-5J. Пристрій за пунктом 141541-1J, у якому один або більше процесорів додатково виконані з можливістю, при виборі однієї з множини кодових книг, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент прогнозується з наступного просторового компонента.

[1030] Пункт 141541-6J. Пристрій за пунктом 141541-1J, у якому один або більше процесорів додатково виконані з можливістю, при виборі однієї з множини кодових книг, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент представляє синтетичний аудіооб'єкт у звуковому полі.

[1031] Пункт 141541-7J. Пристрій за пунктом 141541-1J, у якому синтетичний аудіооб'єкт містить аудіооб'єкт, підданий імпульсно-кодовій модуляції (IKM).

[1032] Пункт 141541-8J. Пристрій за пунктом 141541-1J, у якому один або більше процесорів додатково виконані з можливістю, при виборі однієї з множини кодових книг, визначення оцінки кількості бітів, використовуваних для представлення просторового компонента з використанням однієї з множини кодових книг, що підлягають використанню, коли просторовий компонент представляє записаний аудіооб'єкт у звуковому полі.

[1033] Пункт 141541-9J. Пристрій за будь-якою комбінацією пунктів 141541-1J-141541-8J, у якому векторний синтез включає розкладання на сингулярні значення.

[1034] У кожному з різних вищеописаних випадків, варто розуміти, що пристрій 510 аудіокодування може здійснювати спосіб або іншим чином містити засіб для здійснення кожного етапу способу, для якого пристрій 510 аудіокодування виконаний з можливістю здійснення в ряді випадків, цей засіб може містити один або більше процесорів. У ряді випадків, один або більше процесорів можуть представляти процесор спеціального призначення, що виконує інструкції, які зберігаються на довгостроковому зчитуваному комп'ютером носії даних. Іншими словами, різні аспекти методів у кожному з наборів прикладів кодування можуть передбачати довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати спосіб, для виконання якого призначений пристрій 510 аудіокодування.

[1035] На фіг. 40J показана блок-схема, що демонструє ілюстративний пристрій 510J аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 510J аудіокодування може бути аналогічним пристрою 510G аудіокодування тим, що пристрій 510J аудіокодування включає в себе блок 512 стиснення аудіосигналу, блок 514 аудіокодування і блок 516 генерації бітового потоку. Крім того, блок 512 стиснення аудіосигналу пристрою 510J аудіокодування може бути аналогічним блоку пристрою 510G аудіокодування тим, що блок 512 стиснення аудіосигналу включає в себе блок 518 розкладання і блок 520 витягання компонентів звукового поля, що можуть діяти аналогічно відповідним блокам пристрою 510I аудіокодування. У деяких прикладах, пристрій 510J аудіокодування може включати в себе блок 534 квантування, як описано з посиланням на фіг. 40D-40E, для квантування одного або більше векторів з будь-яких векторів $525C U_{DIST}$, векторів $525D U_{BG}$, векторів $525E V_{DIST}^T$ і векторів $525J V_{BG}^T$.

[1036] Блок 512 стиснення аудіосигналу пристрою 510J аудіокодування, однак, може відрізнятися від блока 512 стиснення аудіосигналу пристрою 510G аудіокодування тим, що блок 512 стиснення аудіосигналу пристрою 510J аудіокодування включає в себе додатковий блок, позначений як блок 550 інтерполяції. Блок 550 інтерполяції може представляти блок, що інтерполює підкадри першого кадру аудіо з підкадрів першого кадру аудіо і другого в часі наступного або попереднього кадру аудіо, як більш докладно описано нижче з посиланням на фіг. 45 і 45B. Блок 550 інтерполяції може, при здійсненні цієї інтерполяції, знижувати обчислювальну складність (відносно циклів обробки і/або споживання пам'яті) шляхом потенційного зниження міри, у якій блок 518 розкладання потрібний для розкладання SHC 511. Блок 550 інтерполяції може діяти аналогічно описаному вище відносно блока 550 інтерполяції пристроїв 510H і 510I аудіокодування, показаних у прикладах, наведених на фіг. 40H і 40I.

[1037] У ході експлуатації, блок інтерполяції 200 може інтерполювати один або більше підкадрів першого кадру аудіо з першого розкладання, наприклад матриці 519A' V , частини

першої множини коефіцієнтів 11 сферичних гармонік, включених у перший кадр, і другого розкладання, наприклад матриці $519A' V$, частини другої множини коефіцієнтів 11 сферичних гармонік, включених у другий кадр, для генерації розкладених інтерпольованих коефіцієнтів сферичних гармонік для одного або більше підкадрів.

5 [1038] Блок 550 інтерполяції може одержувати розкладені інтерпольовані коефіцієнти сферичних гармонік для відрізка часу, щонайменше частково, шляхом здійснення інтерполяції відносно першого розкладання першої множини коефіцієнтів сферичних гармонік і другого розкладання другої множини коефіцієнтів сферичних гармонік. Блок 554 згладжування може застосовувати розкладені інтерпольовані коефіцієнти сферичних гармонік для згладжування
10 щонайменше одного з просторових компонентів і часових компонентів першої множини коефіцієнтів сферичних гармонік і другої множини коефіцієнтів сферичних гармонік. Блок 554 згладжування може генерувати згладжені матриці $525C' U_{DIST}$, як описано вище з посиланням на фіг. 37-39. Перше і друге розкладання можуть означати $556 V_1^T$, $556B V_2^T$ на фіг. 40J.

15 [1039] У ряді випадків, V^T або інші V-вектори або V-матриці можуть виводитися в квантованій версії для інтерполяції. Таким чином, вектори V для інтерполяції можуть бути ідентичні векторам V на декодері, що також здійснює інтерполяцію вектора V, наприклад, для відновлення багатовимірного сигналу.

20 [1040] У деяких прикладах, перше розкладання містить першу матрицю $519A' V$, яка представляє праві сингулярні вектори частини першої множини коефіцієнтів 511 сферичних гармонік. Аналогічно, у деяких прикладах, друге розкладання містить другу матрицю $519A' V$, яка представляє праві сингулярні вектори частини другої множини коефіцієнтів сферичних гармонік.

25 [1041] Блок 550 інтерполяції може здійснювати часову інтерполяцію відносно одного або більше підкадрів на основі першої матриці $519A' V$ і другої матриці $519A' V$. Таким чином, блок 550 інтерполяції може інтерпольовати в часі, наприклад, другий, третій і четвертий підкадри з чотирьох повних підкадрів для першого кадру аудіо на основі матриці $519A' V$, розкладеної з першого підкадру першого кадру аудіо, і матриці $519A' V$, розкладеної з першого підкадру другого кадру аудіо. У деяких прикладах, ця часова інтерполяція є лінійною часовою
30 інтерполяцією, де матриця $519A' V$, розкладена з першого підкадру першого кадру аудіо, одержує більшу вагу при інтерполяції другого підкадру першого кадру аудіо, ніж при інтерполяції четвертого підкадру першого кадру аудіо. При інтерполяції третього підкадру, матриці $519A' V$ можуть одержувати однакову вагу. При інтерполяції четвертого підкадру, матриця $519A' V$, розкладена з першого підкадру другого кадру аудіо, може одержувати більшу вагу, ніж матриця $519A' V$, розкладена з першого підкадру першого кадру аудіо.

35 [1042] Іншими словами, лінійна часова інтерполяція може зважувати матриці $519A' V$ з урахуванням близькості одного з підкадрів першого кадру аудіо, що підлягає інтерполяції. Для другого підкадру, що підлягає інтерполяції, матриця $519A' V$, розкладена з першого підкадру першого кадру аудіо, одержує більшу вагу з урахуванням його близькості до другого підкадру, що підлягає інтерполяції, ніж матриця $519A' V$, розкладена з першого підкадру другого кадру аудіо. Вагові коефіцієнти, з цієї причини, можуть бути еквівалентні при інтерполяції третього
40 підкадру на основі матриць $519A' V$. Ваговий коефіцієнт, застосовуваний до матриці $519A' V$, розкладеної з першого підкадру другого кадру аудіо, може бути більше, ніж застосовуваний до матриці $519A' V$, розкладеної з першого підкадру першого кадру аудіо, за умови, що четвертий підкадр, що підлягає інтерполяції, ближче до першого підкадру другого кадру аудіо, ніж перший
45 підкадр першого кадру аудіо.

[1043] У деяких прикладах, блок 550 інтерполяції може проектувати першу матрицю $519A' V$ розкладену з першого підкадру першого кадру аудіо в просторову область для генерації перших проєкційних розкладань. У деяких прикладах, ця проєкція включає в себе проєкцію на сферу (наприклад, з використанням проєкційної матриці, наприклад T-подібної матриці). Потім блок
50 550 інтерполяції може проектувати другу матрицю $519A' V$, розкладену з першого підкадру другого кадру аудіо, в просторову область для генерації других проєкційних розкладань. Потім блок 550 інтерполяції може просторово інтерпольовати (знову ж, за допомогою лінійної інтерполяції) перші проєкційні розкладання і другі проєкційні розкладання для генерації першого просторово інтерпольованого проєкційного розкладання і другого просторово інтерпольованого проєкційного розкладання. Потім блок 550 інтерполяції може інтерпольовати в часі один або
55 більше підкадрів на основі першого просторово інтерпольованого проєкційного розкладання і другого просторово інтерпольованого проєкційного розкладання.

[1044] У прикладах, коли блок 550 інтерполяції проектує просторово і потім у часі матриці $519A' V$, блок 550 інтерполяції може проектувати інтерпольовані в часі коефіцієнти сферичних

гармонік, обумовлені інтерполяцією одного або більше підкадрів, зворотно в область сферичних гармонік, таким чином, генеруючи матрицю 519A V, матрицю 519B S і матрицю 519C U.

[1045] У деяких прикладах, частина першої множини коефіцієнтів сферичних гармонік містить єдиний підкадр першої множини коефіцієнтів 511 сферичних гармонік. У деяких прикладах, частина другої множини коефіцієнтів сферичних гармонік містить єдиний підкадр другої множини коефіцієнтів 511 сферичних гармонік. У деяких прикладах, цей єдиний підкадр, з якого розкладаються матриці 519A' V, є першим підкадром.

[1046] У деяких прикладах, перший кадр поділяється на чотири підкадри. У цих і інших прикладах, частина першої множини коефіцієнтів сферичних гармонік містить тільки перший підкадр із множини коефіцієнтів 511 сферичних гармонік. У цих і інших прикладах, другий кадр поділяється на чотири підкадри, і частина другої множини коефіцієнтів 511 сферичних гармонік містить тільки перший підкадр другої множини коефіцієнтів 511 сферичних гармонік.

[1047] Хоча, у деяких прикладах, тільки перший підкадр кожного кадру аудіо використовується для здійснення інтерполяції, частина першої множини коефіцієнтів сферичних гармонік може містити два з чотирьох підкадрів першої множини коефіцієнтів 511 сферичних гармонік. У цих і інших прикладах, частина другої множини коефіцієнтів 511 сферичних гармонік містить два з чотирьох підкадрів другої множини коефіцієнтів 511 сферичних гармонік.

[1048] Як згадано вище, єдиний пристрій, наприклад пристрій 510J аудіокодування, може здійснювати інтерполяцію, при цьому також розкладаючи частину першої множини коефіцієнтів сферичних гармонік для генерації перших розкладань частини першої множини коефіцієнтів сферичних гармонік. У цих і інших прикладах, блок 518 розкладання може розкласти частину другої множини коефіцієнтів сферичних гармонік для генерації других розкладань частини другої множини коефіцієнтів сферичних гармонік. Хоча вони описані відносно єдиного пристрою, два або більше пристроїв можуть здійснювати методи, описані в цьому винаході, де один із двох пристроїв здійснює розкладання й інший з пристроїв здійснює інтерполяцію відповідно до методів, описаних в цьому винаході.

[1049] У деяких прикладах, блок 518 розкладання може здійснювати розкладання на сингулярні значення відносно частини першої множини коефіцієнтів 511 сферичних гармонік для генерації матриці 519A' V (а також матриці 519B' S і матриці 519C' U, які не показані для спрощення ілюстрації), яка представляє праві сингулярні вектори першої множини коефіцієнтів 511 сферичних гармонік. У цих і інших прикладах, блок 518 розкладання може здійснювати розкладання на сингулярні значення відносно частини другої множини коефіцієнтів 511 сферичних гармонік для генерації матриці 519A' V (а також матриці 519B' S і матриці 519C' U, що не показані для спрощення ілюстрації), яка представляє праві сингулярні вектори другої множини коефіцієнтів сферичних гармонік.

[1050] У деяких прикладах, як згадано вище, перша і друга множини коефіцієнтів сферичних гармонік представляють плоскохвильове представлення звукового поля. У цих і інших прикладах, кожне з першої і другої множини коефіцієнтів 511 сферичних гармонік представляє один або більше змішаних один з одним монофонічних аудіооб'єктів.

[1051] Іншими словами, 3D аудіо на основі сферичних гармонік може бути параметричним представленням 3D поля тиску відносно ортогональних базисних функцій на сфері. Чим вище порядок N представлення, тим потенційно вище просторове розрізнення, і, часто, тим більше кількість коефіцієнтів сферичних гармонік (SH) (для усього $(N+1)^2$ коефіцієнтів). Для багатьох застосувань, може знадобитися, щоб стиснення смуги коефіцієнтів дозволяло ефективно передавати і зберігати коефіцієнти. Ці методи, передбачені в цьому винаході, можуть забезпечувати процес зниження розмірності на кадровій основі з використанням розкладання на сингулярні значення (SVD). Аналіз SVD може розкласти кожен кадр коефіцієнтів на три матриці U, S і V. У деяких прикладах, методи можуть обробляти деякі з векторів у U як направлені компоненти звукового поля, що лежить в основі. Однак, при такій обробці, ці вектори (у U) є розривними від кадру до кадру, хоча вони представляють один і той же характерний аудіокомпонент. Ці порушення неперервності можуть приводити до значних артефактів, коли компонента надходять через перетворюючі аудіокодуери.

[1052] Методи, описані в цьому винаході, дозволяють подолати ці порушення неперервності. Таким чином, методи можуть ґрунтуватися на тому спостереженні, що матрицю V можна інтерпретувати як ортогональні просторові осі в області сферичних гармонік. Матриця U може представляти проекцію даних сферичних гармонік (HOA) відносно цих базисних функцій, де порушення неперервності може приписуватися базисним функціям (V), що змінюють кожен кадр і, таким чином, самі є розривними. У цьому полягає відмінність від аналогічного розкладання, наприклад перетворення Фур'є, де базисні функції, у деяких прикладах, постійні від кадру до кадру. У цьому відношенні, SVD можна розглядати як алгоритм переслідування збігу. Методи,

описані в цьому винаході, можуть передбачати, що блок 550 інтерполяції підтримує неперервність між базисними функціями (V) від кадру до кадру шляхом інтерполяції між ними.

[1053] У деяких прикладах, методи передбачають, що блок 550 інтерполяції розділяє кадр даних SH на чотири підкадри, як описано вище і додатково описано нижче з посиланням на фіг. 45 і 45B. Потім блок 550 інтерполяції може обчислювати SVD для першого підкадру. Аналогічно обчислюється SVD для першого підкадру другого кадру. Для кожного з першого кадру і другого кадру, блок 550 інтерполяції може перетворювати вектори в V у просторову карту шляхом проектування векторів на сферу (з використанням проекційної матриці, наприклад T -подібної матриці). Потім блок 550 інтерполяції може інтерпретувати вектори в V як форми на сфері. Для інтерполяції матриць V для трьох підкадрів між першим підкадром першого кадру і першим підкадром наступного кадру, блок 550 інтерполяції потім може інтерполювати ці просторові форми і потім перетворювати їх зворотно у SH вектори за допомогою оберненої проекційної матриці. Методи цього винаходу, таким чином, можуть забезпечувати плавний перехід між матрицями V .

[1054] На фіг. 41A-41D показані блок-схеми, що демонструють ілюстративний пристрій 540A-540D аудіодекодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для декодування коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 540A аудіодекодування може представляти будь-який пристрій, здатний декодувати аудіодані, наприклад настільний комп'ютер, портативний комп'ютер, робочу станцію, планшет або графічний планшет, спеціалізований звукозаписувальний пристрій, стільниковий телефон (включаючи так називані "смартфони"), пристрій персонального медіапрогравача, персональний ігровий пристрій або пристрій будь-якого іншого типу, здатний декодувати аудіодані.

[1055] У деяких прикладах, пристрій 540A аудіодекодування здійснює процес аудіодекодування, зворотний процесу аудіокодування, здійснюваному будь-яким із пристроїв 510 або 510B аудіокодування за винятком здійснення зниження порядку (як описано вище відносно прикладів, наведених на фіг. 40B-40J), який, у деяких прикладах, використовується пристроями 510B-510J аудіокодування для полегшення видалення сторонніх несуттєвих даних.

[1056] Хоча він показаний як єдиний пристрій, тобто пристрій 540A у прикладі, наведеному на фіг. 41, різні компоненти або блоки, згадані нижче як включені в пристрій 540A, можуть формувати окремі пристрої, зовнішні відносно пристрою 540A. Іншими словами, хоча вони описані в цьому винаході як здійснювані єдиним пристроєм, тобто пристроєм 540A у прикладі, наведеному на фіг. 41, методи можуть реалізуватися або інакше здійснюватися системою, що містить множинні пристрої, причому кожний з цих пристроїв може включати в себе один або більше з різних компонентів або блоків, більш докладно описаних нижче. Відповідно, методи не підлягають обмеженню в цьому відношенні прикладом, наведеним на фіг. 41.

[1057] Як показано в прикладі, наведеному на фіг. 41, пристрій 540A аудіодекодування містить блок 542 витягання, блок 544 аудіодекодування, математичний блок 546 і блок 548 аудіорендеризації. Блок 542 витягання представляє блок, виконаний з можливістю витягання кодованих приведених коефіцієнтів 515B сферичних гармонік фону, кодованих векторів 515A $U_{DIST} * S_{DIST}$ і векторів 525E V_{DIST}^T з бітового потоку 517. Блок 542 витягання виводить кодовані приведені коефіцієнти 515B сферичних гармонік фону і кодовані вектори 515A $U_{DIST} * S_{DIST}$ на блок 544 аудіодекодування, також виводячи і матрицю 525E V_{DIST}^T на математичний блок 546. У цьому відношенні, блок 542 витягання може діяти аналогічно блоку 72 витягання пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5.

[1058] Блок 544 аудіодекодування представляє блок для декодування кодованих аудіоданих (часто згідно зі зворотною схемою аудіодекодування, наприклад схемою декодування AAC) для відновлення векторів 527 $U_{DIST} * S_{DIST}$ і приведених коефіцієнтів 529 сферичних гармонік фону. Блок 544 аудіодекодування виводить вектори 527 $U_{DIST} * S_{DIST}$ і приведені коефіцієнти 529 сферичних гармонік фону на математичний блок 546. У цьому відношенні, блок 544 аудіодекодування може діяти аналогічно блоку 80 психоакустичного декодування пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5.

[1059] Математичний блок 546 може представляти блок, виконаний з можливістю здійснення матричного множення і додавання (а також, у деяких прикладах, будь-якої іншої математичної операції над матрицями). Математичний блок 546 може спочатку здійснювати матричне множення векторів 527 $U_{DIST} * S_{DIST}$ на матрицю 525E V_{DIST}^T . Потім математичний блок 546 може підсумовувати результат множення векторів 527 $U_{DIST} * S_{DIST}$ на матрицю 525E V_{DIST}^T з приведеними коефіцієнтами 529 сферичних гармонік фону (які, знову ж, можуть означати результат множення матриці 525D U_{BG} на матрицю 525B S_{BG} і потім на матрицю 525F V_{BG}^T) з результатом матричного множення векторів 527 $U_{DIST} * S_{DIST}$ на матрицю 525E V_{DIST}^T для

генерації приведеної версії вихідних коефіцієнтів 11 сферичних гармонік, що позначаються як відновлені коефіцієнти 547 сферичних гармонік. Математичний блок 546 може виводити відновлені коефіцієнти 547 сферичних гармонік на блок 548 аудіорендеризації. У цьому відношенні, математичний блок 546 може діяти аналогічно блоку 78 формування переднього плану і блоку 82 формування коефіцієнтів НОА пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5.

[1060] Блок 548 аудіорендеризації представляє блок, виконаний з можливістю рендеризації каналів 549А-549N ("каналів 549", які також можуть, у загальному випадку, іменуватися "багатоканальними аудіоданими 549" або "підведеннями 549 гучномовця"). Блок 548 аудіорендеризації може застосовувати перетворення (часто виражене у формі матриці) до відновлених коефіцієнтів 547 сферичних гармонік. Оскільки відновлені коефіцієнти 547 сферичних гармонік описують звукове поле в трьох вимірах, відновлені коефіцієнти 547 сферичних гармонік представляють аудіоформат, який полегшує рендеризацію багатоканальних аудіоданих 549А у режимі, що допускає адаптацію до більшості декодерно-локальних геометрій гучномовців (яка може означати геометрію гучномовців, що будуть відтворювати багатоканальні аудіодані 549). Додаткова інформація, що стосується рендеризації багатоканальних аудіоданих 549А, описана вище з посиланням на фіг. 48.

[1061] Хоча вони описані в контексті багатоканальних аудіоданих 549А, які є багатоканальними аудіоданими 549 об'ємного звуку, блок 48 аудіорендеризації також може здійснювати форму бінауралізації для бінауралізації відновлених коефіцієнтів сферичних гармонік 549А і, таким чином, генерувати два бінаурально рендеризованих канали 549. Відповідно, методи не підлягають обмеженню формами об'ємного звуку багатоканальних аудіоданих, але можуть включати в себе бінауралізовані багатоканальні аудіодані.

[1062] Різні перераховані нижче пункти можуть представляти різні аспекти методів, описаних у цьому винаході.

[1063] Пункт 132567-1В. Пристрій, наприклад пристрій 540 аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю визначення одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік.

[1064] Пункт 132567-2В. Пристрій за пунктом 132567-1В, у якому один або більше перших векторів містять один або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S , причому матриця U і матриця S генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів додатково виконані з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$ для генерації аудіодекодованої версії одного або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$.

[1065] Пункт 132567-3В. Пристрій за пунктом 132567-1В, у якому один або більше перших векторів містять один або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів додатково виконані з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$ для генерації аудіодекодованої версії одного або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$.

[1066] Пункт 132567-4В. Пристрій за пунктом 132567-3В, у якому один або більше процесорів додатково виконані з можливістю множення векторів $U_{DIST} * S_{DIST}$ на вектори V_{DIST}^T для відновлення тих з множини сферичних гармонік, які представляють характерні компоненти звукового поля.

[1067] Пункт 132567-5В. Пристрій за пунктом 132567-1В, у якому один або більше других векторів містять один або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$, які, до аудіокодування, генерувалися шляхом множення векторів U_{BG} , включених у матрицю U , на вектори S_{BG} , включені в матрицю S , і потім на вектори V_{BG}^T , включені в транспоновану матрицю V , і при цьому матриця S і кожна з матриці U і матриці V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік.

[1068] Пункт 132567-6В. Пристрій за пунктом 132567-1В, у якому один або більше других векторів містять один або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$, які, до аудіокодування,

генерувалися шляхом множення векторів U_{BG} , включених у матрицю U , на вектори S_{BG} , включені в матрицю S , і потім на вектори V_{BG}^T , включені в транспоновану матрицю V , і при цьому матриця S , матриця U і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів додатково виконані з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$ для генерації одного або більше аудіодекодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$.

[1069] Пункт 132567-7B. Пристрій за пунктом 132567-1B, у якому один або більше перших векторів містять один або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів додатково виконані з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{DIST} * S_{DIST}$ для генерації одного або більше векторів $U_{DIST} * S_{DIST}$ і множення векторів $U_{DIST} * S_{DIST}$ на вектори V_{DIST}^T для відновлення тих з множини коефіцієнтів сферичних гармонік, які описують характерні компоненти звукового поля, причому один або більше других векторів містять один або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$, які, до аудіокодування, генерувалися шляхом множення векторів U_{BG} , включених у матрицю U , на вектори S_{BG} , включені в матрицю S , і потім на вектори V_{BG}^T , включені в транспоновану матрицю V , і при цьому один або більше процесорів додатково виконані з можливістю аудіодекодування одного або більше аудіокодованих векторів $U_{BG} * S_{BG} * V_{BG}^T$ для відновлення щонайменше частини з множини коефіцієнтів сферичних гармонік, що описують фонові компоненти звукового поля, і додавання множини коефіцієнтів сферичних гармонік, що описують характерні компоненти звукового поля, до щонайменше частини з множини коефіцієнтів сферичних гармонік, що описують фонові компоненти звукового поля, для генерації реконструйованої версії з множини коефіцієнтів сферичних гармонік.

[1070] Пункт 132567-8B. Пристрій за пунктом 132567-1B, у якому один або більше перших векторів містять один або більше векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів додатково виконані з можливістю визначення значення D , що вказує кількість векторів, які підлягають витяганню з бітового потоку, для формування одного або більше векторів $U_{DIST} * S_{DIST}$ і одного або більше векторів V_{DIST}^T .

[1071] Пункт 132567-9B. Пристрій за пунктом 132567-10B, у якому один або більше перших векторів містять один або більше векторів $U_{DIST} * S_{DIST}$, які, до аудіокодування, згенеровані шляхом множення одного або більше аудіокодованих векторів U_{DIST} матриці U на один або більше векторів S_{DIST} матриці S і один або більше векторів V_{DIST}^T транспонованої матриці V , причому матриця U , матриця S і матриця V згенеровані щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і при цьому один або більше процесорів додатково виконані з можливістю визначення значення D на основі кадрів аудіо, що вказує кількість векторів, які підлягають витяганню з бітового потоку, для формування одного або більше векторів $U_{DIST} * S_{DIST}$ і одного або більше векторів V_{DIST}^T .

[1072] Пункт 132567-1G. Пристрій, наприклад пристрій 540 аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю визначення одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно багатоканальних аудіоданих, що представляють щонайменше частину звукового поля.

[1073] Пункт 132567-2G. Пристрій за пунктом 132567-1G, у якому багатоканальні аудіодані містять множину коефіцієнтів сферичних гармонік.

[1074] Пункт 132567-3G. Пристрій за пунктом 132567-2G, у якому один або більше процесорів додатково виконані з можливістю здійснення будь-якої комбінації пунктів 132567-2B-132567-9B.

[1075] З кожного з різних описаних вище пунктів випливає, що будь-який з пристроїв 540A-540D аудіодекодування може здійснювати спосіб або іншим чином містити засіб для здійснення кожного етапу способу, для якого пристрої 540A-540D аудіодекодування виконані з можливістю

здійснення в ряді випадків, цей засіб може містити один або більше процесорів. У ряді випадків, один або більше процесорів можуть представляти процесор спеціального призначення, який виконує інструкції, що зберігаються на довгостроковому зчитуваному комп'ютером носії даних. Іншими словами, різні аспекти методів у кожному з наборів прикладів кодування можуть передбачати довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати спосіб, для здійснення якого призначені пристрої 540A-540D аудіодекодування.

[1076] Наприклад, пункт 132567-10B може бути виведений з вищенаведеного пункту 132567-1B як спосіб, що включає спосіб, який включає: визначення одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, що представляють звукове поле.

[1077] У порядку іншого прикладу, пункт 132567-11B може бути виведений з вищенаведеного пункту 132567-1B як пристрій, наприклад пристрій 540 аудіодекодування, який містить: засіб для визначення одного або більше перших векторів, що описують характерні компоненти звукового поля, і одного або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік; і засіб для збереження одного або більше перших векторів і одного або більше других векторів.

[1078] У порядку ще одного прикладу, пункт 132567-12B може бути виведений з вищенаведеного пункту 132567-1B як довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам визначати один або більше перших векторів, що описують характерні компоненти звукового поля, і один або більше других векторів, що описують фонові компоненти звукового поля, причому один або більше перших векторів і один або більше других векторів генеруються щонайменше шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, включених в аудіодані амбіофонії більш високого порядку, що описують звукове поле.

[1079] Різні пункти аналогічно можуть бути виведені з пунктів 132567-2B-132567-9B для різних пристроїв, способів і довгострокових зчитуваних комп'ютером носіїв даних, виведених, як згадано вище. Те ж саме можна здійснювати для різних інших пунктів, перерахованих протягом цього розкриття.

[1080] На фіг. 41B показана блок-схема, що демонструє приклад пристрою 540B аудіодекодування, який може здійснювати різні аспекти методів, описаних у цьому винаході, для декодування коефіцієнтів сферичних гармонік, що описують дво- або тривимірні звукові поля. Пристрій 540B аудіодекодування може бути аналогічним пристрою 540 аудіодекодування, за винятком того, що, у деяких прикладах, блок 542 витягання може витягати переупорядковані вектори 539 V_{DIST}^T замість векторів 525E V_{DIST}^T . В інших прикладах, блок 542 витягання може витягати вектори 525E V_{DIST}^T і потім переупорядковувати ці вектори 525E V_{DIST}^T на основі інформації переупорядкування, зазначеної в бітовому потоці або виведеної (за допомогою аналізу інших векторів), для визначення переупорядкованих векторів 539 V_{DIST}^T . У цьому відношенні, блок 542 витягання може діяти аналогічно блоку 72 витягання пристрою 24 аудіодекодування, показаного в прикладі, наведеному на фіг. 5. У будь-якому випадку, блок 542 витягання може виводити переупорядковані вектори 539 V_{DIST}^T на математичний блок 546, де процес, описаний вище відносно відновлення коефіцієнтів сферичних гармонік, може здійснюватися відносно цих переупорядкованих векторів 539 V_{DIST}^T .

[1081] Таким чином, методи можуть передбачати, що пристрій 540B аудіодекодування аудіодекодує переупорядковані один або більше векторів, що представляють характерні компоненти звукового поля, причому переупорядковані один або більше векторів переупорядковуються для полегшення стиснення одного або більше векторів. У цих і інших прикладах, пристрій 540B аудіодекодування може рекомбінувати переупорядковані один або більше векторів з переупорядкованими одним або більше додатковими векторами для відновлення коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля. У цих і інших прикладах, пристрій 540B аудіодекодування може потім відновлювати множину коефіцієнтів сферичних гармонік на основі коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля, і коефіцієнтів сферичних гармонік, що представляють фонові компоненти звукового поля.

[1082] Таким чином, різні аспекти методів можуть передбачати, що пристрій 540В аудіодекодування виконаний з можливістю декодування переупорядкованих одного або більше векторів відповідно до наступних пунктів.

5 [1083] Пункт 133146-1F. Пристрій, наприклад пристрій 540В аудіокодування, який містить: один або більше процесорів, виконаних з можливістю визначення кількості векторів, що відповідають компонентам у звуковому полі.

[1084] Пункт 133146-2F. Пристрій за пунктом 133146-1F, у якому один або більше процесорів виконані з можливістю визначення кількості векторів після здійснення зниження порядку відповідно до будь-якої комбінації вищеописаних випадків.

10 [1085] Пункт 133146-3F. Пристрій за пунктом 133146-1F, у якому один або більше процесорів додатково виконані з можливістю здійснення зниження порядку відповідно до будь-якої комбінації вищеописаних випадків.

[1086] Пункт 133146-4F. Пристрій за пунктом 133146-1F, у якому один або більше процесорів виконані з можливістю визначення кількості векторів зі значення, зазначеного в бітовому потоці, і при цьому один або більше процесорів додатково виконані з можливістю виділення бітового потоку на основі визначеної кількості векторів для ідентифікації одного або більше векторів у бітовому потоці, що представляють характерні компоненти звукового поля.

20 [1087] Пункт 133146-5F. Пристрій за пунктом 133146-1F, у якому один або більше процесорів виконані з можливістю визначення кількості векторів зі значення, зазначеного в бітовому потоці, і при цьому один або більше процесорів додатково виконані з можливістю виділення бітового потоку на основі визначеної кількості векторів для ідентифікації одного або більше векторів у бітовому потоці, що представляють фонові компоненти звукового поля.

[1088] Пункт 133143-1C. Пристрій, наприклад пристрій 540В аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю переупорядковування переупорядкованих одного або більше векторів, що представляють характерні компоненти звукового поля.

25 [1089] Пункт 133143-2C. Пристрій за пунктом 133143-1C, у якому один або більше процесорів додатково виконані з можливістю визначення переупорядкованих одного або більше векторів і визначення інформації переупорядковування, що описує, як були переупорядковані переупорядковані один або більше векторів, у якому один або більше процесорів додатково виконані з можливістю, при переупорядкуванні переупорядкованих одного або більше векторів, переупорядковування переупорядкованих одного або більше векторів на основі визначеної інформації переупорядковування.

30 [1090] Пункт 133143-3C. Пристрій за пунктом 1C, у якому переупорядковані один або більше векторів містять один або більше переупорядкованих перших векторів відповідно до будь-якої комбінації за пунктами 1A-18A або будь-якої комбінації за пунктами 1B-19B, і при цьому один або більше перших векторів визначаються згідно зі способом відповідно до будь-якої комбінації за пунктами 1A-18A або будь-якої комбінації за пунктами 1B-19B.

40 [1091] Пункт 133143-4D. Пристрій, наприклад пристрій 540В аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю аудіодекодування переупорядкованих одного або більше векторів, що представляють характерні компоненти звукового поля, причому переупорядковані один або більше векторів переупорядковуються для полегшення стиснення одного або більше векторів.

45 [1092] Пункт 133143-5D. Пристрій за пунктом 133143-4D, у якому один або більше процесорів додатково виконані з можливістю рекомбінувати переупорядковані один або більше векторів з переупорядкованими одним або більш додатковими векторами для відновлення коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля.

[1093] Пункт 133143-6D. Пристрій за пунктом 133143-5D, у якому один або більше процесорів додатково виконані з можливістю відновлення множини коефіцієнтів сферичних гармонік на основі коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля, і коефіцієнтів сферичних гармонік, що представляють фонові компоненти звукового поля.

50 [1094] Пункт 133143-1E. Пристрій, наприклад пристрій 540В аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю переупорядковування одного або більше векторів для генерації переупорядкованих одного або більше перших векторів і, таким чином, полегшення кодування успадкованим аудіокодером, причому один або більше векторів представляють характерні компоненти звукового поля, і аудіокодування переупорядкованих одного або більше векторів з використанням успадкованого аудіокодера для генерації кодованої версії переупорядкованих одного або більше векторів.

60 [1095] Пункт 133143-2E. Пристрій за пунктом 1E, у якому переупорядковані один або більше векторів містять один або більше переупорядкованих перших векторів відповідно до будь-якої

комбінації за пунктами 1A-18A або будь-якої комбінації за пунктами 1B-19B, і при цьому один або більше перших векторів визначаються згідно зі способом відповідно до будь-якої комбінації за пунктами 1A-18A або будь-якої комбінації за пунктами 1B-19B.

[1096] На фіг. 41C показана блок-схема, що демонструє інший ілюстративний пристрій 540C аудіокодування. Пристрій 540C аудіодекодування може представляти будь-який пристрій, здатний декодувати аудіодані, наприклад настільний комп'ютер, портативний комп'ютер, робочу станцію, планшет або графічний планшет, спеціалізований звукозаписувальний пристрій, стільниковий телефон (включаючи так називані "смартфони"), пристрій персонального медіапрогравача, персональний ігровий пристрій або пристрій будь-якого іншого типу, здатний декодувати аудіодані.

[1097] У прикладі, наведеному на фіг. 41C, пристрій 540C аудіодекодування здійснює процес аудіодекодування, зворотний процесу аудіокодування, здійснюваному будь-яким із пристроїв 510B-510E аудіокодування, за винятком здійснення зниження порядку (як описано вище відносно прикладів, наведених на фіг. 40B-40J), яке, у деяких прикладах, використовується пристроєм 510B-510J аудіокодування для полегшення видалення сторонніх несуттєвих даних.

[1098] Хоча він показаний як єдиний пристрій, тобто пристрій 540C у прикладі, наведеному на фіг. 41C, різні компоненти або блоки, згадані нижче як включені в пристрій 540C, можуть формувати окремі пристрої, зовнішні відносно пристрою 540C. Іншими словами, хоча вони описані в цьому винаході як здійснювані єдиним пристроєм, тобто пристроєм 540C у прикладі, наведеному на фіг. 41C, методи можуть реалізовуватися або інакше здійснюватися системою, що містить множинні пристрої, причому кожний з цих пристроїв може включати в себе один або більше з різних компонентів або блоків, більш докладно описаних нижче. Відповідно, методи не підлягають обмеженню в цьому відношенні прикладом, наведеним на фіг. 41C.

[1099] Крім того, пристрій 540C аудіокодування може бути аналогічним пристрою 540B аудіокодування. Однак блок 542 витягання може визначати один або більше векторів $521 V_{SMALL}^T$ з бітового потоку 517 замість переупорядкованих векторів $539 V_{Q_DIST}^T$ або векторів $525 E V_{DIST}^T$ (як у випадку, описаному відносно пристрою 510 аудіокодування на фіг. 40). У результаті, блок 542 витягання може передавати вектори $521 V_{SMALL}^T$ на математичний блок 546.

[1100] Крім того, блок 542 витягання може визначати аудіокодовані модифіковані коефіцієнти $515B'$ сферичних гармонік фону з бітового потоку 517, передаючи ці коефіцієнти $515B'$ на блок 544 аудіодекодування, який може аудіодекодувати кодовані модифіковані коефіцієнти $515B$ сферичних гармонік фону для відновлення модифікованих коефіцієнтів 537 сферичних гармонік фону. Блок 544 аудіодекодування може передавати ці модифіковані коефіцієнти 537 сферичних гармонік фону на математичний блок 546.

[1101] Потім математичний блок 546 може множити аудіодекодовані (і, можливо, неупорядковані) вектори $527 U_{DIST}^T S_{DIST}'$ на один або більше векторів $521 V_{SMALL}^T$ для відновлення характерних коефіцієнтів сферичних гармонік більш високого порядку. Потім математичний блок 546 може підсумовувати характерні коефіцієнти сферичних гармонік більш високого порядку з модифікованими коефіцієнтами 537 сферичних гармонік фону для відновлення множини коефіцієнтів 511 сферичних гармонік або деякої їх похідної (яка може бути похідною внаслідок зниження порядку, здійснюваного на блоці 510E кодера).

[1102] Таким чином, методи можуть передбачати, що пристрій 540C аудіодекодування визначає, з бітового потоку, щонайменше один з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, які рекомбінували з коефіцієнтами сферичних гармонік фону, для зниження кількості бітів, які необхідно виділяти одному або більше векторам у бітовому потоці, причому коефіцієнти сферичних гармонік описують звукове поле і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[1103] Різні аспекти методів можуть, у цьому відношенні, передбачати, що пристрій 540C аудіодекодування, у ряді випадків, виконаний з можливістю визначення, з бітового потоку, щонайменше одного з одного або більше векторів, розкладених з коефіцієнтів сферичних гармонік, які рекомбінували з коефіцієнтами сферичних гармонік фону, причому коефіцієнти сферичних гармонік описують звукове поле і при цьому коефіцієнти сферичних гармонік фону описують один або більше фонових компонентів одного і того ж звукового поля.

[1104] У цих і інших випадках, пристрій 540C аудіодекодування виконаний з можливістю одержання, з бітового потоку, першої частини коефіцієнтів сферичних гармонік, які мають порядок, що дорівнює N_{BG} .

[1105] У цих і інших випадках, пристрій 540C аудіодекодування додатково виконаний з можливістю одержання, з бітового потоку, першої аудіокодованої частини коефіцієнтів сферичних гармонік, які мають порядок, що дорівнює N_{BG} , і аудіодекодування аудіокодованої

першої частини коефіцієнтів сферичних гармонік для генерації першої частини коефіцієнтів сферичних гармонік.

[1106] У цих і інших випадках, щонайменше один з одного або більше векторів містить один або більше векторів V_{SMALL}^T , причому один або більше векторів V_{SMALL}^T визначені із
5 транспонованої матриці V , генерованої шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік.

[1107] У цих і інших випадках, щонайменше один з одного або більше векторів містить один або більше векторів V_{SMALL}^T , причому один або більше векторів V_{SMALL}^T визначені із
10 транспонованої матриці V , генерованої шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і пристрій 540°C аудіодекодування додатково виконаний з можливістю одержання, з бітового потоку, одного або більше векторів $U_{DIST} * S_{DIST}$, що виводяться з матриці U і матриці S , які обидві згенеровані шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік і
множення векторів $U_{DIST} * S_{DIST}$ на вектори V_{SMALL}^T .

[1108] У цих і інших випадках, щонайменше один з одного або більше векторів містить один або більше векторів V_{SMALL}^T , причому один або більше векторів V_{SMALL}^T визначені із
20 транспонованої матриці V , генерованої шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і пристрій 540C аудіодекодування додатково виконаний з можливістю одержання, з бітового потоку, одного або більше векторів $U_{DIST} * S_{DIST}$, що виводяться з матриці U і матриці S , які обидві згенеровані шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, множення векторів $U_{DIST} * S_{DIST}$ на вектори V_{SMALL}^T для відновлення характерних коефіцієнтів сферичних гармонік фону більш високого порядку і додавання коефіцієнтів сферичних гармонік фону, що включають в себе характерні коефіцієнти сферичних гармонік фону більш низького
25 порядку, до характерних коефіцієнтів сферичних гармонік фону більш високого порядку для відновлення, щонайменше частково, множини коефіцієнтів сферичних гармонік.

[1109] У цих і інших випадках, щонайменше один з одного або більше векторів містить один або більше векторів V_{SMALL}^T , причому один або більше векторів V_{SMALL}^T визначені із
30 транспонованої матриці V , генерованої шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, і пристрій 540C аудіодекодування додатково виконаний з можливістю одержання, з бітового потоку, одного або більше векторів $U_{DIST} * S_{DIST}$, виведених з матриці U і матриці S , які обидві згенеровані шляхом здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік, множення векторів $U_{DIST} * S_{DIST}$ на вектори V_{SMALL}^T для відновлення характерних коефіцієнтів
35 сферичних гармонік фону більш високого порядку, додавання коефіцієнтів сферичних гармонік фону, що включають в себе характерні коефіцієнти сферичних гармонік фону більш низького порядку, до характерних коефіцієнтів сферичних гармонік фону більш високого порядку для відновлення, щонайменше частково, множини коефіцієнтів сферичних гармонік і рендеризації відновленої множини коефіцієнтів сферичних гармонік.

[1110] На фіг. 41D показана блок-схема, що демонструє інший ілюстративний пристрій 540D аудіокодування. Пристрій 540D аудіокодування може представляти будь-який пристрій,
40 здатний декодувати аудіодані, наприклад настільний комп'ютер, портативний комп'ютер, робочу станцію, планшет або графічний планшет, спеціалізований звукозаписувальний пристрій, стільниковий телефон (включаючи так називані "смартфони"), пристрій персонального медіапрогравача, персональний ігровий пристрій або пристрій будь-якого іншого типу, здатний декодувати аудіодані.

[1111] У прикладі, наведеному на фіг. 41D, пристрій 540D аудіодекодування здійснює процес аудіодекодування, зворотний процесу аудіокодування, здійснюваному будь-яким із пристроїв 510B-510J аудіокодування за винятком здійснення зниження порядку (як описано вище відносно
50 прикладів, наведених на фіг. 40B-40J), що, у деяких прикладах, використовується пристроями 510B-510J аудіокодування для полегшення видалення сторонніх несуттєвих даних.

[1112] Хоча він показаний як єдиний пристрій, тобто пристрій 540D у прикладі, наведеному на фіг. 41D, різні компоненти або блоки, згадані нижче як включені в пристрій 540D, можуть формувати окремі пристрої, зовнішні відносно пристрою 540D. Іншими словами, хоча вони
55 описані в цьому винаході як здійснювані єдиним пристроєм, тобто пристроєм 540D у прикладі, наведеному на фіг. 41D, методи можуть реалізуватися або інакше здійснюватися системою, що містить множинні пристрої, причому кожний з цих пристроїв може включати в себе один або більше з різних компонентів або блоків, більш докладно описаних нижче. Відповідно, методи не підлягають обмеженню в цьому відношенні прикладом, наведеним на фіг. 41D.

[1113] Крім того, пристрій 540D аудіодекодування може бути аналогічним пристрою 540B аудіодекодування, за винятком того, що пристрій 540D аудіодекодування здійснює додаткове зняття стиснення V, у загальному випадку, зворотне стисненню, здійснюваному блоком 552 стиснення V, описаним вище з посиланням на фіг. 40I. У прикладі, наведеному на фіг. 41D, блок 542 витягання включає в себе блок 555 зняття стиснення V, який здійснює це зняття стиснення V стиснених просторових компонентів 539', включених у бітовий потік 517 (і, у загальному випадку, зазначених відповідно до прикладу, показаного на одній з фіг. 10B і 10C). Блок 555 зняття стиснення V може знімати стиснення векторів 539 V_{DIST}^T на основі наступного рівняння:

$$\hat{v}_q = \begin{cases} 0, & \text{if } cid = 0 \\ sgn * (2^{cid-1} + residual), & \text{if } cid \neq 0. \end{cases}$$

Іншими словами, блок 555 зняття стиснення V може спочатку виділяти значення nbits з бітового потоку 517 і ідентифікувати належний набір з п'яти таблиць кодів Хаффмана для використання при декодуванні коду Хаффмана, що представляє cid. На підставі інформації режиму прогнозування і хаффманівського кодування, зазначеної в бітовому потоці 517, і, можливо, порядку елемента просторового компонента відносно інших елементів просторового компонента, блок 555 зняття стиснення V може ідентифікувати правильну одну з п'яти таблиць Хаффмана, заданих для виділеного значення nbits. З використанням цієї таблиці Хаффмана, блок 555 зняття стиснення V може декодувати значення cid з коду Хаффмана. Блок 555 зняття стиснення V може потім виділяти знаковий біт і залишковий блоковий код, декодувати залишковий блоковий код для ідентифікації залишку. Відповідно до вищенаведеного рівняння, блок 555 зняття стиснення V може декодувати один з векторів 539 V_{DIST}^T .

[1114] Вищевикладене можна коротко представити в нижченаведеній таблиці синтаксису:

Таблиця декодованих векторів

Синтаксис	Кільк. бітів	Мнемоніка
decodeVVec(i)		
{		
switch CodedVVecLength {		
case 0: //повний вектор		
VVecLength=NumOfHoaCoeffs;		
for (m=0; m< VVecLength; ++m){		
VecCoeff[m]=m+1; }		
break;		
case 1: // більш низькі порядки видаляються		
VVecLength=NumOfHoaCoeffs	-	
MinNumOfCoeffsForAmbHOA;		
for (m=0; m< VVecLength; ++m){		
VecCoeff[m]=m+MinNumOfCoeffsForAmbHOA+1;		
}		
break;		
case 2:		
VVecLength=NumOfHoaCoeffs	-	
MinNumOfCoeffsForAmbHOA	-	
NumOfAddAmbHoaChan;		
n=0;		

for (m=0; m<NumOfHoaCoeffs- MinNumOfCoeffsForAmbHOA; ++m){		
c=m+MinNumOfCoeffsForAmbHOA+1;		
if (ismember(c, AmbCoeffIdx) == 0){		
VecCoeff[n]=c;		
n++;		
}		
}		
break;		
case 3:		
VVecLength=NumOfHoaCoeffs - NumOfAddAmbHoaChan;		
n=0;		
for (m=0; m<NumOfHoaCoeffs; ++m){		
c=m+1;		
if (ismember(c, AmbCoeffIdx) == 0){		
VecCoeff[n]=c;		
n++;		
}		
}		
}		
if (NbBitsQ[i] == 5) { /* однорідний квантователь */		
for (m=0; m< VVecLength; ++m){		
VVec(k)[i][m]=(VecValue/128.0) - 1.0;	8	uimsbf
}		
}		
else { /* хатфманівське декодування */		
for (m=0; m< VVecLength; ++m){		
Idx=5;		
If (CbFlag[i] == 1) {		
idx=(min(3, max(1, ceil(sqrt(VecCoeff[m]) - 1))));		
}		
else if (PFlag[i] == 1) {idx=4;}		
cid=huffDecode(huffmannTable[NbBitsQ], codebook[idx]; huffVal);	динамічна	huffDecode

if (cid > 0) {		
aVal=sgn=(sgnVal * 2) - 1;	1	bslbf
if (cid > 1) {		
aVal=sgn * (2.0^(cid -1)+intAddVal);	cid-1	uimsbf
}		
} else {aVal=0.0;}		
}		
}		
}		
Примітка:		
- функція кодера для однорідного кантувача:		$\min(255, \text{round}((x+1.0) * 128.0));$
- кільк. бітів для мнемоніки huffDecode є динамічною.		

[1115] У вищенаведеній таблиці синтаксису, перший оператор переключення з чотирма випадками (випадками 0-3) забезпечує можливість визначення довжини вектора V_{DIST}^T відносно числа коефіцієнтів. Перший випадок, case 0, указує, що задані всі коефіцієнти для векторів V_{DIST}^T . Другий випадок, case 1, указує, що задані тільки коефіцієнти вектора V_{DIST}^T , що відповідають порядку, більшому, ніж MinNumOfCoeffsForAmbHOA, що може позначати те, що іменується вище $(N_{\text{DIST}}+1)-(N_{\text{BG}}+1)$. Третій випадок, case 2, аналогічний другому випадку, але додатково віднімає коефіцієнти, ідентифіковані NumOfAddAmbHoaChan, що позначає змінну для вказання додаткових каналів (де "канали" означають конкретний коефіцієнт, що відповідає визначеному порядку, комбінації підпорядків), які відповідають порядку, що перевищує порядок N_{BG} . Четвертий випадок, case 3, указує, що задані тільки коефіцієнти вектора V_{DIST}^T , що залишилися після видалення коефіцієнтів, ідентифікованих за допомогою NumOfAddAmbHoaChan.

[1116] Після цього оператора переключення, рішення, чи здійснювати уніфіковане деквантування, приймається залежно від NbitsQ (або, як позначено вище, pbits), яке, якщо не дорівнює 5, приводить до застосування хафманівського декодування. Вищезгадане значення cid дорівнює двом молодшим бітам значення NbitsQ. Розглянутий вище режим прогнозування позначається у вищенаведеній таблиці синтаксису як PFlag, тоді як біт інформації HT позначається у вищенаведеній таблиці синтаксису як CbFlag. Синтаксис, що залишився, указує, як декодування виникає по суті, аналогічним чином описаному вище.

[1117] Таким чином, методи цього винаходу можуть передбачати, що пристрій 540D аудіодекодування одержує бітовий потік, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік і зняття стиснення стисненої версії просторового компонента для одержання просторового компонента.

[1118] Крім того, методи можуть передбачати, що пристрій 540D аудіодекодування знімає стиснення стисненої версії просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[0100] Таким чином, пристрій 540D аудіокодування може здійснювати різні аспекти методів, викладених нижче відносно наступних пунктів.

[1119] Пункт 141541-1B. Пристрій, який містить:

[1120] один або більше процесорів, виконаних з можливістю одержання бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік, і зняття стиснення стисненої версії просторового компонента для одержання просторового компонента.

[1121] Пункт 141541-2B. Пристрій за пунктом 141541-1B, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента, і при цьому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, знімати стиснення стисненої версії

просторового компонента, щонайменше частково, на основі режиму прогнозування для одержання просторового компонента.

[1122] Пункт 141541-3B. Пристрій за будь-якою комбінацією пунктів 141541-1B і 141541-2B, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента, і при цьому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, знімати стиснення стисненої версії просторового компонента, щонайменше частково, на основі інформації таблиці Хаффмана.

[1123] Пункт 141541-4B. Пристрій за будь-якою комбінацією пунктів 141541-1B-141541-3B, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента, і при цьому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, знімати стиснення стисненої версії просторового компонента, щонайменше частково, на основі значення.

[1124] Пункт 141541-5B. Пристрій за пунктом 141541-4B, у якому значення містить значення nbits.

[1125] Пункт 141541-6B. Пристрій за будь-якою комбінацією пунктів 141541-4B і 141541-5B, у якому бітовий потік містить стиснену версію з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, при цьому значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів, і при цьому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, знімати стиснення множини стисненої версії просторового компонента, щонайменше частково, на основі значення.

[1126] Пункт 141541-7B. Пристрій за будь-якою комбінацією пунктів 141541-1B-141541-6B, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент, і при цьому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, знімати стиснення стисненої версії просторового компонента, щонайменше частково, на основі коду Хаффмана.

[1127] Пункт 141541-8B. Пристрій за будь-якою комбінацією пунктів 141541-1B-141541-7B, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення, і при цьому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, знімати стиснення стисненої версії просторового компонента, щонайменше частково, на основі знакового біта.

[1128] Пункт 141541-9B. Пристрій за будь-якою комбінацією пунктів 141541-1B-141541-8B, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення залишкового значення просторового компонента, і при цьому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, знімати стиснення стисненої версії просторового компонента, щонайменше частково, на основі коду Хаффмана.

[1129] Пункт 141541-10B. Пристрій за будь-якою комбінацією пунктів 141541-1B-141541-10B, у якому векторний синтез включає розкладання на сингулярні значення.

[1130] Крім того, пристрій 540D аудіодекодування може бути виконаний з можливістю здійснення різних аспектів методів, викладених нижче відносно наступних пунктів.

[1131] Пункт 141541-1C. Пристрій, наприклад пристрій 540D аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю зняття стиснення стисненої версії просторового компонента звукового поля, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

[1132] Пункт 141541-2C. Пристрій за будь-якою комбінацією пунктів 141541-1C і 141541-2C, у якому один або більше процесорів додатково виконані з можливістю, при знятті стиснення стисненої версії просторового компонента, одержувати ідентифікатор категорії, що ідентифікує категорію, до якої був віднесений просторовий компонент при стисненні, одержувати знак, що ідентифікує, чи має просторовий компонент позитивне або негативне значення, одержувати залишкове значення, пов'язане зі стисненою версією просторового компонента, і знімати

стиснення стисненої версії просторового компонента на основі ідентифікатора категорії, знака і залишкового значення.

5 [1133] Пункт 141541-3C. Пристрій за пунктом 141541-2C, у якому один або більше процесорів додатково виконані з можливістю, при одержанні ідентифікатора категорії, одержувати код Хаффмана, що представляє ідентифікатор категорії, і декодувати код Хаффмана для одержання ідентифікатора категорії.

10 [1134] Пункт 141541-4C. Пристрій за пунктом 141541-3C, у якому один або більше процесорів додатково виконані з можливістю, при декодуванні коду Хаффмана, ідентифікувати таблицю Хаффмана, використовувану для декодування коду Хаффмана, на основі, щонайменше частково, відносної позиції просторового компонента у векторі, що вказує множину просторових компонентів.

15 [1135] Пункт 141541-5C. Пристрій за будь-якою комбінацією пунктів 141541-3C і 141541-4C, у якому один або більше процесорів додатково виконані з можливістю, при декодуванні коду Хаффмана, ідентифікувати таблицю Хаффмана, використовувану для декодування коду Хаффмана, на основі, щонайменше частково, режиму прогнозування, використовуваного при стисненні просторового компонента.

20 [1136] Пункт 141541-6C. Пристрій за будь-якою комбінацією пунктів 141541-3C-141541-5C, у якому один або більше процесорів додатково виконані з можливістю, при декодуванні коду Хаффмана, ідентифікувати таблицю Хаффмана, використовувану для декодування коду Хаффмана, на основі, щонайменше частково, інформації таблиці Хаффмана, пов'язаної зі стисненою версією просторового компонента.

25 [1137] Пункт 141541-7C. Пристрій за пунктом 141541-3C, у якому один або більше процесорів додатково виконані з можливістю, при декодуванні коду Хаффмана, ідентифікувати таблицю Хаффмана, використовувану для декодування коду Хаффмана, на основі, щонайменше частково, відносної позиції просторового компонента у векторі, що вказує множину просторових компонентів, режим прогнозування, використовуваний при стисненні просторового компонента, і інформацію таблиці Хаффмана, пов'язану зі стисненою версією просторового компонента.

30 [1138] Пункт 141541-8C. Пристрій за пунктом 141541-2C, у якому один або більше процесорів додатково виконані з можливістю, при одержанні залишкового значення, декодувати блоковий код, що представляє залишкове значення, для одержання залишкового значення.

[1139] Пункт 141541-9C. Пристрій за будь-якою комбінацією пунктів 141541-1C-141541-8C, у якому векторний синтез включає розкладання на сингулярні значення.

35 [1140] Крім того, пристрій 540D аудіодекодування може бути виконаний з можливістю здійснення різних аспектів методів, викладених нижче відносно наступних пунктів.

40 [1141] Пункт 141541-1G. Пристрій, наприклад пристрій 540D аудіодекодування, який містить: один або більше процесорів, виконаних з можливістю ідентифікації кодової книги Хаффмана для використання, при знятті стиснення стисненої версії поточного просторового компонента, множини стиснених просторових компонентів на основі порядку стисненої версії поточного просторового компонента відносно стиснених просторових компонентів, що залишилися з множини, причому просторовий компонент генерується шляхом здійснення векторного синтезу відносно множини коефіцієнтів сферичних гармонік.

45 [1142] Пункт 141541-2G. Пристрій за пунктом 141541-1G, у якому один або більше процесорів додатково виконані з можливістю здійснення будь-якої комбінації етапів, згаданих у пунктах 141541-1D-141541-10D і пунктах 141541-1E-141541-9E.

50 [1143] На фіг. 42A-42C показані блок-схеми, що більш детально демонструють блок 528A зниження порядку, показаний у прикладах, наведених на фіг. 40B-40J. На фіг. 42 показана блок-схема, що демонструє блок 528 зниження порядку, який може представляти один приклад блока 528A зниження порядку, показаного на фіг. 40B-40J. Блок 528A зниження порядку може приймати або інакше визначати цільову бітову швидкість 535 і здійснювати зниження порядку відносно коефіцієнтів 531 сферичних гармонік фону тільки на основі цієї цільової бітової швидкості 535. У деяких прикладах, блок 528A зниження порядку може звертатися до таблиці або іншої структури даних з використанням цільової бітової швидкості 535 для ідентифікації тих порядків і/або підпорядків, що підлягають видаленню з коефіцієнтів 531 сферичних гармонік фону, для генерації приведених коефіцієнтів 529 сферичних гармонік фону.

55 [1144] Таким чином, методи можуть передбачати, що пристрій аудіокодування, наприклад пристрої аудіокодування 510B-410J, здійснює, на основі цільової бітової швидкості 535, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань, наприклад коефіцієнтів 531 сферичних гармонік фону, для генерації приведених коефіцієнтів

529 сферичних гармонік або їх приведених розкладань, причому множина коефіцієнтів сферичних гармонік представляє звукове поле.

[1145] У кожному з різних вищеописаних випадків, варто розуміти, що пристрій 540 аудіодекодування може здійснювати спосіб або іншим чином містити засіб для здійснення кожного етапу способу, для якого пристрій 540 аудіодекодування виконаний з можливістю здійснення в ряді випадків, цей засіб може містити один або більше процесорів. У ряді випадків, один або більше процесорів можуть представляти процесор спеціального призначення, який виконує інструкції, що зберігаються на довгостроковому зчитуваному комп'ютером носії даних. Іншими словами, різні аспекти методів у кожному з наборів прикладів кодування можуть передбачати довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати спосіб, для здійснення якого призначений пристрій 540 аудіодекодування.

[1146] На фіг. 42В показана блок-схема, що демонструє блок 528В зниження порядку, який може представляти один приклад блока 528А зниження порядку, показаного на фіг. 40В-40J. У прикладі, наведеному на фіг. 42В, замість здійснення зниження порядку тільки на основі цільової бітової швидкості 535, блок 528В зниження порядку може здійснювати зниження порядку на основі аналізу контенту коефіцієнтів 531 сферичних гармонік фону. Блок 528В зниження порядку може включати в себе блок 536 аналізу контенту, що здійснює цей аналіз контенту.

[1147] У деяких прикладах, блок 536 аналізу контенту може включати в себе блок 536А просторового аналізу, що здійснює форму аналізу контенту, іменованого просторовим аналізом. Просторовий аналіз може передбачати аналіз коефіцієнтів 531 сферичних гармонік фону для ідентифікації просторової інформації, що описує форму або інші просторові властивості фонових компонентів звукового поля. На основі цієї просторової інформації, блок 528В зниження порядку може ідентифікувати ті порядки і/або підпорядки, що підлягають видаленню з коефіцієнтів 531 сферичних гармонік фону для генерації приведених коефіцієнтів 529 сферичних гармонік фону.

[1148] У деяких прикладах, блок 536 аналізу контенту може включати в себе блок 536В аналізу дифузії, який здійснює форму аналізу контенту, іменованого аналізом дифузії. Аналіз дифузії може передбачати аналіз коефіцієнтів 531 сферичних гармонік фону для ідентифікації інформації дифузії, що описує дифузійність фонових компонентів звукового поля. На основі цієї інформації дифузії, блок 528В зниження порядку може ідентифікувати ті порядки і/або підпорядки, що підлягають видаленню з коефіцієнтів 531 сферичних гармонік фону для генерації приведених коефіцієнтів 529 сферичних гармонік фону.

[1149] Хоча він показаний як такий, що включає в себе блок 536А просторового аналізу і блок 536В аналізу дифузії, блок 536 аналізу контенту може включати в себе тільки блок просторового аналізу 536А, тільки блок 536В аналізу дифузії або і блок 536А просторового аналізу, і блок 536В аналізу дифузії. У деяких прикладах, блок 536 аналізу контенту може здійснювати інші форми аналізу контенту крім або замість одного або обох із просторового аналізу й аналізу дифузії. Відповідно, методи, описані в цьому винаході, не підлягають обмеженню в цьому відношенні.

[1150] Таким чином, методи можуть передбачати, що пристрій аудіокодування, наприклад пристрої 510В-510J аудіокодування, здійснює, на основі аналізу контенту множини коефіцієнтів сферичних гармонік або їх розкладань, що описують звукове поле, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

[1151] Іншими словами, методи можуть передбачати, що пристрій, наприклад пристрої 510В-510J аудіокодування, сконфігуровані згідно з наступними пунктами.

[1152] Пункт 133146-1Е. Пристрій, наприклад кожний з пристроїв 510В-510J аудіокодування, який містить: один або більше процесорів, виконаних з можливістю здійснення, на основі аналізу контенту множини коефіцієнтів сферичних гармонік або їх розкладань, що описують звукове поле, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

[1153] Пункт 133146-2Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів додатково виконані з можливістю, до здійснення зниження порядку, здійснення розкладання на сингулярні значення відносно множини коефіцієнтів сферичних гармонік для ідентифікації одного або більше перших векторів, що описують характерні компоненти звукового поля, й одного або більше других векторів, що ідентифікують фонові компоненти звукового поля, і при цьому один або більше процесорів виконані з можливістю здійснення

зниження порядку відносно одного або більше перших векторів, одного або більше других векторів або одного або більше перших векторів і одного або більше других векторів.

5 [1154] Пункт 133146-3Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів додатково виконані з можливістю здійснення аналізу контенту відносно множини коефіцієнтів сферичних гармонік або їх розкладань.

[1155] Пункт 133146-4Е. Пристрій за пунктом 133146-3Е, у якому один або більше процесорів виконані з можливістю здійснення просторового аналізу відносно множини коефіцієнтів сферичних гармонік або їх розкладань.

10 [1156] Пункт 133146-5Е. Пристрій за пунктом 133146-3Е, у якому здійснення аналізу контенту включає здійснення аналізу дифузії відносно множини коефіцієнтів сферичних гармонік або їх розкладань.

[1157] Пункт 133146-6Е. Пристрій за пунктом 133146-3Е, у якому один або більше процесорів виконані з можливістю здійснення просторового аналізу й аналізу дифузії відносно множини коефіцієнтів сферичних гармонік або їх розкладань.

15 [1158] Пункт 133146-7Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів виконані з можливістю здійснення, на основі аналізу контенту з множини коефіцієнтів сферичних гармонік або їх розкладань і цільової бітової швидкості, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

20 [1159] Пункт 133146-8Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів додатково виконані з можливістю аудіокодування приведених коефіцієнтів сферичних гармонік або їх розкладань.

25 [1160] Пункт 133146-9Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів додатково виконані з можливістю аудіокодування приведених коефіцієнтів сферичних гармонік або їх приведених розкладань і генерації бітового потоку для включення приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

[1161] Пункт 133146-10Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів додатково виконані з можливістю указання одного або більше порядків і/або одного або більше підпорядків сферичних базисних функцій, яким відповідають ті з приведених коефіцієнтів сферичних гармонік або їх приведених розкладань у бітовому потоці, що включає в себе приведені коефіцієнти сферичних гармонік або їх приведених розкладань.

[1162] Пункт 133146-11Е. Пристрій за пунктом 133146-1Е, у якому приведені коефіцієнти сферичних гармонік або їх приведених розкладань мають менше значень, ніж множина коефіцієнтів сферичних гармонік або їх розкладань.

35 [1163] Пункт 133146-12Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів додатково виконані з можливістю видалення тих з множини коефіцієнтів сферичних гармонік або векторів їх розкладань, що мають зазначений порядок і/або підпорядок, для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

40 [1164] Пункт 133146-13Е. Пристрій за пунктом 133146-1Е, у якому один або більше процесорів виконані з можливістю обнулення тих з множини коефіцієнтів сферичних гармонік або тих векторів їх розкладань, що мають зазначений порядок і/або підпорядок, для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

45 [1165] На фіг. 42С показана блок-схема, що демонструє блок 528С зниження порядку, який може представляти один приклад блока 528А зниження порядку, показаного на фіг. 40В-40J. Блок 528С зниження порядку, показаний на фіг. 42С, по суті, ідентичний блоку 528В зниження порядку, але може приймати або інакше визначати цільову бітову швидкість 535 вищеописаним чином відносно блока 528А зниження порядку, показаного на фіг. 42А, і при цьому також здійснювати аналіз контенту вищеописаним чином відносно блока 528В зниження порядку, показаного на фіг. 42В. Потім блок 528С зниження порядку може здійснювати зниження порядку відносно коефіцієнтів 531 сферичних гармонік фону на основі цієї цільової бітової швидкості 535 і аналізу контенту.

50 [1166] Таким чином, методи можуть передбачати, що пристрій аудіокодування, наприклад пристрої 510В-510J аудіокодування, здійснює аналіз контенту відносно множини коефіцієнтів сферичних гармонік або їх розкладань. При здійсненні зниження порядку, пристрої 510В-510J аудіокодування можуть здійснювати, на основі цільової бітової швидкості 535 і аналізу контенту, зниження порядку відносно множини коефіцієнтів сферичних гармонік або їх розкладань для генерації приведених коефіцієнтів сферичних гармонік або їх приведених розкладань.

[1167] За умови, що один або більше векторів видаляються, пристрої 510В-510J аудіокодування можуть указувати кількість векторів у бітовому потоці як дані керування.

Пристрої 510B-510J аудіокодування можуть указувати цю кількість векторів у бітовому потоці для полегшення витягання векторів з бітового потоку пристроєм аудіодекодування.

[1168] На фіг. 44 показана схема, що демонструє ілюстративні операції, здійснювані пристроєм аудіокодування 410D для компенсації помилки квантування відповідно до різних аспектів методів, описаних у цьому винаході. У прикладі, наведеному на фіг. 44, математичний блок 526 пристрою 510D аудіокодування показаний пунктирним прямокутником для позначення, що математичні операції можуть здійснюватися математичним блоком 526 пристрою 510D аудіодекодування.

[1169] Як показано в прикладі, наведеному на фіг. 44, математичний блок 526 може спочатку множити вектори $527 U_{DIST} * S_{DIST}$ на вектори $525E V_{DIST}^T$ для генерації характерних коефіцієнтів сферичних гармонік (позначених як "вектори 630 H_{DIST} "). Потім математичний блок 526 може ділити вектори $630 H_{DIST}$ на квантовану версію векторів $525E V_{DIST}^T$ (які позначені, знову ж, як "вектори $525G V_{Q_DIST}^T$ "). Математичний блок 526 може здійснювати це ділення шляхом визначення псевдооберненого векторів $525G V_{Q_DIST}^T$ і потім множення векторів H_{DIST} на псевдообернене векторів $525G V_{Q_DIST}^T$ і виводити версію з компенсацією помилки $U_{DIST} * S_{DIST}$ (яка може скорочено іменуватися " US_{DIST} " або "вектори US_{DIST} "). Версія з компенсацією помилки US_{DIST} може бути позначена як вектори $527' US_{DIST}^*$ у прикладі, наведеному на фіг. 44. Таким чином, методи можуть ефективно проектувати помилку квантування, щонайменше частково, у вектори $527 US_{DIST}$, генеруючи вектори $527' US_{DIST}^*$.

[1170] Потім математичний блок 526 може віднімати вектори $527' US_{DIST}^*$ з векторів $527 U_{DIST} * S_{DIST}$ для визначення векторів $634 US_{ERR}$ (які можуть представляти щонайменше частину помилки внаслідок квантування, проєктованої у вектори $527 U_{DIST} * S_{DIST}$). Потім математичний блок 526 може множити вектори $634 US_{ERR}$ на вектори $525G V_{Q_DIST}^T$ для визначення векторів $636 H_{ERR}$. Математично, вектори $636 H_{ERR}$ можуть бути еквівалентні векторам $527 US_{DIST}$ - векторам $527' US_{DIST}^*$, результат чого потім множиться на вектори $525E V_{DIST}^T$. Потім математичний блок 526 може підсумовувати вектори $636 H_{ERR}$ з коефіцієнтами 531 сферичних гармонік фону (позначеними як вектори $531 H_{BG}$ у прикладі, наведеному на фіг. 44), обчисленими шляхом множення векторів $525D U_{BG}$ на вектори $525B S_{BG}$ і потім на вектори $525F V_{BG}^T$. Математичний блок 526 може додавати вектори $636 H_{ERR}$ до векторів $531 H_{BG}$, ефективно проєктуючи щонайменше частину помилки квантування у вектори $531 H_{BG}$ для генерації скомпенсованих векторів $531' H_{BG}$. Таким чином, методи можуть проєктувати щонайменше частину помилки квантування у вектори $531 H_{BG}$.

[1171] На фіг. 45A і 45B показані схеми, що демонструють інтерполяцію підкадрів з ділянок двох кадрів відповідно до різних аспектів методів, описаних у цьому винаході. У прикладі, наведеному на фіг. 45A, показані перший кадр 650 і другий кадр 652. Перший кадр 650 може включати в себе коефіцієнти сферичних гармонік (" $SH[1]$ "), які можна розкласти на матриці $U[1]$, $S[1]$ і $V[1]$. Другий кадр 652 може включати в себе коефіцієнти сферичних гармонік (" $SH[2]$ "). Ці $SH[1]$ і $SH[2]$ можуть ідентифікувати різні кадри SHC 511, описаного вище.

[1172] У прикладі, наведеному на фіг. 45B, блок 518 розкладання пристрою 510H аудіокодування, показаного в прикладі, наведеному на фіг. 40H, може розділяти кожний з кадрів 650 і 652 на чотири відповідних підкадри $651A-651D$ і $653A-653D$. Потім блок 518 розкладання може розкласти перший підкадр $651A$ (позначений як " $SH[1,1]$ ") кадру 650 на матриці $U[1,1]$, $S[1,1]$ і $V[1,1]$, виводячи матрицю $519A' V[1,1]$ на блок 550 інтерполяції. Потім блок 518 розкладання може розкласти другий підкадр $653A$ (позначений як " $SH[2,1]$ ") кадру 652 на матриці $U[1,1]$, $S[1,1]$ і $V[1,1]$, виводячи матрицю $519A' V[2,1]$ на блок 550 інтерполяції. Блок 518 розкладання також може виводити $SH[1,1]$, $SH[1,2]$, $SH[1,3]$ і $SH[1,4]$ SHC 11 і $SH[2,1]$, $SH[2,2]$, $SH[2,3]$ і $SH[2,4]$ SHC 511 на блок 550 інтерполяції.

[1173] Потім блок 550 інтерполяції може здійснювати інтерполяції, ідентифіковані внизу ілюстрації, показаної в прикладі, наведеному на фіг. 45B. Таким чином, блок 550 інтерполяції може інтерполювати $V'[1,2]$ на основі $V'[1,1]$ і $V'[2,1]$. Блок 550 інтерполяції також може інтерполювати $V'[1,3]$ на основі $V'[1,1]$ і $V'[2,1]$. Крім того, блок 550 інтерполяції також може інтерполювати $V'[1,4]$ на основі $V'[1,1]$ і $V'[2,1]$. Ці інтерполяції можуть передбачати проєкцію $V'[1,1]$ і $V'[2,1]$ у просторову область, як показано в прикладі, наведеному на фіг. 46A-46E, після якої іде часова інтерполяція і потім проєкція зворотно в область сферичних гармонік.

[1174] Потім блок 550 інтерполяції може виводити $U[1,2]S[1,2]$ шляхом множення $SH[1,2]$ на $(V'[1,2])^{-1}$, $U[1,3]S[1,3]$ шляхом множення $SH[1,3]$ на $(V'[1,3])^{-1}$ і $U[1,4]S[1,4]$ шляхом множення $SH[1,4]$ на $(V'[1,4])^{-1}$. Потім блок 550 інтерполяції може перетворювати кадр в розкладену форму, виводячи матрицю $519A V$, матрицю $519B S$ і матрицю $519C U$.

[1175] На фіг. 46A-46E показані схеми, що демонструють поперечний переріз проєкції одного або більше векторів розкладеної версії множини коефіцієнтів сферичних гармонік,

інтерпольованих відповідно до методів, описаних в цьому винаході. Фіг. 46А демонструє поперечний переріз проекції одного або більше перших векторів першої матриці 519А' V, що розкладається з SHC 511 першого підкадру з першого кадру за допомогою процесу SVD. Фіг. 46В демонструє поперечний переріз проекції одного або більше других векторів другої матриці 519А' V, що розкладається з SHC 511 першого підкадру з другого кадру за допомогою процесу SVD.

[1176] Фіг. 46С демонструє поперечний переріз проекції одного або більше інтерпольованих векторів для матриці 519А' V, що представляє другий підкадр із першого кадру, причому ці вектори інтерпольовані відповідно до методів, описаних в цьому винаході, з матриці 519А' V, розкладеної з першого підкадру першого кадру SHC 511 (тобто один або більше векторів матриці 519А' V, показаної в прикладі, наведеному на фіг. 46 у цьому прикладі) і першого підкадру другого кадру SHC 511 (тобто один або більше векторів матриці 519А' V, показаної в прикладі, наведеному на фіг. 46В у цьому прикладі).

[1177] Фіг. 46D демонструє поперечний переріз проекції одного або більше інтерпольованих векторів для матриці 519А' V, що представляє третій підкадр із першого кадру, причому ці вектори інтерпольовані відповідно до методів, описаних в цьому винаході, з матриці 519А' V, розкладеної з першого підкадру першого кадру SHC 511 (тобто один або більше векторів матриці 519А' V, показаної в прикладі, наведеному на фіг. 46 у цьому прикладі) і першого підкадру другого кадру SHC 511 (тобто один або більше векторів матриці 519А' V, показаної в прикладі, наведеному на фіг. 46В у цьому прикладі).

[1178] Фіг. 46Е демонструє поперечний переріз проекції одного або більше інтерпольованих векторів для матриці 519А' V, що представляє четвертий підкадр із першого кадру, причому ці вектори інтерпольовані відповідно до методів, описаних в цьому винаході, з матриці 519А' V, розкладеної з першого підкадру першого кадру SHC 511 (тобто один або більше векторів матриці 519А' V, показаної в прикладі, наведеному на фіг. 46 у цьому прикладі) і першого підкадру другого кадру SHC 511 (тобто один або більше векторів матриці 519А' V, показаної в прикладі, наведеному на фіг. 46В у цьому прикладі).

[1179] На фіг. 47 показана блок-схема, що демонструє, більш детально, блок 542 витягання пристроїв 540А-540D аудіодекодування, показаних у прикладах, наведених на фіг. 41А-41D. У деяких прикладах, блок 542 витягання може представляти вхідний каскад так названого "вбудованого декодера", що може здійснювати дві або більше схем декодування (де шляхом здійснення цих двох або більше схем декодер можна розглядати як "об'єднуючий" дві або більше схем). Як показано в прикладі, наведеному на фіг. 44, блок 542 витягання включає в себе мультиплексор 620 і підблоки 622А і 622В витягання ("підблоки 622 витягання"). Мультиплексор 620 ідентифікує матриці 547А-547N кодованих кадрованих SHC, що підлягають відправленню на підблок 622А витягання і підблок 622В витягання на основі відповідного вказання, чи генеруються відповідні матриці 547А-547N кодованих кадрованих SHC із синтетичного аудіооб'єкта або запису. Кожний з підблоків 622А витягання може здійснювати окрему схему декодування (яке може іменуватися "зняттям стиснення"), тобто, у деяких прикладах, адаптовану або до SHC, генерованого із синтетичного аудіооб'єкта, або до SHC, генерованого з запису. Кожний з підблоків 622А витягання може здійснювати відповідну одну з цих схем зняття стиснення для генерації кадрів SHC 547, що виводяться в SHC 547.

[1180] Наприклад, блок 622А витягання може здійснювати схему зняття стиснення для реконструкції SA з переважного сигналу (PS) з використанням наступної формули:

$$HOA = \text{DirV} \times \text{PS},$$

де DirV - направлений вектор (що представляє різні напрямки і значення ширини), який може передаватися через побічний канал. Блок 622В витягання може, у цьому прикладі, здійснювати схему зняття стиснення, що реконструює матрицю HOA з PS з використанням наступної формули:

$$HOA = \sqrt{4\pi} \cdot Y_{nm}(\text{тета}, \text{фі}) \cdot \text{PS},$$

де Y_{nm} - сферична гармонічна функція, і інформація тета і фі може відправлятися через побічний канал.

[1181] У цьому відношенні, методи передбачають, що блок 538 витягання вибирає одну з множини схем зняття стиснення на основі вказання, чи генеруються стиснена версія коефіцієнтів сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, і знімає стиснення стисненої версії коефіцієнтів сферичних гармонік з використанням вибраної однієї з множини схем зняття стиснення. У деяких прикладах, пристрій містить вбудований декодер.

[1182] На фіг. 48 показана блок-схема, що більш детально демонструє блок 48 аудіорендеризації пристрою 540А-540D аудіодекодування, показаного в прикладах, наведених

на фіг. 41A-41D. Фіг. 48 демонструє перетворення з відновлених коефіцієнтів 547 сферичних гармонік у багатоканальні аудіодані 549A, сумісні з декодерно-локальною геометрією гучномовців. Для деяких локальних геометрій гучномовців (які, знову ж, можуть означати геометрію гучномовців на декодері), деякі перетворення, що гарантують оборотність, можуть

приводити до небажано низької якості аудіозображення. Таким чином, звуковідтворення не завжди приводить до правильної локалізації звуків у порівнянні з захоплюваним аудіосигналом. Для корекції цієї небажано низької якості зображення, методи можна додатково доповнювати для введення поняття, яке може іменуватися "віртуальними гучномовцями".

[1183] Замість вимоги, щоб один або більше гучномовців переміщалися або розташовувалися в конкретних або заданих областях простору, що мають визначені кутові допуски, установлені стандартом, наприклад вищезгаданим ITU-R BS.775-1, вищеописану конфігурацію можна модифікувати для включення деякої форми панорамування, наприклад амплітудного панорамування на векторній основі (VBAP), амплітудного панорамування на основі відстані або інших форм панорамування. Розглядаючи VBAP з метою ілюстрації, VBAP може ефективно вводити так називані "віртуальні гучномовці". VBAP може змінювати підведення до одного або більше гучномовців, завдяки чому ці один або більше гучномовців ефективно виводять звук, який сприймається як такий, що виходить від віртуального гучномовця, одне або більше з положення і кута якого відрізняється від щонайменше одного з положення і/або кута одного або більше гучномовців, які підтримують віртуальний гучномовець.

[1184] Для ілюстрації, наступне рівняння для визначення підведень гучномовця відносно SHC може виглядати наступним чином:

$$\begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{-(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX \\ M \times N \end{bmatrix} \begin{bmatrix} D \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix}.$$

[1185] У вищенаведеному рівнянні, матриця VBAP має розмір M рядків на N стовпців, де M позначає кількість гучномовців (яка дорівнює п'яти у вищенаведеному рівнянні) і N позначає кількість віртуальних гучномовців. Матрицю VBAP можна обчислювати як функцію векторів із заданого положення слухача до кожної з позицій гучномовців і векторів із заданого положення слухача до кожної з позицій віртуальних гучномовців. Матриця D у вищенаведеному рівнянні може мати розмір N рядків на (порядок+1)² стовпців, де порядок може означати порядок функцій SH. Матриця D може представляти наступну матрицю:

$$\begin{bmatrix} h_0^{(2)}(kr_1)Y_0^{0*}(\theta_1, \varphi_1) & h_0^{(2)}(kr_2)Y_0^{0*}(\theta_2, \varphi_2) & \dots & \dots \\ h_1^{(2)}(kr_1)Y_1^{1*}(\theta_1, \varphi_1) & \dots & \dots & \dots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

[1186] Матриця g (або вектор, за умови, що існує один єдиний стовпець) може представляти коефіцієнт підсилення для підведень гучномовця для гучномовців, розміщених у декодерно-локальній геометрії. У рівнянні, матриця g має розмір M. Матриця A (або вектор, за умови, що існує один єдиний стовпець) може позначати SHC 520 і має розмір (порядок+1)(порядок+1), що також може бути позначений як (порядок+1)².

[1187] Фактично, матриця VBAP є матрицею M×N, що забезпечує так назване "регулювання коефіцієнта підсилення", яка розкладається на положення гучномовців і позицію віртуальних гучномовців. Введення панорамування, таким чином, може приводити до поліпшення відтворення багатоканального аудіосигналу, що приводить до підвищення якості зображення при відтворенні локальною геометрією гучномовців. Крім того, завдяки включенню VBAP у це рівняння, методи можуть долати недоліки геометрій гучномовців, які не вирівнюються з зазначеними в різних стандартах.

[1188] На практиці, рівняння можна інвертувати і застосовувати для перетворення SHC зворотно у багатоканальні підведення для конкретної геометрії або конфігурації гучномовців, яка, знову ж, може іменуватися декодерно-локальною геометрією в цьому винаході. Таким

чином, рівняння можна інвертувати для рішення відносно матриці g . Інвертоване рівняння може виглядати наступним чином:

$$\begin{bmatrix} g_1(\omega) \\ g_2(\omega) \\ g_3(\omega) \\ \dots \\ g_M(\omega) \end{bmatrix} = -ik \begin{bmatrix} VBAP \\ MATRIX^{-1} \\ M \times N \end{bmatrix} \begin{bmatrix} D^{-1} \\ Nx(Order+1)^2 \end{bmatrix} \begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ \dots \\ A_{(Order+1)(Order+1)}^{-(Order+1)(Order+1)}(\omega) \end{bmatrix}.$$

5

[1189] Матриця g може представляти коефіцієнт підсилення гучномовця для, у цьому прикладі, кожного з п'яти гучномовців у конфігурації гучномовців 5.1. Положення віртуальних гучномовців, використовуваних у цій конфігурації, може відповідати положенням, заданим у специфікації або стандарті багатоканального формату 5.1. Положення гучномовців, яке може підтримувати кожний з цих віртуальних гучномовців, можна визначити з використанням будь-якої кількості відомих методів локалізації аудіосигналу, багато які з яких передбачають відтворення тону, що має конкретну частоту, для визначення положення кожного гучномовця відносно основного блока (наприклад, приймача аудіо/відеосигналу (A/V приймача), телевізора, ігрової системи, система цифрового відеодиска або інших типів основних систем).

10 Альтернативно, користувач основного блока може вручну вказувати положення кожного з гучномовців. У будь-якому випадку, на основі цих відомих положень і можливих кутів, основний блок може знаходити коефіцієнти підсилення, передбачаючи ідеальну конфігурацію віртуальних гучномовців за допомогою VBAP.

[1190] У цьому відношенні, пристрій або устаткування може здійснювати амплітудне панорамування на векторній основі або іншу форму панорамування на множині віртуальних каналів для створення множини каналів, що стимулюють гучномовці в декодерно-локальній геометрії випромінювати звуки, які сприймаються як такі, що виходять з віртуальних гучномовців, сконфігурованих в іншій локальній геометрії. Методи, таким чином, можуть передбачати, що пристрій 40 аудіодекодування здійснює перетворення на множині коефіцієнтів сферичних гармонік, наприклад відновлених коефіцієнтів 47 сферичних гармонік, для створення множини каналів. Кожний з множини каналів може бути пов'язаний з відповідною окремою областю простору. Крім того, кожний з множини каналів може містити множину віртуальних каналів, де множина віртуальних каналів може бути пов'язана з відповідною окремою областю простору. Таким чином, пристрій може здійснювати амплітудне панорамування на векторній основі на віртуальних каналах для створення множини каналів багатоканальних аудіоданих 49.

[1191] На фіг. 49A-49E(ii) показані схеми, що демонструють відповідні системи 560A-560C, 567D, 569D, 571E і 573E аудіокодування, які можуть реалізувати різні аспекти методів, описаних у цьому винаході. Як показано в прикладі, наведеному на фіг. 49A, система 560A аудіокодування може включати в себе пристрій 562 аудіокодування і пристрій 564 аудіодекодування. Пристрій 562 аудіокодування може бути аналогічним будь-якому з пристроїв 20 і 510A-510D аудіокодування, показаним у прикладі, наведеному на фіг. 4 і 40A-40D, відповідно. Пристрій 564 аудіодекодування може бути аналогічним пристрою 24 і 540 аудіодекодування, показаному в прикладі, наведеному на фіг. 5 і 41.

[1192] Як описано вище, амбіофонія більш високого порядку (НОА) дозволяє описувати всю інформацію направленості звукового поля на основі просторового перетворення Фур'є. У деяких прикладах, чим вище порядок амбіофонії, N , тим вище просторове розрізнення і тим більше кількість коефіцієнтів сферичних гармонік (SH) $(N+1)^2$. Таким чином, більш високий порядок амбіофонії N , у деяких прикладах, приводить до підвищених вимог до ширини смуги для передачі і збереження коефіцієнтів. Оскільки вимоги до ширини смуги НОА досить високі, в порівнянні, наприклад, з аудіоданими об'ємного звуку 5.1 або 7.1, зменшення ширини смуги може бути бажаним для багатьох застосувань.

[1193] Відповідно до методів, описаних в цьому винаході, система 560A аудіокодування може здійснювати спосіб на основі відділення характерних елементів (переднього плану) від нехарактерних елементів (фону або оточення) у просторовій звуковій сцені. Цей поділ дозволяє системі 560A аудіокодування обробляти елементи переднього плану і фону незалежно один від одного. У цьому прикладі, система 560A аудіокодування використовує ту особливість, що елементи переднього плану можуть більше привертати увагу (слухача) і легше локалізуються (знову ж слухачем) у порівнянні з фоновими елементами. У результаті, система 560A аудіокодування може більш ефективно зберігати або передавати контент НОА.

[1194] У деяких прикладах, система 560A аудіокодування може досягати цього поділу шляхом застосування процесу розкладання на сингулярні значення (SVD). Процес SVD може розділяти кадр коефіцієнтів HOA на 3 матриці (U, S, V). Матриця U містить ліві сингулярні вектори і матриця V містить праві сингулярні вектори. Діагональна матриця S містить неенегативні, сортовані сингулярні значення на своїй діагоналі. У загальному випадку гарна (або, у ряді випадків, відмінна, передбачаючи необмежену точність представлення коефіцієнтів HOA) реконструкція коефіцієнтів HOA буде задана за допомогою $U \cdot S \cdot V^T$. Реконструюючи тільки підпростір з D найбільшими сингулярними значеннями: $U(:, 1:D) \cdot S(1:D, :) \cdot V^T$, система 560A аудіокодування може витягати найбільш помітну просторову інформацію з цього кадру HOA, тобто звукові елементи переднього плану (і, можливо, деякі сильні ранні кімнатні відображення). Залишок $U(:, D+1:end) \cdot S(D+1:end, :) \cdot V^T$ може реконструювати фонові елементи і реверберацію з контенту.

[1195] Система 560A аудіокодування може визначати значення D, що розділяє два підпростори, шляхом аналізу нахилу кривої, створеної зменшуваними діагональними значеннями S: великі сингулярні значення представляють звуки переднього плану, низькі сингулярні значення представляють фонові значення. Система 560A аудіокодування може використовувати першу і другу похідні кривої сингулярних значень. Система 560A аудіокодування також може обмежувати число D величиною між одиницею і п'ятьма. Альтернативно, система 560A аудіокодування може попередньо задавати число D, наприклад значення чотири. У будь-якому випадку, оцінивши число D, система 560A аудіокодування витягає підпростори переднього плану і фону з матриць U і S.

[1196] Система 560A аудіокодування може потім реконструювати коефіцієнти HOA фонові сцени за допомогою $U(:, D+1:end) \cdot S(D+1:end, :) \cdot V^T$, що дає $(N+1)^2$ каналів коефіцієнтів HOA. Оскільки відомо, що фонові елементи, у деяких прикладах, не настільки помітні і не настільки локалізовані як елементи переднього плану, система 560A аудіокодування може зрізати порядок каналів HOA. Крім того, система 560A аудіокодування може стискати ці канали за допомогою втратних або безвтратних аудіокодеків, наприклад AAC, або, у необов'язковому порядку, за допомогою більш агресивного аудіокодека в порівнянні з використанням для стиснення помітних елементів переднього плану. У ряді випадків, для економії ширини смуги, система 560A аудіокодування може по-різному передавати елементи переднього плану. Таким чином, система аудіокодування може передавати ліві сингулярні вектори $U(:, 1:D)$, стиснувши їх за допомогою втратних або безвтратних аудіокодеків (наприклад AAC), і передавати ці стиснені ліві сингулярні значення разом з матрицею реконструкції $R = S(1:D, :) \cdot V^T$. R може представляти матрицю $D \times (N+1)^2$, що може розрізнятися по кадрах.

[1197] На стороні приймача системи 560 аудіокодування, система аудіокодування може множити ці дві матриці для реконструкції кадру $(N+1)^2$ каналів HOA. Після підсумовування одного з одним каналів HOA фону і переднього плану, система 560A аудіокодування може рендеризуватися на будь-яку акустичну установку з використанням будь-якого належного рендеризатора амбіофонії. Оскільки методи забезпечують відділення елементів переднього плану (прямого або характерного звуку) від фонових елементів, слабкочуючі можуть керувати мікшуванням елементів переднього плану і фону для збільшення розбірливості. Застосовні також інші аудіоефекти, наприклад динамічний компресор тільки на елементах переднього плану.

[1198] На фіг. 49B показана блок-схема, що більш детально демонструє систему 560B аудіокодування. Як показано в прикладі, наведеному на фіг. 49B, система 560B аудіокодування може включати в себе пристрій 566 аудіокодування і пристрій 568 аудіокодування. Пристрій 566 аудіокодування може бути аналогічним пристроєм 24 і 510E аудіокодування, показаним у прикладі, наведеному на фіг. 4 і 40E. Пристрій 568 аудіокодування може бути аналогічним пристроєм 24 і 540B аудіокодування, показаним в прикладі, наведеному на фіг. 5 і 41B.

[1199] Відповідно до методів, описаних в цьому винаході, при використанні SVD на кадровій основі (або споріднених способів, наприклад KLT і PCA) розкладання на сигналах HOA, з метою зменшення ширини смуги, пристрій 66 аудіокодування може квантувати перші декілька векторів матриці U (помножені на відповідні сингулярні значення матриці S), а також відповідні вектори вектора V^T . Вони будуть містити 'фонові' компоненти звукового поля. Методи можуть передбачати, що пристрій 566 аудіокодування кодує вектор $U_{DIST} \cdot S_{DIST}$ з використанням машини аудіокодування типу "чорний ящик". Вектор V може піддаватися скалярному або векторному квантуванню. Крім того, деякі або усі з векторів, що залишилися, у матриці U можна множити на відповідні сингулярні значення матриці S і матриці V і також кодувати з використанням машини аудіокодування типу "чорний ящик". Вони будуть містити 'фонові' компоненти звукового поля.

[1200] Оскільки найгучніші звукові компоненти розкладаються на 'компоненти переднього плану', пристрій 566 аудіокодування може знижувати порядок амбіофонії 'фонових' компонентів до використання машини аудіокодування типу "чорний ящик", оскільки (передбачається, що) фон не містить важливого локалізованого контенту. Залежно від порядку амбіофонії компонентів переднього плану, блок 566 аудіокодування може передавати відповідний(i) V-вектор(и), який може бути достатньо великий. Наприклад, просте 16-бітове скалярне квантування векторів V буде приводити до службового навантаження приблизно 20 кбіт/с для 4-го порядку (25 коефіцієнтів) і 40 кбіт/с для 6-го порядку (49 коефіцієнтів) для кожного компонента переднього плану. Методи, описані в цьому винаході, можуть забезпечувати спосіб зниження цього службового навантаження V-вектора.

[1201] Для ілюстрації, передбачимо, що порядок амбіофонії елементів переднього плану дорівнює N_{DIST} і порядок амбіофонії фонових елементів дорівнює N_{BG} , як описано вище. Оскільки пристрій 566 аудіокодування може знижувати порядок амбіофонії фонових елементів, як описано вище, N_{BG} може бути менше N_{DIST} . Довжина V-вектора переднього плану, який потрібно передавати для реконструкції елементів переднього плану на стороні приймача, має довжину $(N_{DIST}+1)^2$ для кожного елемента переднього плану, тоді як перші $((N_{DIST}+1)^2) - ((N_{BG}+1)^2)$ коефіцієнтів можна використовувати для реконструкції компонентів переднього плану або характерних компонентів аж до порядку N_{BG} . З використанням методів, описаних у цьому винаході, пристрій 566 аудіокодування може реконструювати передній план аж до порядку N_{BG} і об'єднувати результуючі $(N_{BG}+1)^2$ каналів з каналами фону, з утворенням повного звукового поля аж до порядку N_{BG} . Потім пристрій 566 аудіокодування може скорочувати V-вектор до коефіцієнтів з індексом вище $(N_{BG}+1)^2$ для передачі (і ці вектори можуть іменуватися " V_{SMALL}^T "). На стороні приймача, блок 568 аудіодекодування може реконструювати аудіоканали переднього плану для порядку амбіофонії, більшого N_{BG} , шляхом множення елементів переднього плану на вектори V_{SMALL}^T .

[1202] На фіг. 49C показана блок-схема, що більш детально демонструє систему 560C аудіокодування. Як показано в прикладі, наведеному на фіг. 49C, система 560C аудіокодування може включати в себе пристрій 567 аудіокодування і пристрій 569 аудіодекодування. Пристрій 567 аудіокодування може бути аналогічним пристроєм 20 і 510F аудіокодування, показаним у прикладі, наведеному на фіг. 4 і 40F. Пристрій 569 аудіодекодування може бути аналогічним пристроєм 24 і 540B аудіодекодування, показаним у прикладі, наведеному на фіг. 5 і 41B.

[1203] Відповідно до методів, описаних в цьому винаході, при використанні SVD на кадровій основі (або споріднених способів, наприклад KLT і PCA) розкладання на сигналах НОА, з метою зменшення ширини смуги, пристрій 567 аудіокодування може квантувати перші декілька векторів матриці U (помножені на відповідні сингулярні значення матриці S), а також відповідні вектори вектора V^T . Вони будуть містити 'фонові' компоненти звукового поля. Методи можуть передбачати, що пристрій 567 аудіокодування кодує вектор $U_{DIST} * S_{DIST}$ з використанням машини аудіокодування типу "чорний ящик". Вектор V може піддаватися скалярному або векторному квантуванню. Крім того, деякі або усі з векторів, що залишилися, у матриці U можна множити на відповідні сингулярні значення матриці S і матриці V і також кодувати з використанням машини аудіокодування типу "чорний ящик". Вони будуть містити 'фонові' компоненти звукового поля.

[1204] Оскільки найгучніші звукові компоненти розкладаються на 'компоненти переднього плану', пристрій 567 аудіокодування може знижувати порядок амбіофонії 'фонових' компонентів до використання машини аудіокодування типу "чорний ящик", оскільки (передбачається, що) фон не містить важливого локалізованого контенту. Пристрій 567 аудіокодування може знижувати порядок таким чином, щоб зберігати повну енергію звукового поля відповідно до описаних тут методів. Залежно від порядку амбіофонії компонентів переднього плану, блок 567 аудіокодування може передавати відповідний(i) V-вектор(и), який може бути достатньо великим. Наприклад, просте 16-бітове скалярне квантування векторів V буде приводити до службового навантаження приблизно 20 кбіт/с для 4-го порядку (25 коефіцієнтів) і 40 кбіт/с для 6-го порядку (49 коефіцієнтів) для кожного компонента переднього плану. Методи, описані в цьому винаході, можуть забезпечувати спосіб зниження цього службового навантаження V-вектора(ів).

[1205] Для ілюстрації, передбачимо, що порядок амбіофонії елементів переднього плану дорівнює i і порядок амбіофонії фонових елементів дорівнює N. Пристрій 567 аудіокодування може знижувати порядок амбіофонії фонових елементів V-вектора(ів) від N до? таким чином, що? $<N$. Пристрій 567 аудіокодування додатково застосовує компенсацію для збільшення значень фонових елементів V-вектора(ів) для збереження повної енергії звукового поля, описаного за допомогою SHC. Ілюстративні методи застосування компенсації описані вище з посиланням на фіг. 40F. На стороні приймача, блок 569 аудіодекодування може реконструювати фонові аудіоканали для порядку амбіофонії.

[1206] Фіг. 49D(i) і 49D(ii) ілюструють пристрій 567D аудіокодування і пристрій 569D аудіодекодування, відповідно. Пристрій 567D аудіокодування і пристрій 569D аудіодекодування можуть бути виконані з можливістю здійснення одного або більше визначень характерності на основі направленості, відповідно до аспектів цього винаходу. Амбіофонія більш високого

порядку (HOA) є способом опису всієї інформації направленості звукового поля на основі просторового перетворення Фур'є. Чим вище порядок амбіофонії N , тим вище просторове розрізнення, тим більше кількість коефіцієнтів сферичних гармонік (SH) $(N+1)^2$, тим більше необхідна ширина смуги для передачі і збереження даних. Оскільки вимоги до ширини смуги HOA достатньо високі, для багатьох застосувань бажане зменшення ширини смуги.

[1207] Вище були наведені описи, як SVD (розкладання на сингулярні значення) або споріднені процеси можна використовувати для просторового стиснення аудіосигналу. Описані тут методи представляють поліпшений алгоритм для вибору помітних елементів, тобто елементів переднього плану. Після розкладання на основі SVD кадру HOA аудіо на його матриці U , S і V , методи основують вибір K помітних елементів винятково на перших K каналах матриці U [$U(:,1:K)*S(1:K, 1:K)$]. Це приводить до вибору аудіоелементів з найвищою енергією. Однак не гарантується, що ці елементи також є направленими. Таким чином, методи стосуються знаходження звукових елементів, що мають високу енергію і також є направленими. Це потенційно досягається шляхом зважування матриці V матрицею S . Потім, для кожного рядка цієї результуючої матриці, елементи з більш високими індексами (які пов'язані з коефіцієнтами більш високого порядку HOA) підносяться до квадрата і підсумовуються, що дає одне значення для кожного рядка [$\text{sum}(VS^2)$ у псевдокоді, описаному з посиланням на фіг. 40H]. Відповідно до послідовності операцій, представленої в псевдокоді, розглядаються коефіцієнти амбіофонії більш високого порядку, починаючи з 5-го індексу. Ці значення сортуються відповідно до їх розміру і індекс сортування використовується для переупорядкування вихідних матриць U , S і V , відповідно. Алгоритм стиснення на основі SVD, описаний вище в цьому винаході, потім можна застосовувати без додаткової зміни.

[1208] На фіг. 49E(i) і 49E(ii) показана блок-схема, що демонструє пристрій 571E аудіокодування і пристрій 573E аудіодекодування, відповідно. Пристрій 571E аудіокодування і пристрій 573E аудіодекодування можуть здійснювати різні аспекти методів, описані вище відносно прикладів, наведених на фіг. 49A-49D(ii), за винятком того, що пристрій 571E аудіокодування може здійснювати розкладання на сингулярні значення відносно матриці спектральної щільності потужності (PDS) коефіцієнтів HOA для генерації матриці S^2 і матриці V . Матриця S^2 може позначати квадрат матриці S , після чого матриця S^2 може піддаватися операції витягання квадратного кореня для одержання матриці S . Пристрій 571E аудіокодування може, у ряді випадків, здійснювати квантування відносно матриці V для одержання квантованої матриці V (яка може бути позначена як матриця V').

[1209] Пристрій 571E аудіокодування може одержувати матрицю U , спочатку множачи матрицю S на квантовану матрицю V' , для генерації матриці SV' . Пристрій 571E аудіокодування може потім одержувати псевдообернену матрицю SV' і потім множити коефіцієнти HOA на псевдообернену матрицю SV' для одержання матриці U . Шляхом здійснення SVD відносно спектральної щільності потужності коефіцієнтів HOA замість самих коефіцієнтів, пристрій 571E аудіокодування може потенційно знижувати обчислювальну складність здійснення SVD відносно одного або більше з циклів процесора і простору збереження, досягаючи при цьому такої ж ефективності аудіокодування джерела, як якби SVD застосовувалося безпосередньо до коефіцієнтів HOA.

[1210] Пристрій 573E аудіодекодування може бути аналогічним пристроям аудіодекодування, описаним вище, за винятком того, що пристрій 573E аудіодекодування може реконструювати коефіцієнти HOA з розкладань коефіцієнтів HOA, що досягаються шляхом застосування SVD до спектральної щільності потужності коефіцієнтів HOA, а не безпосередньо до коефіцієнтів HOA.

[1211] На фіг. 50A і 50B показані блок-схеми, що демонструють один із двох різних підходів для потенційного зниження порядку фонового контенту відповідно до методів, описаних в цьому винаході. Як показано в прикладі, наведеному на фіг. 50, перший підхід може використовувати зниження порядку відносно векторів $U_{BG} * S_{BG} * V^T$ для зниження порядку від N до \hat{n} , де \hat{n} менше, ніж ($<$) N . Таким чином, блок 528A зниження порядку, показаний у прикладах, наведених на фіг. 40B-40J, може здійснювати зниження порядку для зрізання або іншого зниження порядку N векторів $U_{BG} * S_{BG} * V^T$ до \hat{n} , де \hat{n} менше, ніж ($<$) N .

[1212] Як альтернативний підхід, блок 528A зниження порядку може, як показано в прикладі, наведеному на фіг. 50B, здійснювати це зрізання відносно V^T , усуваючи рядки до кількості $(\hat{n}+1)^2$, що не проілюстровано в прикладі, наведеному на фіг. 40B, для спрощення ілюстрації.

Іншими словами, блок 528A зниження порядку може видаляти один або більше рядків матриці V^T для ефективної генерації матриці V_{BG} . Розмір цієї матриці V_{BG} дорівнює $(\hat{n}+1)^2 \times (N+1)^2 - D$, де ця матриця V_{BG} потім використовується замість матриці V^T при генерації векторів $U_{BG} * S_{BG} * V^T$, ефективно здійснюючи зрізання для генерації векторів $U_{BG} * S_{BG} * V^T$ розміром $M \times (\hat{n}+1)^2$.

[1213] На фіг. 51 показана блок-схема, що демонструє приклади тракту стиснення характерних компонентів пристрою 700A аудіокодування, який може реалізувати різні аспекти методів, описаних у цьому винаході, для стиснення коефіцієнтів 701 сферичних гармонік. У прикладі, наведеному на фіг. 51, тракт стиснення характерних компонентів може означати тракт обробки пристрою 700A аудіокодування, який стискає характерні компоненти звукового поля, представленого SHC 701. Інший тракт, що може іменуватися трактом стиснення фонового компонента, може представляти тракт обробки пристрою 700A аудіокодування, що стискає фонові компоненти SHC 701.

[1214] Хоча це не показано для спрощення ілюстрації, тракт стиснення фонового компонента може діяти відносно безпосередньо SHC 701 замість розкладань SHC 701. Це аналогічно описаному вище з посиланням на фіг. 49A-49C, за винятком того, що замість повторного збирання фонових компонентів з матриць U_{BG} , S_{BG} і V_{BG} і потім здійснення деякої форми психоакустичного кодування (наприклад, з використанням кодера AAC) цих реконструйованих фонових компонентів, тракт обробки фонового компонента може діяти безпосередньо відносно SHC 701 (як описано вище відносно пристрою 20 аудіокодування, показаного в прикладі, наведеному на фіг. 4), стискаючи ці фонові компоненти з використанням психоакустичного кодера. Завдяки здійсненню психоакустичного кодування безпосередньо відносно SHC 701, порушення неперервності можуть знижуватися, поряд зі зниженням обчислювальної складності (відносно операцій, необхідних для стиснення фонових компонентів) у порівнянні зі здійсненням психоакустичного кодування відносно реконструйованих фонових компонентів. Хоча використовуються терміни "характерний" і "фон", термін "помітний" можна використовувати замість "характерний" і термін "оточення" можна використовувати замість "фон" у цьому винаході.

[1215] У будь-якому випадку, коефіцієнти 701 сферичних гармонік ("SHC 701") можуть містити матрицю коефіцієнтів, що має розмір $M \times (N+1)^2$, де M позначає кількість вибірок (i , у деяких прикладах, дорівнює 1024) у кадрі аудіо і N позначає найвищий порядок базисної функції, якій відповідають коефіцієнти. Як згадано вище, N звичайно задається рівним чотирьом (4) для всього 1024×25 коефіцієнтів. Кожний з SHC 701, відповідних конкретному порядку, комбінації підпорядків, може іменуватися каналом. Наприклад, усі M коефіцієнтів вибірки, відповідних базисній функції першого порядку, нульового підпорядку, можуть представляти канал, тоді як коефіцієнти, відповідні базисній функції нульового порядку, нульового підпорядку, можуть представляти інший канал, і т. д. SHC 701 також можуть іменуватися в цьому винаході контентом 701 амбіофонії більш високого порядку (HOA) або сигналом 701 SH.

[1216] Як показано в прикладі, наведеному на фіг. 51, пристрій 700A аудіокодування включає в себе блок 702 аналізу, блок 704 синтезу на векторній основі, блок 706 приведення векторів, блок 708 психоакустичного кодування, блок 710 приведення коефіцієнтів і блок 712 стиснення ("блок 712 стиснення"). Блок 702 аналізу може представляти блок, виконаний з можливістю здійснення аналізу відносно SHC 701 для ідентифікації характерних компонентів звукового поля (D) 703 і сумарної кількості фонових компонентів (BG_{TOT}) 705. У порівнянні з пристроями аудіокодування, описаними вище, пристрій 700A аудіокодування не здійснює це визначення відносно розкладань SHC 701, але безпосередньо відносно SHC 701.

[1217] Блок 704 синтезу на векторній основі представляє блок, виконаний з можливістю здійснення деякої форми векторного синтезу відносно SHC 701, наприклад SVD, KLT, PCA або будь-якого іншого векторного синтезу, для генерації, у випадках SVD, матриці 707 [US], що має розмір $M \times (N+1)^2$, і матриці 709 [V], що має розмір $(N+1)^2 \times (N+1)^2$. Матриця 707 [US] може представляти матрицю, одержану матричним множенням матриці [U] і матриці [S], генерованої шляхом застосування SVD до SHC 701.

[1218] Блок 706 приведення векторів може представляти блок, виконаний з можливістю зниження кількості векторів матриці 707 [US] і матриці 709 [V] таким чином, що кожний з векторів, що залишилися, матриці 707 [US] і матриці 709 [V] ідентифікує характерний або помітний компонент звукового поля. Блок 706 приведення векторів може здійснювати це приведення на основі кількості характерних компонентів D 703. Кількість характерних компонентів D 703 може, фактично, представляти масив чисел, де кожне число ідентифікує різні характерні вектори матриць 707 і 709. Блок 706 приведення векторів може виводити приведену матрицю 711 [US] розміром $M \times D$ і приведену матрицю 713 [V] розміром $(N+1)^2 \times D$.

[1219] Хоча це не показано для спрощення ілюстрації, інтерполяція матриці 709 [V] може відбуватися до приведення матриці 709 [V] аналогічно тому, що більш детально описано вище. Крім того, хоча це не показано для спрощення ілюстрації, переупорядковування приведеної матриці 711 [US] і/або приведеної матриці 712 [V], як більш детально описано вище. Відповідно, методи не підлягають обмеженню цими й іншими відношеннями (наприклад, проекцією помилки або будь-яким іншим аспектом вищенаведених методів, описаних вище, але не показаних у прикладі, наведеному на фіг. 51).

[1220] Блок 708 психоакустичного кодування представляє блок, виконаний з можливістю здійснення психоакустичного кодування відносно матриці 711 [US] для генерації бітового потоку 715. Блок 710 приведення коефіцієнтів може представляти блок, виконаний з можливістю зниження кількості каналів приведеної матриці 713 [V]. Іншими словами, блок 710 приведення коефіцієнтів може представляти блок, виконаний з можливістю виключення коефіцієнтів характерних векторів V (які утворюють приведену матрицю 713 [V]), що мають мало або не мають інформації направленості. Як описано вище, у деяких прикладах, коефіцієнти характерних векторів V, відповідні базисним функціям першого і нульового порядку (позначеного вище як N_{BG}), забезпечують мало інформації направленості і, таким чином, можуть бути видалені з характерних векторів V (що іменуються вище "зниженням порядку"). У цьому прикладі, підвищена гнучкість може забезпечуватися не тільки для ідентифікації цих коефіцієнтів, що відповідають N_{BG} , але і для ідентифікації додаткових каналів НОА (які можуть позначатися змінною $TotalOfAddAmbHOAChan$) з набору $[(NBG+1)^2+1, (N+1)^2]$. Блок 702 аналізу може аналізувати SHC 701 для визначення BG_{TOT} , що може ідентифікувати не тільки $(N_{BG}+1)^2$, але і $TotalOfAddAmbHOAChan$. Блок 710 приведення коефіцієнтів потім може видаляти коефіцієнти, що відповідають $(N_{BG}+1)^2$, і $TotalOfAddAmbHOAChan$ із приведеної матриці 713 [V] для генерації малої матриці 717 [V] розміром $((N+1)^2-(BG_{TOT})) \times D$.

[1221] Блок 712 стиснення потім може здійснювати вищезгадане скалярне квантування і/або хафманівське кодування для стиснення малої матриці 717 [V], виводячи стиснену малу матрицю 717 [V] як інформацію 719 побічного каналу ("інформацію 719 побічного каналу"). Блок 712 стиснення може виводити інформацію 719 побічного каналу аналогічно показаному в прикладі, наведеному на фіг. 10A-10O(ii). У деяких прикладах, блок генерації бітового потоку, аналогічний описаним вище, може включати інформацію 719 побічного каналу в бітовий потік 715. Крім того, хоча він описано відносно бітового потоку 715, пристрій 700A аудіокодування може, як згадано вище, включати в себе тракт обробки фонових компонента, що забезпечує інший бітовий потік, де блок генерації бітового потоку, аналогічний описаним вище, може генерувати бітовий потік, аналогічний бітовому потоку 17, описаному вище, що включає в себе бітовий потік 715 і бітовий потік, що виводиться трактом обробки фонових компонента.

[1222] Відповідно до методів, описаних в цьому винаході, блок 702 аналізу може бути виконаний з можливістю визначення першого ненульового набору коефіцієнтів вектора, тобто векторів приведеної матриці 713 [V] у цьому прикладі, що підлягають використанню для представлення характерного компонента звукового поля. У деяких прикладах, блок 702 аналізу може визначити, що всі коефіцієнти кожного вектора, що утворює приведену матрицю 713 [V], підлягають включенню в інформацію 719 побічного каналу. Таким чином блок 702 аналізу може задавати BG_{TOT} рівним нулю.

[1223] Пристрій 700A аудіокодування може, таким чином, ефективно діяти зворотно описаному вище відносно таблиці, позначеної як "декодовані вектори". Крім того, пристрій 700A аудіокодування може вказувати елемент синтаксису в заголовку одиниці доступу (який може включати в себе один або більше кадрів), який з множини режимів конфігурації вибраний. Хоча він описаний як зазначений для кожної одиниці доступу, блок 702 аналізу може вказувати цей елемент синтаксису на покадровій основі або будь-якій іншій періодичній основі або неперіодичній основі (наприклад, один раз для всього бітового потоку). У будь-якому випадку, цей елемент синтаксису може містити два біти, що вказують, який з чотирьох режимів конфігурації був вибраний для вказання ненульового набору коефіцієнтів приведеної матриці 713 [V] для представлення аспектів направленості цього характерного компонента. Елемент синтаксису може бути позначений як "codedVVecLength". Таким чином, пристрій 700A аудіокодування може сигналізувати або інакше вказувати в бітовому потоці, який з чотирьох режимів конфігурації використовувався для вказання малої матриці 717 [V] у бітовому потоці. Хоча вони описані відносно чотирьох режимів конфігурації, методи не підлягають обмеженню чотирма режимами конфігурації, але передбачають будь-яку кількість режимів конфігурації, що включає в себе одиничний режим конфігурації або множину режимів конфігурації.

[1224] Різні аспекти методів, таким чином, можуть передбачати, що пристрій 700A аудіокодування виконаний з можливістю діяти згідно з наступними пунктами.

[1225] Пункт 133149-1F. Пристрій, який містить: один або більше процесорів, виконаних з можливістю вибору одного з множини режимів конфігурації, для указання ненульового набору коефіцієнтів вектора, причому вектор розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле і представляють характерний компонент звукового поля, і указання ненульового набору коефіцієнтів вектора на основі вибраного одного з множини режимів конфігурації.

[1226] Пункт 133149-2F. Пристрій за пунктом 133149-1F, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти.

[1227] Пункт 133149-3F. Пристрій за пунктом 133149-1F, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, що відповідають порядку, який перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

[1228] Пункт 133149-4F. Пристрій за пунктом 133149-1F, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, що відповідають порядку, який перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік, і не включає в себе щонайменше один з коефіцієнтів, що відповідають порядку, який перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

[1229] Пункт 133149-5F. Пристрій за пунктом 133149-1F, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти за винятком щонайменше одного з коефіцієнтів.

[1230] Пункт 133149-6F. Пристрій за пунктом 133149-1F, у якому один або більше процесорів додатково виконані з можливістю указання вибраного одного з множини режимів конфігурації в бітовому потоці.

[1231] Пункт 133149-1G. Пристрій, який містить: один або більше процесорів, виконаних з можливістю визначення одного з множини режимів конфігурації, що дозволяє витягати ненульовий набір коефіцієнтів вектора відповідно до одного з множини режимів конфігурації, причому вектор розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле і представляють характерний компонент звукового поля, і витягання ненульового набору коефіцієнтів вектора на основі одержаного одного з множини режимів конфігурації.

[1232] Пункт 133149-2G. Пристрій за пунктом 133149-1G, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти.

[1233] Пункт 133149-3G. Пристрій за пунктом 133149-1G, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, що відповідають порядку, який перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

[1234] Пункт 133149-4G. Пристрій за пунктом 133149-1G, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе ті з коефіцієнтів, що відповідають порядку, який перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік, і не включає в себе щонайменше один з коефіцієнтів, що відповідають порядку, який перевищує порядок базисної функції, якій відповідають один або більше з множини коефіцієнтів сферичних гармонік.

[1235] Пункт 133149-5G. Пристрій за пунктом 133149-1G, у якому один з множини режимів конфігурації вказує, що ненульовий набір коефіцієнтів включає в себе всі коефіцієнти за винятком щонайменше одного з коефіцієнтів.

[1236] Пункт 133149-6G. Пристрій за пунктом 133149-1G, у якому один або більше процесорів додатково виконані з можливістю, при визначенні одного з множини режимів конфігурації, визначати один з множини режимів конфігурації на основі значення, сигналізованого в бітовому потоці.

[1237] На фіг. 52 показана блок-схема, що демонструє інший приклад пристрою 750A аудіодекодування, який може реалізувати різні аспекти методів, описаних у цьому винаході, для реконструкції або зразкової реконструкції SHC 701. У прикладі, наведеному на фіг. 52, пристрій 750A аудіодекодування аналогічний пристрою 540D аудіодекодування, показаному в прикладі, наведеному на фіг. 41D, за винятком того, що блок 542 витягання приймає бітовий потік 715' (який аналогічний бітовому потоку 715, описаному вище відносно прикладу, наведеного на фіг. 51, за винятком того, що бітовий потік 715' також включає в себе аудіокодовану версію SHC_{BG} 752) і інформацію 719 побічного каналу. З цієї причини, блок витягання позначається "блок 542' витягання".

[1238] Крім того, блок 542' витягання відрізняється від блока 542 витягання тим, що блок 542' витягання включає в себе модифіковану форму блока 555 зняття стиснення V (який

показаний як "блок 555' зняття стиснення V" у прикладі, наведеному на фіг. 52). Блок 555' зняття стиснення V приймає інформацію 719 побічного каналу й елемент синтаксису, позначений codedVVecLength 754. Блок 542' витягання виділяє codedVVecLength 754 з бітового потоку 715' (і, в одному прикладі, із заголовка одиниці доступу, включеного в бітовий потік 715'). Блок 555' зняття стиснення V включає в себе блок 756 конфігурування режиму ("блок 756 конфігурування режиму") і блок 758 виділення, виконаний з можливістю діяти відповідно до будь-якого з вищеописаних режимів 760 конфігурації.

[1239] Блок 756 конфігурування режиму приймає елемент 754 синтаксису і вибирає один з режимів 760 конфігурації. Потім блок 756 конфігурування режиму конфігурує блок 758 виділення вибраним одним з режимів 760 конфігурації. Блок 758 виділення представляє блок, виконаний з можливістю діяти відповідно до будь-якого з режимів 760 конфігурації для виділення стисненої форми малих векторів 717 [V] з інформації 719 побічного каналу. Блок 758 виділення може діяти відповідно до оператора переключення, представленого у нижченаведеній таблиці.

Таблиця декодованих векторів

Синтаксис	Кільк. бітів	Мнемоніка
<code>decodeVVec(i)</code>		
<code>{</code>		
<code>switch CodedVVecLength {</code>		
<code>case 0: // повний вектор</code>		
<code>VVecLength=NumOfHoaCoeffs;</code>		
<code>for (m=0; m< VVecLength; ++m){</code>		
<code>VecCoeff[m]=m+1; }</code>		
<code>break;</code>		
<code>case 1: // більш низькі порядки видаляються</code>		
<code>VVecLength=NumOfHoaCoeffs</code>	-	
<code>MinNumOfCoeffsForAmbHOA;</code>		
<code>for (m=0; m< VVecLength; ++m){</code>		
<code>VecCoeff[m]=m+MinNumOfCoeffsForAmbHOA+1;</code>		
<code>}</code>		
<code>break;</code>		
<code>case 2:</code>		
<code>VVecLength=NumOfHoaCoeffs</code>	-	
<code>MinNumOfCoeffsForAmbHOA</code>	-	
<code>NumOfAddAmbHoaChan;</code>		
<code>n=0;</code>		
<code>for (m=0;m<NumOfHoaCoeffs-</code>		
<code>MinNumOfCoeffsForAmbHOA; ++m){</code>		
<code>c=m+MinNumOfCoeffsForAmbHOA+1;</code>		
<code>if (ismember(c, AmbCoeffIdx) == 0){</code>		
<code>VecCoeff[n]=c;</code>		
<code>n++;</code>		
<code>}</code>		
<code>}</code>		
<code>break;</code>		
<code>case 3:</code>		
<code>VVecLength=NumOfHoaCoeffs</code>	-	
<code>NumOfAddAmbHoaChan;</code>		
<code>n=0;</code>		
<code>for (m=0; m<NumOfHoaCoeffs; ++m){</code>		
<code>c=m+1;</code>		

if (ismember(c, AmbCoeffIdx) == 0){		
VecCoeff[n]=c;		
n++;		
}		
}		
}		
if (NbitsQ[i] == 5) { /* донорідний квантувач */		
for (m=0; m< VVecLength; ++m){		
VVec(k)[i][m]=(VecValue/128.0) - 1.0;	8	uimsbf
}		
}		
else { /* хатфманівське декодування */		
for (m=0; m< VVecLength; ++m){		
Idx=5;		
If (CbFlag[i] == 1) {		
idx=(min(3, max(1, ceil(sqrt(VecCoeff[m]) - 1)))		
);		
}		
else if (PFlag[i] == 1) {idx=4;}		
cid=huffDecode(huffmannTable[NbitsQ]. codebook[idx]; huffVal);	динамічна	huffDecode
if (cid > 0) {		
aVal=sgn=(sgnVal * 2) - 1;	1	bslbf
if (cid > 1) {		
aVal=sgn * (2.0^(cid -1)+intAddVal);	cid-1	uimsbf
}		
} else {aVal=0.0;}		
}		
}		
}		
Примітка:		
- функція кодера для однорідного квантувача:		min(255,
round((x+1.0) * 128.0));		
- кільк. бітів для мнемоніки huffDecode є динамічною.		

- 5 [1240] У вищенаведеній таблиці синтаксису, перший оператор переключення з чотирма випадками (випадками 0-3) забезпечує можливість визначення довжини кожного вектора малої матриці 717 [V] відносно кількості коефіцієнтів. Перший випадок, case 0, указує, що задані всі коефіцієнти для векторів V_{DIST}^T . Другий випадок, case 1, указує, що задані тільки коефіцієнти вектора V_{DIST}^T , що відповідають порядку, більшому, ніж MinNumOfCoeffsForAmbHOA, що може позначати те, що іменується вище $(N_{DIST}+1)-(N_{BG}+1)$. Третій випадок, case 2, аналогічний

другому випадку, але додатково віднімає коефіцієнти, ідентифіковані NumOfAddAmbHoaChan, який позначає змінну для вказання додаткових каналів (де "канали" означають конкретний коефіцієнт, що відповідає визначеному порядку, комбінації підпорядків), які відповідають порядку, що перевищує порядок N_{BG} . Четвертий випадок, case 3, указує, що задані тільки

5 коефіцієнти вектора V_{DIST}^T , що залишилися після видалення коефіцієнтів, ідентифікованих за допомогою NumOfAddAmbHoaChan.

[1241] У цьому відношенні, пристрій 750A аудіодекодування може діяти відповідно до методів, описаних в цьому винаході, для визначення першого ненульового набору коефіцієнтів вектора, що представляють характерний компонент звукового поля, причому вектор

10 розкладений з множини коефіцієнтів сферичних гармонік, що описують звукове поле.

[1242] Крім того, пристрій 750A аудіодекодування може бути виконаний з можливістю діяти відповідно до методів, описаних в цьому винаході, для визначення одного з множини режимів конфігурації, що дозволяє витягати ненульовий набір коефіцієнтів вектора відповідно до одного з множини режимів конфігурації, причому вектор розкладений з множини коефіцієнтів

15 сферичних гармонік, що описують звукове поле і представляють характерний компонент звукового поля, і витягання ненульового набору коефіцієнтів вектора на основі одержаного одного з множини режимів конфігурації.

[1243] На фіг. 53 показана блок-схема, що демонструє інший приклад пристрою 570 аудіокодування, який може здійснювати різні аспекти методів, описаних у цьому винаході. У прикладі, наведеному на фіг. 53, пристрій 570 аудіокодування може бути аналогічним одному або більше із пристроїв 510A-510J аудіокодування (де блок 528A зниження порядку передбачається включеним у блок 20 витягання компонентів звукового поля, але не показаний для спрощення ілюстрації). Однак пристрій 570 аудіокодування може включати в себе більш загальний блок 572 перетворення, який може містити блок 518 розкладання в деяких

25 прикладах.

[1244] На фіг. 54 показана блок-схема, що демонструє, більш детально, ілюстративну реалізацію пристрою 570 аудіокодування, показаного в прикладі, наведеному на фіг. 53. Як показано в прикладі, наведеному на фіг. 54, блок 572 перетворення пристрою 570 аудіокодування включає в себе блок 654 повороту. Блок 520 витягання компонентів звукового

30 поля пристрою 570 аудіокодування включає в себе блок 650 просторового аналізу, блок 652 аналізу характеристик контенту, блок 656 витягання когерентних компонентів і блок 658 витягання дифузних компонентів. Блок 514 аудіокодування пристрою 570 аудіокодування включає в себе машину 660 кодування AAC і машину 162 кодування AAC. Блок 516 генерації бітового потоку пристрою 570 аудіокодування включає в себе мультиплексор (MUX) 164.

[1245] Ширина смуги, вимірювана в бітах на секунду, необхідна для представлення 3D аудіоданих у формі SHC, може бути недостатньою для використання споживачем. Наприклад, при використанні частота дискретизації 48 кГц, і з тим же розрізненням 32 біти, представлення SHC четвертого порядку представляє ширину смуги 36 Мбіт/с ($25 \times 48000 \times 32$ біт/с). У порівнянні з традиційним аудіокодуванням для стереосигналів, що звичайно складає близько 100 кбіт/с, це велике значення. Методи, реалізовані в прикладі, наведеному на фіг. 54, можуть знижувати ширину смуги представлень 3D аудіо.

40

[1246] Блок 650 просторового аналізу, блок 652 аналізу характеристик контенту і блок 654 повороту можуть приймати SHC 511. Як описано в іншому місці цього розкриття, SHC 511 може представляти звукове поле. У прикладі, наведеному на фіг. 54, блок 650 просторового аналізу, блок 652 аналізу характеристик контенту і блок 654 повороту можуть приймати двадцять п'ять SHC для представлення звукового поля четвертого порядку ($n=4$).

45

[1247] Блок 650 просторового аналізу може аналізувати звукове поле, представлене SHC 511, для ідентифікації характерних компонентів звукового поля і дифузних компонентів звукового поля. Характерні компоненти звукового поля являють собою звуки, які сприймаються як такі, що надходять з ідентифікованого напрямку, або інакше відрізняються від фонових або дифузних компонентів звукового поля. Наприклад, звук, генерований окремим музичним інструментом, може сприйматися як такий, що надходить з ідентифікованого напрямку. Навпаки, дифузні або фонові компоненти звукового поля не сприймаються як такі, що надходять з ідентифікованого напрямку. Наприклад, звук вітру в лісі може бути дифузним компонентом

50 звукового поля.

[1248] Блок 650 просторового аналізу може ідентифікувати один або більше характерних компонентів у спробі ідентифікувати оптимальний кут повороту звукового поля для вирівнювання характерних компонентів, що мають найбільшу енергію, з вертикальною і/або горизонтальною віссю (відносно передбачуваного мікрофона, що записував це звукове поле). Блок 650 просторового аналізу може ідентифікувати цей оптимальний кут, завдяки чому звукове

60

поле може повертатися таким чином, що ці характерні компоненти краще вирівнюються зі сферичними базисними функціями, що лежать в основі, показаними в прикладах, наведених на фіг. 1 і 2.

[1249] У деяких прикладах, блок 650 просторового аналізу може представляти блок, виконаний з можливістю здійснення форми аналізу дифузії для ідентифікації відсотка звукового поля, представленого SHC 511, що включає в себе дифузні звуки (які можуть означати звуки, що мають низькі рівні направленості або SHC більш низького порядку, тобто SHC 511, що має порядок, який менше або дорівнює одиниці). У порядку одного прикладу, блок 650 просторового аналізу може здійснювати аналіз дифузії аналогічно описаному в статті Ville Pulkki, за назвою "Spatial Sound Reproduction with Directional Audio Coding", опублікованій в J. Audio Eng. Soc., Vol. 55, № 6, датованій червнем 2007 р. У ряді випадків, блок 650 просторового аналізу може аналізувати тільки ненульовий піднабір коефіцієнтів НОА, наприклад коефіцієнти НОА нульового і першого порядку SHC 511, при здійсненні аналізу дифузії для визначення відсотка дифузії.

[1250] Блок 652 аналізу характеристик контенту може визначати, щонайменше частково, на основі SHC 511, чи згенеровані SHC 511 за допомогою природного запису звукового поля або створені штучно (тобто синтетично), у порядку одного прикладу, з аудіооб'єкта, наприклад об'єкта IKM. Крім того, блок 652 аналізу характеристик контенту може потім визначати, щонайменше частково, на основі того, чи згенеровані SHC 511 за допомогою фактичного запису звукового поля або зі штучного аудіооб'єкта, сумарну кількість каналів для включення в бітовий потік 517. Наприклад, блок 652 аналізу характеристик контенту може визначати, щонайменше частково, на основі того, чи згенеровані SHC 511 із запису фактичного звукового поля або зі штучного аудіооб'єкта, що в бітовий потік 517 потрібно включати шістнадцять каналів. Кожний з каналів може бути моноканалом. Блок 652 аналізу характеристик контенту може додатково здійснювати визначення сумарної кількості каналів для включення в бітовий потік 517 на основі вихідної бітової швидкості бітового потоку 517, наприклад 1,2 Мбіт/с.

[1251] Крім того, блок 652 аналізу характеристик контенту може визначати, щонайменше частково, на основі того, чи згенеровані SHC 511 із запису фактичного звукового поля або зі штучного аудіооб'єкта, скільки каналів виділяти когерентним або, іншими словами, характерним компонентам звукового поля і скільки каналів виділяти дифузним або, іншими словами, фоновим компонентам звукового поля. Наприклад, коли SHC 511 згенеровані з запису фактичного звукового поля з використанням, у порядку одного прикладу, Eigenmics, блок 652 аналізу характеристик контенту може виділяти три канали когерентним компонентам звукового поля і може виділяти канали, що залишилися, дифузним компонентам звукового поля. У цьому прикладі, коли SHC 511 згенеровані зі штучного аудіооб'єкта, блок 652 аналізу характеристик контенту може виділяти п'ять каналів когерентним компонентам звукового поля і може виділяти канали, що залишилися, дифузним компонентам звукового поля. Таким чином, блок аналізу контенту (тобто блок 652 аналізу характеристик контенту) може визначати тип звукового поля (наприклад, дифузний/направлений і т. д.) і по черзі визначати кількість когерентних/дифузних компонентів для витягання.

[1252] Цільова бітова швидкість може впливати на кількість компонентів і бітову швидкість окремих машин кодування AAC (наприклад, машин 660, 662 кодування AAC). Іншими словами, блок 652 аналізу характеристик контенту може додатково здійснювати визначення, скільки каналів виділяти когерентним компонентам і скільки каналів виділяти дифузним компонентам на основі вихідної бітової швидкості бітового потоку 517, наприклад 1,2 Мбіт/с.

[1253] У деяких прикладах, канали, виділені когерентним компонентам звукового поля, можуть мати більш високі бітові швидкості, ніж канали, виділені дифузним компонентам звукового поля. Наприклад, максимальна бітова швидкість бітового потоку 517 може дорівнювати 1,2 Мбіт/с. У цьому прикладі, може існувати чотири канали, виділені когерентним компонентам, і 16 каналів, виділених дифузним компонентам. Крім того, у цьому прикладі, кожен з каналів, виділених когерентним компонентам, може мати максимальну бітову швидкість 64 кбіт/с. У цьому прикладі, кожен з каналів, виділених дифузним компонентам, може мати максимальну бітову швидкість 48 кбіт/с.

[1254] Як згадано вище, блок 652 аналізу характеристик контенту може визначати, чи згенеровані SHC 511 із запису фактичного звукового поля або зі штучного аудіооб'єкта. Блок 652 аналізу характеристик контенту може робити це визначення по-різному. Наприклад, пристрій 570 аудіокодування може використовувати SHC 4-го порядку. У цьому прикладі, блок 652 аналізу характеристик контенту може кодувати 24 канали і прогнозувати 25-й канал (який може бути представлений як вектор). Блок 652 аналізу характеристик контенту може застосовувати скаляри до щонайменше деяких з 24 каналів і додавати результуючі значення

для визначення 25-го вектора. Крім того, у цьому прикладі, блок 652 аналізу характеристик контенту може визначати точність прогнозованого 25-го каналу. У цьому прикладі, якщо точність прогнозованого 25-го каналу відносно висока (наприклад, точність перевищує конкретний поріг), SHC 511, імовірно, генерується із синтетичного аудіооб'єкта. Навпаки, якщо точність прогнозованого 25-го каналу відносно низька (наприклад, точність нижче конкретного порога), SHC 511 з більшою мірою імовірності представляють записане звукове поле. Наприклад, у цьому прикладі, якщо відношення сигнал/шум (SNR) 25-го каналу більше 100 децибелів (дБ), SHC 511 з більшою мірою імовірності представляють звукове поле, генероване із синтетичного аудіооб'єкта. Навпаки, SNR звукового поля, записаного з використанням власного мікрофона, може складати від 5 до 20 дБ. Таким чином, може існувати очевидна відмінність у відношенні SNR між звуковим полем, представленим SHC 511, генерованими з фактичного прямого запису і із синтетичного аудіооб'єкта.

[1255] Крім того, блок 652 аналізу характеристик контенту може вибирати, щонайменше частково, на основі того, чи згенеровані SHC 511 із запису фактичного звукового поля або зі штучного аудіооб'єкта, кодові книги для квантування вектора V . Іншими словами, блок 652 аналізу характеристик контенту може вибирати різні кодові книги для використання при квантуванні вектора V , залежно від того, чи є звукове поле, представлене коефіцієнтами НОА, записаним або синтетичним.

[1256] У деяких прикладах, блок 652 аналізу характеристик контенту може визначати, на повторюваній основі, чи згенеровані SHC 511 із запису фактичного звукового поля або зі штучного аудіооб'єкта. У деяких таких прикладах, основою повторення може бути кожен кадр. В інших прикладах, блок 652 аналізу характеристик контенту може здійснювати це визначення один раз. Крім того, блок 652 аналізу характеристик контенту може визначати, на повторюваній основі, сумарну кількість каналів і виділення каналів когерентного компонента і каналів дифузного компонента. У деяких таких прикладах, основою повторення може бути кожен кадр. В інших прикладах, блок 652 аналізу характеристик контенту може здійснювати це визначення один раз. У деяких прикладах, блок 652 аналізу характеристик контенту може вибирати, на повторюваній основі, кодові книги для використання при квантуванні вектора V . У деяких таких прикладах, основою повторення може бути кожен кадр. В інших прикладах, блок 652 аналізу характеристик контенту може здійснювати це визначення один раз.

[1257] Блок 654 повороту може здійснювати операцію повороту коефіцієнтів НОА. Як розглянуто в іншому місці цього розкриття (наприклад, з посиланням на фіг. 55 і 55B), здійснення операції повороту може знижувати кількість бітів, необхідну для представлення SHC 511. У деяких прикладах, аналіз повороту, здійснюваний блоком повороту 652, є прикладом аналізу розкладання на сингулярні значення ("SVD"). Аналіз головних компонентів ("PCA"), аналіз незалежних компонентів ("ICA") і перетворення Карунена-Лева ("KLT") є спорідненими методами, які можна застосовувати.

[1258] У прикладі, наведеному на фіг. 54, блок 656 витягання когерентних компонентів приймає повернений SHC 511 від блока 654 повороту. Крім того, блок 656 витягання когерентних компонентів витягає, з повернених SHC 511, ті з повернених SHC 511, що пов'язані з когерентними компонентами звукового поля.

[1259] Крім того, блок 656 витягання когерентних компонентів генерує один або більше каналів когерентного компонента. Кожний з каналів когерентного компонента може включати в себе окремий піднабір повернених SHC 511, пов'язаних з когерентними коефіцієнтами звукового поля. У прикладі, наведеному на фіг. 54, блок 656 витягання когерентних компонентів може генерувати від одного до 16 каналів когерентного компонента. Кількість каналів когерентного компонента, генерованих блоком 656 витягання когерентних компонентів, можна визначити по кількості каналів, виділених блоком 652 аналізу характеристик контенту когерентним компонентам звукового поля. Бітові швидкості каналів когерентного компонента, генерованих блоком 656 витягання когерентних компонентів, можуть визначатися блоком 652 аналізу характеристик контенту.

[1260] Аналогічно, у прикладі, наведеному на фіг. 54, блок 658 витягання дифузних компонентів приймає повернений SHC 511 від блока 654 повороту. Крім того, блок 658 витягання дифузних компонентів витягає, з повернених SHC 511, ті з повернених SHC 511, що пов'язані з дифузними компонентами звукового поля.

[1261] Крім того, блок 658 витягання дифузних компонентів генерує один або більше каналів дифузного компонента. Кожний з каналів дифузного компонента може включати в себе окремий піднабір повернених SHC 511, пов'язаних з дифузними коефіцієнтами звукового поля. У прикладі, наведеному на фіг. 54, блок 658 витягання дифузних компонентів може генерувати від одного до 9 каналів дифузного компонента. Кількість каналів дифузного компонента,

генерованих блоком 658 витягання дифузних компонентів, можна визначити по кількості каналів, виділених блоком 652 аналізу характеристик контенту дифузним компонентам звукового поля. Бітові швидкості каналів дифузного компонента, генерованих блоком 658 витягання дифузних компонентів, можуть визначатися блоком 652 аналізу характеристик контенту.

[1262] У прикладі, наведеному на фіг. 54, блок 660 кодування AAC може використовувати кодек AAC для кодування каналів когерентного компонента, генерованих блоком 656 витягання когерентних компонентів. Аналогічно, блок 662 кодування AAC може використовувати кодек AAC для кодування каналів дифузного компонента, генерованих блоком 658 витягання дифузних компонентів. Мультіплексор 664 ("MUX 664") може мультіплексувати кодовані канали когерентного компонента і кодовані канали дифузного компонента, разом з побічними даними (наприклад, оптимальний кут, визначений блоком 650 просторового аналізу), для генерації бітового потоку 517.

[1263] Таким чином, методи можуть передбачати, що пристрій 570 аудіокодування визначає, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

[1264] У деяких прикладах, пристрій 570 аудіокодування може визначати, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, піднабір коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля. У цих і інших прикладах, пристрій 570 аудіокодування може генерувати бітовий потік для включення піднабору коефіцієнтів сферичних гармонік. Пристрій 570 аудіокодування може, у ряді випадків, аудіокодувати піднабір коефіцієнтів сферичних гармонік і генерувати бітовий потік для включення аудіокодованого піднабору коефіцієнтів сферичних гармонік.

[1265] У деяких прикладах, пристрій 570 аудіокодування може визначати, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, піднабір коефіцієнтів сферичних гармонік, що представляють фонові компоненти звукового поля. У цих і інших прикладах, пристрій 570 аудіокодування може генерувати бітовий потік для включення піднабору коефіцієнтів сферичних гармонік. У цих і інших прикладах, пристрій 570 аудіокодування може аудіокодувати піднабір коефіцієнтів сферичних гармонік і генерувати бітовий потік для включення аудіокодованого піднабору коефіцієнтів сферичних гармонік.

[1266] У деяких прикладах, пристрій 570 аудіокодування може здійснювати просторовий аналіз відносно коефіцієнтів сферичних гармонік для ідентифікації кута повороту звукового поля, представленого коефіцієнтами сферичних гармонік, і здійснення операції повороту для повороту звукового поля на ідентифікований кут для генерації повернених коефіцієнтів сферичних гармонік.

[1267] У деяких прикладах, пристрій 570 аудіокодування може визначати, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, перший піднабір коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля, і визначати, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, другий піднабір коефіцієнтів сферичних гармонік, що представляють фонові компоненти звукового поля. У цих і інших прикладах, пристрій 570 аудіокодування може аудіокодувати перший піднабір коефіцієнтів сферичних гармонік, що мають більш високу цільову бітову швидкість, ніж використовується для аудіокодування другого піднабору коефіцієнтів сферичних гармонік.

[1268] Таким чином, різні аспекти методів можуть передбачати, що пристрій 570 аудіокодування визначає, чи генеруються SCH 511 із синтетичного аудіооб'єкта, згідно з наступними пунктами.

[1269] Пункт 132512-1. Пристрій, наприклад пристрій 570 аудіокодування, у якому один або більше процесорів додатково виконані з можливістю визначення, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта.

[1270] Пункт 132512-2. Пристрій за пунктом 132512-1, у якому один або більше процесорів додатково виконані з можливістю, при визначенні, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, виключення першого вектора з матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше частина коефіцієнтів сферичних гармонік, що представляють звукове поле, для одержання приведеної матриці кадрованих коефіцієнтів сферичних гармонік.

[1271] Пункт 132512-3. Пристрій за пунктом 132512-1, у якому один або більше процесорів додатково виконані з можливістю, при визначенні, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, виключення першого вектора з матриці кадрованих коефіцієнтів сферичних гармонік, де зберігається щонайменше

сферичних гармонік із синтетичного аудіооб'єкта в бітовому потоці, де зберігається стиснена версія коефіцієнтів сферичних гармонік.

[1280] Пункт 132512-12. Пристрій за пунктом 132512-11, у якому указання є одиничним бітом.

5 [1281] Пункт 132512-13. Пристрій за пунктом 132512-1, у якому один або більше процесорів додатково виконані з можливістю визначення, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, піднабору коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля.

10 [1282] Пункт 132512-14. Пристрій за пунктом 132512-13, у якому один або більше процесорів додатково виконані з можливістю генерації бітового потоку для включення піднабору коефіцієнтів сферичних гармонік.

[1283] Пункт 132512-15. Пристрій за пунктом 132512-13, у якому один або більше процесорів додатково виконані з можливістю аудіокодування піднабору коефіцієнтів сферичних гармонік і генерації бітового потоку для включення аудіокодованого піднабору коефіцієнтів сферичних гармонік.

15 [1284] Пункт 132512-16. Пристрій за пунктом 132512-1, у якому один або більше процесорів додатково виконані з можливістю визначення, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, піднабору коефіцієнтів сферичних гармонік, що представляють фонові компоненти звукового поля.

20 [1285] Пункт 132512-17. Пристрій за пунктом 132512-16, у якому один або більше процесорів додатково виконані з можливістю генерації бітового потоку для включення піднабору коефіцієнтів сферичних гармонік.

[1286] Пункт 132512-18. Пристрій за пунктом 132512-15, у якому один або більше процесорів додатково виконані з можливістю аудіокодування піднабору коефіцієнтів сферичних гармонік і генерації бітового потоку для включення аудіокодованого піднабору коефіцієнтів сферичних гармонік.

[1287] Пункт 132512-18. Пристрій за пунктом 132512-1, у якому один або більше процесорів додатково виконані з можливістю здійснення просторового аналізу відносно коефіцієнтів сферичних гармонік для ідентифікації кута повороту звукового поля, представленого коефіцієнтами сферичних гармонік, і здійснення операції повороту для повороту звукового поля на ідентифікований кут для генерації повернених коефіцієнтів сферичних гармонік.

[1288] Пункт 132512-20. Пристрій за пунктом 132512-1, у якому один або більше процесорів додатково виконані з можливістю визначення, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, першого піднабору коефіцієнтів сферичних гармонік, що представляють характерні компоненти звукового поля, і визначення, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, другого піднабору коефіцієнтів сферичних гармонік, що представляють фонові компоненти звукового поля.

40 [1289] Пункт 132512-21. Пристрій за пунктом 132512-20, у якому один або більше процесорів додатково виконані з можливістю аудіокодування першого піднабору коефіцієнтів сферичних гармонік, що мають більш високу цільову бітову швидкість, ніж використовується для аудіокодування другого піднабору коефіцієнтів сферичних гармонік.

45 [1290] Пункт 132512-22. Пристрій за пунктом 132512-1, у якому один або більше процесорів додатково виконані з можливістю здійснення розкладання на сингулярні значення відносно коефіцієнтів сферичних гармонік для генерації матриці U , яка представляє ліві сингулярні вектори з множини коефіцієнтів сферичних гармонік, матриці S , яка представляє сингулярні значення з множини коефіцієнтів сферичних гармонік, і матриці V , яка представляє праві сингулярні вектори з множини коефіцієнтів сферичних гармонік.

50 [1291] Пункт 132512-23. Пристрій за пунктом 132512-22, у якому один або більше процесорів додатково виконані з можливістю визначення, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, ділянок однієї або більше з матриці U , матриці S і матриці V , що представляють характерні компоненти звукового поля.

55 [1292] Пункт 132512-24. Пристрій за пунктом 132512-22, у якому один або більше процесорів додатково виконані з можливістю визначення, на основі того, чи генеруються коефіцієнти сферичних гармонік із синтетичного аудіооб'єкта, ділянок однієї або більше з матриці U , матриці S і матриці V , що представляють фонові компоненти звукового поля.

60 [1293] Пункт 132512-1C. Пристрій, наприклад пристрій 570 аудіокодування, який містить: один або більше процесорів, виконаних з можливістю визначення, чи генеруються коефіцієнти сферичних гармонік, що представляють звукове поле, із синтетичного аудіооб'єкта, на основі відношення, обчисленого як функція щонайменше енергії вектора коефіцієнтів сферичних

гармонік і помилки, виведеної на основі прогнозованої версії вектора коефіцієнтів сферичних гармонік, і вектора коефіцієнтів сферичних гармонік.

[1294] У кожному з різних вищеописаних випадків, варто розуміти, що пристрій 570 аудіокодування може здійснювати спосіб або іншим чином містити засіб для здійснення кожного етапу способу, для якого пристрій 570 аудіокодування виконаний з можливістю здійснення в ряді випадків, цей засіб може містити один або більше процесорів. У ряді випадків, один або більше процесорів можуть представляти процесор спеціального призначення, що виконує інструкції, які зберігаються на довгостроковому зчитуваному комп'ютером носії даних. Іншими словами, різні аспекти методів у кожному з наборів прикладів кодування можуть передбачати довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам здійснювати спосіб, для виконання якого призначений пристрій 570 аудіокодування.

[1295] На фіг. 55A і 55B показані схеми, що демонструють приклад здійснення різних аспектів методів, описаних у цьому винаході, для повороту звукового поля 640. На фіг. 55A показана схема, що демонструє звукове поле 640 до повороту відповідно до різних аспектів методів, описаних у цьому винаході. У прикладі, наведеному на фіг. 55A, звукове поле 640 включає в себе два положення високого тиску, позначені як положення 642A і 642B. Ці положення 642A і 642B ("положення 642") розташовуються уздовж лінії 644, яка має ненульовий нахил (що є іншим способом указання лінії, яка не є горизонтальною, оскільки горизонтальні лінії мають нульовий нахил). За умови, що положення 642 мають координату z крім координат x і y, можуть знадобитися сферичні базисні функції більш високого порядку для правильного представлення цього звукового поля 640 (оскільки ці сферичні базисні функції більш високого порядку описують верхні і нижні або негоризонтальні ділянки звукового поля). Замість приведення звукового поля 640 безпосередньо до SHC 511, пристрій 570 аудіокодування може повертати звукове поле 640 поки лінія 644, що з'єднує положення 642, не стане горизонтальною.

[1296] На фіг. 55B показана схема, що демонструє звукове поле 640 після повороту, поки лінія 644, що з'єднує положення 642, не стане горизонтальною. Таким чином, в результаті повороту звукового поля 640, SHC 511 можна виводити таким чином, що SHC 511 більш високого порядку задаються рівними нулю, за умови, що повернене звукове поле 640 більше не має ніяких положень тиску (або енергії) з координатами z. Таким чином, пристрій 570 аудіокодування може повертати, переміщати або, у більш загальному випадку, регулювати звукове поле 640 для зниження кількості SHC 511, що мають ненульові значення. Потім, спільно з різними іншими аспектами методів, пристрій 570 аудіокодування може, замість того, щоб сигналізувати 32-бітове знакове число, що ідентифікує, що ці SHC 511 більш високого порядку мають нульові значення, сигналізувати в полі бітового потоку 517, що ці SHC 511 більш високого порядку не сигналізуються. Пристрій 570 аудіокодування також може вказувати інформацію повороту в бітовому потоці 517, яка вказує, як було повернене звукове поле 640, часто за допомогою вираження азимута й підвищення вищеописаним чином. Потім пристрій витягання, наприклад пристрій аудіокодування, може виводити, що ці несигналізовані SHC 511 мають нульове значення і, при відтворенні звукового поля 640 на основі SHC 511, здійснювати поворот для повороту звукового поля 640, завдяки чому звукове поле 640 подібне зі звуковим полем 640, показаним у прикладі, наведеному на фіг. 55. Таким чином, пристрій 570 аудіокодування може знижувати кількість SHC 511, які необхідно вказувати в бітовому потоці 517 відповідно до методів, описаних в цьому винаході.

[1297] Алгоритм 'просторового ущільнення' можна використовувати для визначення оптимального повороту звукового поля. В одному варіанті здійснення, пристрій 570 аудіокодування може здійснювати алгоритм для ітерації по всіх можливих комбінаціях азимута й підвищення (тобто комбінаціях 1024×512 у вищенаведеному прикладі), повертаючи звукове поле для кожної комбінації і обчислюючи кількість SHC 511, що перевищує порогове значення. Комбінацію можливих азимутів/підвищень, яка створює найменшу кількість SHC 511, що перевищує порогове значення, можна розглядати як так називаний "оптимальний поворот". У цій поверненій формі, звукове поле може вимагати найменшої кількості SHC 511 для представлення звукового поля і потім може розглядатися як ущільнене. У ряді випадків, регулювання може включати цей оптимальний поворот, і інформація регулювання, описана вище, може включати в себе цю інформацію повороту (яка може іменуватися "оптимальним поворотом") (відносно азимутальних кутів і кутів підвищення).

[1298] У ряді випадків, замість того, щоб вказувати тільки азимутальний кут і кут підвищення, пристрій 570 аудіокодування може вказувати додаткові кути у формі, у порядку одного прикладу, кутів Ейлера. Кути Ейлера вказують кут повороту навколо осі z, згаданої осі x і

згаданої осі z . Хоча вони описані в цьому винаході відносно комбінацій азимутальних кутів і кутів підвищення, методи цього винаходу не підлягають обмеженню указанням тільки азимутальних кутів і кутів підвищення, але можуть включати в себе указання будь-якої кількості кутів, у тому числі трьох вищезгаданих кутів Ейлера. У цьому розумінні, пристрій 570 аудіокодування може повертати звукове поле для зниження кількості множини ієрархічних елементів, що забезпечують інформацію, яка стосується опису звукового поля, і вказувати кути Ейлера як інформацію повороту в бітовому потоці. Кути Ейлера, як згадано вище, можуть описувати, як повернене звукове поле. При використанні кутів Ейлера, пристрій витягання бітового потоку може виділяти бітовий потік для визначення інформації повороту, що включає в себе кути Ейлера і, при відтворенні звукового поля на основі тих з множини ієрархічних елементів, що забезпечують інформацію, яка стосується опису звукового поля, повертати звукове поле на основі кутів Ейлера.

[1299] Крім того, у ряді випадків, замість явного указання цих кутів у бітовому потоці 517, пристрій 570 аудіокодування може вказувати індекс (який може іменуватися "індексом повороту"), пов'язаний із попередньо заданими комбінаціями одного або більше кутів, що вказують поворот. Іншими словами, інформація повороту може, у ряді випадків, включати в себе індекс повороту. У цих випадках, дане значення індексу повороту, наприклад значення нуль, може вказувати, що поворот не здійснювався. Цей індекс повороту можна використовувати в зв'язку з таблицею повороту. Таким чином, пристрій 570 аудіокодування може включати в себе таблицю повороту, що містить елемент для кожної з комбінацій азимутального кута і кута підвищення.

[1300] Альтернативно, таблиця повороту може включати в себе елемент для кожного матричного перетворення, що представляє кожну комбінацію азимутального кута і кута підвищення. Таким чином, пристрій 570 аудіокодування може зберігати таблицю повороту, що має елемент для кожного матричного перетворення для повороту звукового поля за допомогою кожної з комбінацій азимутальних кутів і кутів підвищення. Звичайний пристрій 570 аудіокодування приймає SHC 511 і виводить SHC 511', коли поворот здійснюється, відповідно до наступного рівняння:

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

У вищенаведеному рівнянні, SHC 511' обчислюються як функція матриці кодування для кодування звукового поля відносно другої системи координат ($EncMat_2$), матриця інверсії для повернення SHC 511 зворотно у звукове поле відносно першої системи координат ($InvMat_1$) і SHC 511. $EncMat_2$ має розмір 25×32 , тоді як $InvMat_2$ має розмір 32×25 . SHC 511' і SHC 511 мають розмір 25, де SHC 511' може бути додатково приведена внаслідок видалення тих, котрі не вказують помітну аудіоінформацію. $EncMat_2$ може змінюватися для кожної комбінації азимутального кута і кута підвищення, тоді як $InvMat_1$ може залишатися статичною відносно кожної комбінації азимутального кута і кута підвищення. Таблиця повороту може включати в себе елемент, що зберігає результат множення кожної окремої $EncMat_2$ на $InvMat_1$.

[1301] На фіг. 56 показана схема, що демонструє приклад звукового поля, захопленого відносно першої системи координат, що потім повертається відповідно до методів, описаних в цьому винаході, для вираження звукового поля відносно другої системи координат. У прикладі, наведеному на фіг. 56, звукове поле, що оточує власний мікрофон 646, захоплюється, передбачаючи першу систему координат, що позначається осями X_1 , Y_1 і Z_1 у прикладі, наведеному на фіг. 56. SHC 511 описують звукове поле відносно цієї першої системи координат. $InvMat_1$ перетворює SHC 511 зворотно у звукове поле, дозволяючи повертати звукове поле в другу систему координат, позначену осями X_2 , Y_2 і Z_2 у прикладі, наведеному на фіг. 56. Описана вище $EncMat_2$ може повертати звукове поле і генерувати SHC 511', що описує це повернене звукове поле відносно другої системи координат.

[1302] У будь-якому випадку, вищенаведене рівняння можна виводити наступним чином. За умови, що звукове поле записується у визначеній системі координат, таким чином, що фронт вважається напрямком осі x , 32 позиції мікрофонів власного мікрофона (або інші конфігурації мікрофонів) задаються в цій системі координат. Поворот звукового поля потім може розглядатися як поворот цієї системи координат. Для передбачуваної системи координат, SHC 511 можна обчислювати наступним чином:

$$\begin{bmatrix} SHC \\ 27 \end{bmatrix} = \begin{bmatrix} Y_0^0(Pos_1) & Y_0^0(Pos_2) & \cdots & Y_0^0(Pos_{32}) \\ Y_1^{-1}(Pos_1) & \cdots & & Y_1^{-1}(Pos_{32}) \\ \vdots & \ddots & & \\ Y_4^4(Pos_1) & & & Y_4^4(Pos_{32}) \end{bmatrix} \begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix}.$$

У вищенаведеному рівнянні, Y_n^m представляють сферичні базисні функції в позиції (Pos_i) i -го мікрофона (де i може бути 1-32 у цьому прикладі). Вектор mic_i позначає сигнал мікрофона для i -го мікрофона для часу t . Позиції (Pos_i) означають позицію мікрофона в першій системі координат (тобто системі координат до повороту в цьому прикладі).

[1303] Вищенаведене рівняння можна виразити альтернативно відносно математичних виразів, позначених вище як:

$$[SHC \text{ } 27] = [E_s(\theta, \varphi)][m_i(t)]$$

[1304] Для повороту звукового поля (або в другій системі координат), позиція (Pos_i) буде обчислена в другій системі координат. За умови, що присутні вихідні сигнали мікрофона, звукове поле можна довільно повертати. Однак вихідні сигнали мікрофона $(mic_i(t))$ часто бувають недоступні. Задача може полягати в тому, як витягати сигнали мікрофона $(mic_i(t))$ з SHC 511. Якщо Т-подібна використовується (як у 32-мікрофонному власному мікрофоні), вирішити цю задачу можна, розв'язуючи наступне рівняння:

$$\begin{bmatrix} mic_1(t) \\ mic_2(t) \\ \vdots \\ mic_{32}(t) \end{bmatrix} = [InvMat_1] \begin{bmatrix} SHC \\ 27 \end{bmatrix}$$

Ця $InvMat_1$ може вказувати сферичні гармонічні базисні функції, обчислені відповідно до позиції мікрофонів, як зазначено відносно першої системи координат. Це рівняння також може бути виражене як $[m_i(t)] = [E_s(\theta, \varphi)]^{-1}[SHC]$, як згадано вище.

[1305] Після витягання сигналів мікрофона $(mic_i(t))$ відповідно до вищенаведеного рівняння, сигнали мікрофона $(mic_i(t))$, що описують звукове поле, можуть повертатися для обчислення SHC 511', що відповідає другій системі координат, що дає наступне рівняння:

$$\begin{bmatrix} SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (25 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 25) \end{bmatrix} \begin{bmatrix} SHC \\ 27 \end{bmatrix}.$$

$EncMat_2$ указує сферичні гармонічні базисні функції з поверненої позиції (Pos_i') . Таким чином, $EncMat_2$ може ефективно вказувати комбінацію азимутального кута і кута підвищення. Таким чином, коли в таблиці повороту зберігається результат

$$\begin{bmatrix} EncMat_2 \\ (25 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 25) \end{bmatrix}$$

для кожної комбінації азимутальних кутів і кутів підвищення, таблиця повороту ефективно вказує кожну комбінацію азимутальних кутів і кутів підвищення. Вищенаведене рівняння також може бути виражене у вигляді:

$$[SHC \text{ } 27'] = [E_s(\theta_2, \varphi_2)][E_s(\theta_1, \varphi_1)]^{-1}[SHC \text{ } 27],$$

де θ_2, φ_2 представляють другий азимутальний кут і другий кут підвищення, що відрізняються від перших азимутального кута і кута підвищення, представлених як θ_1, φ_1 .

θ_1, φ_1 відповідають першій системі координат, тоді як θ_2, φ_2 відповідають другій системі координат. Таким чином $InvMat_1$ може відповідати $[E_s(\theta_1, \varphi_1)]^{-1}$, тоді як $EncMat_2$ може відповідати $[E_s(\theta_2, \varphi_2)]$.

[1306] Вищеописане може представляти більш спрощену версію обчислення без урахування операції фільтрації, представленої вище в різних рівняннях, що позначають виведення SHC 511 у частотній області функцією $j_n(\cdot)$, що означає сферичну бesselеву функцію порядку n . В часовій області, ця функція $j_n(\cdot)$ представляє операції фільтрації, що стосуються конкретного порядку, n . За допомогою фільтрації, поворот можна здійснювати для кожного порядку. Для ілюстрації, розглянемо нижченаведені рівняння:

$$\begin{aligned} a_n^k(t) &\square b_n(t) * \langle [Y_n^m] [m_i(t)] \rangle, \\ a_n^k(t) &\square \langle [Y_n^m] [b_n(t) * m_i(t)] \rangle. \end{aligned}$$

[1307] З цих рівнянь, повернені SHC 511' для порядків здійснюються окремо, оскільки $b_n(t)$ відрізняються для кожного порядку. У результаті, вищенаведене рівняння можна перетворити наступним чином для обчислення повернених SHC 511' першого порядку:

$$\begin{bmatrix} 1^{st} \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (3 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 3) \end{bmatrix} \begin{bmatrix} 1^{st} \\ Order \\ SHC \\ 27 \end{bmatrix}.$$

За умови наявності трьох SHC 511 першого порядку, кожний з векторів SHC 511' і 511 має розмір три у вищенаведеному рівнянні. Аналогічно, для другого порядку, може застосовуватися наступне рівняння:

$$\begin{bmatrix} 2^{nd} \\ Order \\ SHC \\ 27' \end{bmatrix} = \begin{bmatrix} EncMat_2 \\ (5 \times 32) \end{bmatrix} \begin{bmatrix} InvMat_1 \\ (32 \times 5) \end{bmatrix} \begin{bmatrix} 2^{nd} \\ Order \\ SHC \\ 27 \end{bmatrix}.$$

Знову ж, за умови наявності п'яти SHC 511 другого порядку, кожний з векторів SHC 511' і 511 має розмір п'ять у вищенаведеному рівнянні. Рівняння, що залишилися, для інших порядків, тобто третього і четвертого порядків, можуть бути аналогічні описаними вище, додержуючись одного і того ж шаблону відносно розмірів матриць (у тому, що кількість рядків $EncMat_2$, кількість стовпців $InvMat_1$ і розміри векторів SHC 511 і SHC 511' третього і четвертого порядку дорівнюють кількості підпорядків (m помножити на два плюс 1) кожної зі сферичних гармонічних базисних функцій третього і четвертого порядку.

[1308] Пристрій 570 аудіокодування, таким чином, може здійснювати цю операцію повороту відносно кожної комбінації азимутального кута і кута підвищення в спробі ідентифікації так названого оптимального повороту. Пристрій 570 аудіокодування може, після здійснення цієї операції повороту, обчислювати кількість SHC 511', що перевищує порогове значення. У ряді випадків, пристрій 570 аудіокодування може здійснювати цей поворот для виведення послідовності SHC 511', що представляють звукове поле протягом проміжку часу, наприклад кадру аудіо. Шляхом здійснення цього повороту для виведення послідовності SHC 511', що представляють звукове поле протягом цього проміжку часу, пристрій 570 аудіокодування може знижувати кількість операцій повороту, що підлягають здійсненню, замість того, щоб робити це для кожного набору SHC 511, що описують звукове поле протягом проміжків часу, менших, ніж кадр або інша довжина. У будь-якому випадку, пристрій 570 аудіокодування може зберігати,

протягом цього процесу, SHC 511', що мають найменшу кількість SHC 511', що перевищує порогове значення.

[1309] Однак здійснення цієї операції повороту відносно кожної комбінації азимутального кута і кута підвищення може вимагати великого обсягу ресурсів процесора або великого часу. У результаті, пристрій 570 аудіокодування може не здійснювати так названу реалізацію "грубої сили" алгоритму повороту. Замість цього, пристрій 570 аудіокодування може здійснювати повороти відносно піднабору, можливо, відомих (статистично) комбінацій азимутального кута і кута підвищення, які забезпечують, у загальному випадку, гарне ущільнення, здійснюючи додаткові повороти відносно комбінацій навколо тих з цього піднабору, які забезпечують краще ущільнення в порівнянні з іншими комбінаціями в піднаборі.

[1310] У порядку іншої альтернативи, пристрій 570 аудіокодування може здійснювати цей поворот відносно тільки відомого піднабору комбінацій. У порядку іншої альтернативи, пристрій 570 аудіокодування може додержуватися траєкторії (просторово) комбінацій, здійснюючи повороти відносно цієї траєкторії комбінацій. У порядку іншої альтернативи, пристрій 570 аудіокодування може вказувати поріг ущільнення, що задає максимальну кількість SHC 511', які мають ненульові значення, що перевищують порогове значення. Цей поріг ущільнення може ефективно встановлювати точку зупинки на пошук, таким чином, що, коли пристрій 570 аудіокодування здійснює поворот і визначає, що кількість SHC 511', які мають значення, що перевищує встановлений поріг, менше або дорівнює (або, у ряді випадків, менше) порогу ущільнення, пристрій 570 аудіокодування зупиняє здійснення будь-яких додаткових операцій повороту відносно комбінацій, що залишилися. У порядку ще однієї альтернативи, пристрій 570 аудіокодування може обходити ієрархічно побудоване дерево (або іншу структуру даних) комбінацій, здійснювати операції повороту відносно поточної комбінації й обходити дерево вправо або вліво (наприклад, для двійкових дерев) залежно від кількості SHC 511', які мають ненульове значення, що перевищує порогове значення.

[1311] У цьому розумінні, кожна з цих альтернатив передбачає здійснення першої і другої операцій повороту і порівняння результату здійснення першої і другої операцій повороту для ідентифікації однієї з першої і другої операцій повороту, що дають найменшу кількість SHC 511', які мають ненульове значення, що перевищує порогове значення. Відповідно, пристрій 570 аудіокодування може здійснювати першу операцію повороту на звуковому полі для повороту звукового поля відповідно до першого азимутального кута і першого кута підвищення і визначати першу кількість з множини ієрархічних елементів, що представляють звукове поле, повернене відповідно до першого азимутального кута і першого кута підвищення, що забезпечують інформацію, яка має відношення до опису звукового поля. Пристрій 570 аудіокодування також може здійснювати другу операцію повороту на звуковому полі для повороту звукового поля згідно з другим азимутальним кутом і другим кутом підвищення і визначати другу кількість з множини ієрархічних елементів, що представляють звукове поле, повернене згідно з другим азимутальним кутом і другим кутом підвищення, що забезпечують інформацію, яка має відношення до опису звукового поля. Крім того, пристрій 570 аудіокодування може вибирати першу операцію повороту або другу операцію повороту на основі порівняння першої кількості з множини ієрархічних елементів і другої кількості з множини ієрархічних елементів.

[1312] У ряді випадків, алгоритм повороту можна здійснювати відносно проміжку часу, де наступні виклики алгоритму повороту можуть здійснювати операції повороту на основі попередніх викликів алгоритму повороту. Іншими словами, алгоритм повороту може бути адаптивним на основі попередньої інформації повороту, визначеної при повороті звукового поля протягом попереднього проміжку часу. Наприклад, пристрій 570 аудіокодування може повертати звукове поле протягом першого проміжку часу, наприклад кадру аудіо, для ідентифікації SHC 511' для цього першого проміжку часу. Пристрій 570 аудіокодування може вказувати інформацію повороту і SHC 511' у бітовому потоці 517 будь-яким з вищеописаних способів. Ця інформація повороту може іменуватися першою інформацією повороту в тому, що вона описує поворот звукового поля протягом першого проміжку часу. Потім пристрій 570 аудіокодування може, на основі цієї першої інформації повороту, повертати звукове поле протягом другого проміжку часу, наприклад другого кадру аудіо, для ідентифікації SHC 511' для цього другого проміжку часу. Пристрій 570 аудіокодування може використовувати цю першу інформацію повороту при здійсненні другої операції повороту протягом другого проміжку часу для ініціалізації пошуку "оптимальної" комбінації азимутальних кутів і кутів підвищення, у порядку одного прикладу. Потім пристрій 570 аудіокодування може вказувати SHC 511' і відповідну інформацію повороту протягом другого проміжку часу (яка може іменуватися "другою інформацією повороту") у бітовому потоці 517.

[1313] Хоча вони описані вище відносно декількох різних способів реалізації алгоритму повороту для зниження часу і/або ресурсів обробки, методи можна здійснювати відносно будь-якого алгоритму, який може знижувати або інакше прискорювати ідентифікацію так названого "оптимального повороту". Крім того, методи можна здійснювати відносно будь-якого алгоритму, який ідентифікує неоптимальні повороти, але які можуть поліпшувати продуктивність в інших аспектах, часто вимірюваних відносно швидкості або процесора або використання інших ресурсів.

[1314] На фіг. 57-57E показані схеми, що демонструють бітові потоки 517A-517E, сформовані відповідно до методів, описаних в цьому винаході. У прикладі, наведеному на фіг. 57A, бітовий потік 517A може представляти один приклад бітового потоку 517, показаного вище на фіг. 53. Бітовий потік 517A включає в себе поле 670 присутності SHC і поле, де зберігаються SHC 511' (де поле позначається "SHC 511"). Поле 670 присутності SHC може включати в себе біт, що відповідає кожному з SHC 511. SHC 511' може представляти ті з SHC 511, які задаються в бітовому потоці, яких може бути менше, ніж кількість SHC 511. Звичайно кожний з SHC 511' є тим з SHC 511, що мають ненульові значення. Як згадано вище, для представлення четвертого порядку будь-якого даного звукового поля, потрібно $(1+4)^2$ або 25 SHC. Усунення одного або більше з цих SHC і заміна SHC з нульовими значеннями одиничним бітом можуть заощаджувати 31 біт, що можуть виділятися для більш детального вираження інших ділянок звукового поля або інакше видалятися для полегшення ефективного використання ширини смуги.

[1315] У прикладі, наведеному на фіг. 57B, бітовий потік 517B може представляти один приклад бітового потоку 517, показаного вище на фіг. 53. Бітовий потік 517B включає в себе інформаційне поле 672 перетворення ("інформацію 672 перетворення") і поле, де зберігаються SHC 511' (де поле позначається "SHC 511"). Інформація 672 перетворення, як згадано вище, може містити інформацію перенесення, інформацію повороту і/або будь-яку іншу форму інформації, що позначає регулювання звукового поля. У ряді випадків, інформація 672 перетворення також може вказувати SHC 511 найвищого порядку, які задаються в бітовому потоці 517B як SHC 511'. Таким чином, інформація 672 перетворення може вказувати порядок три, котрий пристрій витягання може зрозуміти як такий, що вказує, що SHC 511' включає в себе SHC 511 аж до і включаючи SHC 511, що мають порядок три. Пристрій витягання може потім бути виконаний з можливістю обнулення SHC 511, що має порядок чотири або більше, таким чином, потенційно усуваючи явну сигналізацію SHC 511 порядку чотири або більш у бітовому потоці.

[1316] У прикладі, наведеному на фіг. 57C, бітовий потік 517C може представляти один приклад бітового потоку 517, показаного вище на фіг. 53. Бітовий потік 517C включає в себе інформаційне поле 672 перетворення ("інформацію 672 перетворення"), поле 670 присутності SHC і поле, де зберігаються SHC 511' (де поле позначається "SHC 511"). Замість того, щоб бути виконаним з можливістю розуміти, SHC 511 якого порядку не сигналізуються, як описано вище з посиланням на фіг. 57B, поле 670 присутності SHC може явно сигналізувати, які з SHC 511 задані в бітовому потоці 517C як SHC 511'.

[1317] У прикладі, наведеному на фіг. 57D, бітовий потік 517D може представляти один приклад бітового потоку 517, показаного вище на фіг. 53. Бітовий потік 517D включає в себе поле 674 порядку ("порядок 60"), поле 670 присутності SHC, прапор 676 азимута ("AZF 676"), прапор 678 підвищення ("ELF 678"), поле 680 азимутального кута ("азимут 680"), поле 682 кута підвищення ("підвищення 682") і поле, де зберігаються SHC 511' (де, знову ж, поле позначається "SHC 511"). Поле 674 порядку вказує порядок SHC 511', тобто порядок, позначений n вище для найвищого порядку сферичної базисної функції, використовуваної для представлення звукового поля. Поле 674 порядку показане як 8-бітове поле, але може мати різні інші бітові розміри, наприклад три (які являють собою кількість бітів, необхідну для вказання четвертого порядку). Поле 670 присутності SHC показане як 25-бітове поле. Знову ж, однак, поле 670 присутності SHC може мати різні інші бітові розміри. Поле 670 присутності SHC показане як 25-бітове для вказання, що поле 670 присутності SHC може включати в себе один біт для кожного з коефіцієнтів сферичних гармонік, що відповідають представленню четвертого порядку звукового поля.

[1318] Прапор 676 азимута представляє однобітовий прапор, що вказує, чи присутнє поле 680 азимута в бітовому потоці 517D. Коли прапор 676 азимута задається рівним одиниці, поле 680 азимута для SHC 511' присутнє у бітовому потоці 517D. Коли прапор 676 азимута задається рівним нулю, поле 680 азимута для SHC 511' відсутнє або інакше не зазначене у бітовому потоці 517D. Аналогічно, прапор 678 підвищення представляє однобітовий прапор, що вказує, чи присутнє поле 682 підвищення в бітовому потоці 517D. Коли прапор 678 підвищення

задається рівним одиниці, поле 682 підвищення для SHC 511' присутнє у бітовому потоці 517D. Коли прапор 678 підвищення задається рівним нулю, поле 682 підвищення для SHC 511' відсутнє або інакше не зазначене у бітовому потоці 517D. Хоча описано, що одиниця сигналізує, що відповідне поле присутнє, і нуль сигналізує, що відповідне поле відсутнє, угода може бути переглянута таким чином, що нуль указує, що відповідне поле вказується в бітовому потоці 517D, і одиниця вказує, що відповідне поле не вказується в бітовому потоці 517D. Таким чином, методи, описані в цьому винаході, не підлягають обмеженню в цьому відношенні.

[1319] Поле 680 азимута представляє 10-бітове поле, що вказує, при наявності в бітовому потоці 517D, азимутальний кут. Хоча воно показано як 10-бітове поле, поле 680 азимута може мати інші бітові розміри. Поле 682 підвищення представляє 9-бітове поле, що вказує, при наявності в бітовому потоці 517D, кут підвищення. Азимутальний кут і кут підвищення, зазначені у полях 680 і 682, відповідно, можуть, разом із прапорами 676 і 678, представляти вищеописану інформацію повороту. Цю інформацію повороту можна використовувати для повороту звукового поля для відновлення SHC 511 у вихідній системі координат.

[1320] Поле SHC 511' показано як змінне поле, що має розмір X. Поле SHC 511' може змінюватися внаслідок кількості SHC 511', зазначеної в бітовому потоці, що позначено полем 670 присутності SHC. Розмір X можна виводити як функцію кількості одиниць у полі 670 присутності SHC, помноженої на 32 біти (що є розміром кожного SHC 511').

[1321] У прикладі, наведеному на фіг. 57E, бітовий потік 517E може представляти інший приклад бітового потоку 517, показаного вище на фіг. 53. Бітовий потік 517E включає в себе поле 674 порядку ("порядок 60"), поле 670 присутності SHC і поле 684 індексу повороту, і поле, де зберігаються SHC 511' (де, знову ж, поле позначається "SHC 511"). Поле 674 порядку, поле 670 присутності SHC і поле SHC 511' можуть бути, по суті, аналогічними описаному вище. Поле 684 індексу повороту може представляти 20-бітове поле, використовуване для указання однієї з комбінацій 1024×512 (або, іншими словами, 524288) кута підвищення й азимутального кута. У ряді випадків, для указання цього поля 684 індексу повороту можна використовувати тільки 19 бітів, і пристрій 570 аудіокодування може вказувати додатковий прапор у бітовому потоці для указання, чи здійснювалася операція повороту (і, таким чином, чи присутнє поле 684 індексу повороту в бітовому потоці). Це поле 684 індексу повороту указує вищезгаданий індекс повороту, який може означати елемент у таблиці повороту, загальний для пристрою 570 аудіокодування і пристрою витягання бітового потоку. Ця таблиця повороту може, у ряді випадків, зберігати різні комбінації азимутальних кутів і кутів підвищення. Альтернативно, таблиця повороту може зберігати вищеописану матрицю, у якій ефективно зберігаються різні комбінації азимутальних кутів і кутів підвищення в матричній формі.

[1322] На фіг. 58 показана блок-схема операцій, що демонструє приклад роботи пристрою 570 аудіокодування, показаного в прикладі, наведеному на фіг. 53, при реалізації аспектів повороту методів, описаних у цьому винаході. Спочатку, пристрій 570 аудіокодування може вибирати комбінацію азимутального кута і кута підвищення відповідно до одного або більше з різних вищеописаних алгоритмів повороту (800). Потім пристрій 570 аудіокодування може повертати звукове поле відповідно до вибраного азимутального кута і кута підвищення (802). Як описано вище, пристрій 570 аудіокодування може спочатку виводити звукове поле з SHC 511 з використанням вищезгаданої $InvMat_1$. Пристрій 570 аудіокодування також може визначати SHC 511', що представляють повернене звукове поле (804). Хоча воно описане у вигляді окремих етапів або операцій, пристрій 570 аудіокодування може застосовувати перетворення (яке може представляти результат $[EncMat_2][InvMat_1]$), що представляє вибір комбінації азимутального кута і кута підвищення, виведення звукового поля з SHC 511, поворот звукового поля і визначення SHC 511', що представляють повернене звукове поле.

[1323] У будь-якому випадку, пристрій 570 аудіокодування потім може обчислювати кількість визначених SHC 511', що більше порогового значення, порівнювати цю кількість з кількістю, обчисленою для попередньої ітерації, відносно попередньої комбінації азимутального кута і кута підвищення (806, 808). У першій ітерації відносно першої комбінації азимутального кута і кута підвищення, це порівняння може виконуватися з попередньо заданою попередньою кількістю (яка може бути задана рівною нулю). У будь-якому випадку, якщо визначена кількість SHC 511' менше, ніж попередня кількість ("Так" 808), пристрій 570 аудіокодування зберігає SHC 511', азимутальний кут і кут підвищення, часто заміняючи попередній SHC 511', азимутальний кут і кут підвищення збережені з попередньої ітерації алгоритму повороту (810).

[1324] Якщо визначена кількість SHC 511' не менше, ніж попередня кількість ("Ні" 808), або після збереження SHC 511', азимутального кута і кута підвищення замість раніше збережених SHC 511', азимутального кута і кута підвищення, пристрій 570 аудіокодування може визначати, чи закінчений алгоритм повороту (812). Таким чином, пристрій 570 аудіокодування може, у

порядку одного прикладу, визначати, чи всі доступні комбінації азимутального кута і кута підвищення оцінені. В інших прикладах, пристрій 570 аудіокодування може визначати, чи задовольняються інші критерії (наприклад, що усі з заданого піднабору комбінацій здійснені, чи пройдена дана траєкторія, чи пройдене ієрархічне дерево до крайового вузла, і т. д.), таким чином, що пристрій 570 аудіокодування закінчив здійснення алгоритму повороту. Якщо не закінчив ("Ні" 812), пристрій 570 аудіокодування може здійснювати вищеописаний процес відносно іншої вибраної комбінації (800-812). Якщо закінчив ("Так" 812), пристрій 570 аудіокодування може вказувати збережені SHC 511', азимутальний кут і кут підвищення в бітовому потоці 517 одним з різних вищеописаних способів (814).

[1325] На фіг. 59 показана блок-схема операцій, що демонструє приклад роботи пристрою 570 аудіокодування, показаного в прикладі, наведеному на фіг. 53, при здійсненні аспектів перетворення методів, описаних у цьому винаході. Спочатку, пристрій 570 аудіокодування може вибирати матрицю, що представляє лінійне оборотне перетворення (820). Одним прикладів матриці, що представляє лінійне оборотне перетворення, може бути показана вище матриця, одержана в результаті $[EncMat_1][IncMat_1]$. Пристрій 570 аудіокодування потім може застосовувати матрицю до звукового поля для перетворення звукового поля (822). Пристрій 570 аудіокодування також може визначати SHC 511', що представляють повернене звукове поле (824). Хоча воно описане у вигляді окремих етапів або операцій, пристрій 570 аудіокодування може застосовувати перетворення (яке може представляти результат $[EncMat_2][InvMat_1]$), виведення звукового поля з SHC 511, перетворення звукового поля і визначення SHC 511', що представляють перетворення звукового поля.

[1326] У будь-якому випадку, пристрій 570 аудіокодування потім може обчислювати кількість визначених SHC 511', що більше порогового значення, порівнювати цю кількість з кількістю, обчисленою для попередньої ітерації, відносно попереднього застосування матриці перетворення (826, 828). Якщо визначена кількість SHC 511' менше, ніж попередня кількість ("Так" 828), пристрій 570 аудіокодування зберігає SHC 511' і матрицю (або деяку їх похідну, наприклад індекс, пов'язаний з матрицею), часто замінюючи попередній SHC 511' і матрицю (або її похідну), збережені з попередньої ітерації алгоритму повороту (830).

[1327] Якщо визначена кількість SHC 511' не менше, ніж попередня кількість ("Ні" 828), або після збереження SHC 511' і матриці замість раніше збережених SHC 511' і матриці, пристрій 570 аудіокодування може визначати, чи закінчений алгоритм перетворення (832). Таким чином, пристрій 570 аудіокодування може, у порядку одного прикладу, визначати, чи всі доступні перетворення матриці оцінені. В інших прикладах, пристрій 570 аудіокодування може визначати, чи задовольняються інші критерії (наприклад, що усі з заданого піднабору доступних матриць перетворення здійснені, чи пройдена дана траєкторія, чи пройдене ієрархічне дерево до крайового вузла, і т. д.) таким чином, що пристрій 570 аудіокодування закінчив здійснення алгоритму перетворення. Якщо не закінчив ("Ні" 832), пристрій 570 аудіокодування може здійснювати вищеописаний процес відносно іншої вибраної матриці перетворення (820-832). Якщо закінчив ("Так" 832), пристрій 570 аудіокодування може вказувати збережені SHC 511' і матрицю в бітовому потоці 517 одним з різних вищеописаних способів (834).

[1328] У деяких прикладах, алгоритм перетворення може здійснювати одиничну ітерацію, оцінюючи одиничну матрицю перетворення. Таким чином, матриця перетворення може містити будь-яку матрицю, що представляє лінійне оборотне перетворення. У ряді випадків, лінійне оборотне перетворення може перетворювати звукове поле з просторової області в частотну область. Приклади такого лінійного оборотного перетворення можуть включати в себе дискретне перетворення Фур'є (DFT). Застосування DFT може передбачати тільки одиничну ітерацію і, таким чином, не буде вимагати включати в себе етапи для визначення, чи закінчений алгоритм перетворення. Відповідно, методи не підлягають обмеженню прикладом, наведеним на фіг. 59.

[1329] Іншими словами, одним прикладом лінійного оборотного перетворення є дискретне перетворення Фур'є (DFT). DFT може здійснюватися на двадцяти п'яти SHC 511' для формування набору з двадцяти п'яти комплексних коефіцієнтів. Пристрій 570 аудіокодування також може заповнювати нулями двадцять п'ять SHC 511', щоб бути цілим кратним 2, щоб потенційно збільшувати розрізнення розміру біна DFT, і потенційно мати більш ефективну реалізацію DFT, наприклад шляхом застосування швидкого перетворення Фур'є (FFT). У ряді випадків, збільшення розрізнення DFT крім 25 точок не завжди потрібне. У перетворюючій області, пристрій 570 аудіокодування може застосовувати поріг для визначення, чи існує яка-небудь спектральна енергія в конкретному біні. Пристрій 570 аудіокодування, у цьому контексті, потім може відкидати або обнулювати енергію спектрального коефіцієнта, яка нижче цього порога, і пристрій 570 аудіокодування може застосовувати обернене перетворення для

відновлення SHC 511', після відкидання або обнулення одного або більше з SHC 511'. Таким чином, після застосування оберненого перетворення, коефіцієнти нижче порога не присутні, і, у результаті, для кодування звукового поля можна використовувати менше бітів.

[1330] В одному або більше прикладах, описані функції можуть бути реалізовані в устаткуванні, програмному забезпеченні, програмно-апаратному забезпеченні або будь-якій їх комбінації. Будучи реалізовані в програмному забезпеченні, функції можуть зберігатися або передаватися у вигляді однієї або більше інструкцій або коду на зчитуваному комп'ютером носії і виконуватися апаратним блоком обробки. Зчитувані комп'ютером Носії можуть включати в себе зчитувані комп'ютером носії даних, які відповідають матеріальному носію, наприклад носіям даних або середовищам передачі даних, що включають в себе будь-яке середовище, яке полегшує перенесення комп'ютерної програми з одного місця в інше, наприклад, відповідно до протоколу зв'язку. Таким чином, зчитувані комп'ютером носії, у загальному випадку, можуть відповідати (1) матеріальним зчитуванням комп'ютером носіям даних, що є довгостроковими, або (2) середовищу передачі даних, наприклад сигналу або несущій хвилі. Носіями даних можуть бути будь-які доступні носії, до яких можуть звертатися один або більше комп'ютерів або один або більше процесорів для витягання інструкцій, коду і/або структур даних для реалізації методів, описаних у цьому винаході. Комп'ютерний програмний продукт може включати в себе зчитуваний комп'ютером носій.

[1331] У порядку прикладу, але не обмеження, такі зчитувані комп'ютером носії даних можуть містити ОЗП, ПЗП, ЕСППЗП, CD-ROM або інше сховище на основі оптичного диска, сховище на основі магнітного диска або інші магнітні запам'ятовуючі пристрої, флеш-пам'ять або будь-який інший носій, який можна використовувати для збереження бажаного програмного коду у формі інструкцій або структур даних, до яких може звертатися комп'ютер. Крім того, будь-яке з'єднання доречно іменувати зчитуванням комп'ютером середовищем. Наприклад, якщо інструкції передаються з веб-сайта, сервера або іншого віддаленого джерела з використанням коаксіального кабелю, оптоволоконного кабелю, виті пари, цифрової абонентської лінії (DSL) або бездротових технологій, наприклад інфрачервоної, радіо- і мікрохвильової, то коаксіальний кабель, оптоволоконний кабель, вита пара, DSL або бездротові технології, наприклад інфрачервона, радіо- і мікрохвильова, включаються у визначення носія. Однак варто розуміти, що зчитувані комп'ютером носії даних і носії даних не включають в себе з'єднання, несучі хвилі, сигнали або інші транзиторні носії, але, замість цього, належать до довгострокових, матеріальних носіїв даних. Використовуваний тут термін диск включає в себе компакт-диск (CD), лазерний диск, оптичний диск, цифровий універсальний диск (DVD), флопі-диск і диск Blu-ray, де диски звичайно відтворюють дані магнітними засобами, тоді як диски відтворюють дані оптичними засобами за допомогою лазерів. Комбінації вищеописаних пристроїв також підлягають включенню в обсяг визначення машиночитаних носіїв.

[1332] Інструкції можуть виконуватися одним або більше процесорами, наприклад одним або більше цифровими сигнальними процесорами (DSP), мікропроцесорами загального призначення, спеціалізованими інтегральними схемами (ASIC), програмованими користувачем вентильними матрицями (FPGA) або іншими еквівалентними інтегральними або дискретними логічними схемами. Відповідно, використовуваний тут термін "процесор" може означати будь-яку з вищенаведених структур або будь-яку іншу структуру, придатну для реалізації описаних тут методів. Крім того, у деяких аспектах, описані тут функціональні можливості можуть забезпечуватися в спеціалізованому устаткуванні і/або програмних модулях, сконфігурованих для кодування і декодування, або впроваджені в комбінований кодек. Також, методи можуть бути повністю реалізовані в одній або більше схемах або логічних елементах.

[1333] Методи цього винаходу можна реалізувати в різноманітних пристроях або апаратах, що включають у себе бездротову телефонну трубку, інтегральну схему (IC) або набір IC (наприклад, чипсет). Різні компоненти, модулі або блоки описані в цьому винаході, щоб підкреслити функціональні аспекти пристроїв, виконаних з можливістю здійснення розкритих методів, але не обов'язково вимагають реалізації різними апаратними блоками. Навпаки, як описано вище, різні блоки можуть бути об'єднані в апаратному блоці кодека або забезпечені сукупністю взаємодіючих апаратних блоків, що включають у себе один або більше процесорів, як описано вище, разом із придатним програмним забезпеченням і/або програмно-апаратним забезпеченням.

[1334] Були описані різні варіанти здійснення методів. Ці й інші аспекти методів входять в обсяг нижченаведеної формули винаходу.

ФОРМУЛА ВИНАХОДУ

1. Спосіб, який включає етап, на якому:
- 5 одержують бітовий потік, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового
- 10 компонента.
2. Спосіб за п. 1, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.
3. Спосіб за п. 1, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує значення, яке
- 15 виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.
4. Спосіб за п. 3, у якому значення містить значення nbits.
5. Спосіб за п. 3, у якому бітовий потік містить стиснену версію з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і
- 20 при цьому значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.
6. Спосіб за п. 1, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення
- 25 ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.
7. Спосіб за п. 1, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.
- 30 8. Спосіб за п. 1, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення залишкового значення просторового компонента.
9. Спосіб за п. 1, у якому одержання бітового потоку включає генерацію бітового потоку за допомогою пристрою генерації бітового потоку.
- 35 10. Спосіб за п. 1, у якому одержання бітового потоку включає одержання бітового потоку за допомогою пристрою витягання бітового потоку.
11. Спосіб за п. 1, у якому розкладання включає розкладання на сингулярні значення.
12. Пристрій, який містить:
- 40 один або більше процесорів, виконаних з можливістю одержання бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента.
- 45 13. Пристрій за п. 12, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.
14. Пристрій за п. 12, у якому стиснена версія просторового компонента додатково
- 50 представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.
15. Пристрій за п. 14, у якому значення містить значення nbits.
16. Пристрій за п. 14, у якому бітовий потік містить стиснену версію з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і
- 55 при цьому значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.
17. Пристрій за п. 12, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для

представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.

18. Пристрій за п. 12, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

19. Пристрій за п. 12, у якому стиснена версія просторового компонента додатково представлена в бітовому потоці з використанням, щонайменше частково, коду Хаффмана для представлення залишкового значення просторового компонента.

20. Пристрій за п. 12, у якому пристрій містить пристрій аудіокодування і пристрій генерації бітового потоку.

21. Пристрій за п. 12, у якому пристрій містить пристрій аудіодекодування.

22. Пристрій за п. 12, у якому розкладання включає розкладання на сингулярні значення.

23. Пристрій, який містить:

засіб для одержання бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента; і

засіб для збереження бітового потоку.

24. Довгостроковий зчитуваний комп'ютером носій даних, де зберігаються інструкції, які, при виконанні, призначають одному або більше процесорам одержувати бітовий потік, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік.

25. Спосіб, який включає етап, на якому:

генерують бітовий потік, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента.

26. Спосіб за п. 25, у якому генерація бітового потоку включає генерацію бітового потоку для включення інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.

27. Спосіб за п. 25, у якому генерація бітового потоку включає генерацію бітового потоку для включення поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.

28. Спосіб за п. 27, у якому значення містить значення *nbits*.

29. Спосіб за п. 27, у якому генерація бітового потоку включає генерацію бітового потоку для включення стисненої версії з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і при цьому значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.

30. Спосіб за п. 25, у якому генерація бітового потоку включає генерацію бітового потоку для включення коду Хаффмана для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.

31. Спосіб за п. 25, у якому генерація бітового потоку включає генерацію бітового потоку для включення знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

32. Спосіб за п. 25, у якому генерація бітового потоку включає генерацію бітового потоку для включення коду Хаффмана для представлення залишкового значення просторового компонента.

33. Спосіб за п. 25, у якому розкладання включає розкладання на сингулярні значення.

34. Пристрій, який містить:

один або більше процесорів, виконаних з можливістю генерації бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік, у якому стиснена версія просторового компонента

представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента.

35. Пристрій за п. 34, у якому один або більше процесорів виконані з можливістю генерації бітового потоку для включення інформації таблиці Хаффмана, що вказує таблицю Хаффмана, використовувану при стисненні просторового компонента.

36. Пристрій за п. 34, у якому один або більше процесорів виконані з можливістю генерації бітового потоку для включення поля, що вказує значення, яке виражає розмір кроку квантування або його змінну, використовувану при стисненні просторового компонента.

37. Пристрій за п. 36, у якому значення містить значення $nbits$.

38. Пристрій за п. 36, у якому один або більше процесорів виконані з можливістю генерації бітового потоку для включення стисненої версії з множини просторових компонентів звукового поля, стиснена версія просторового компонента якого включена, і при цьому значення виражає розмір кроку квантування або його змінну, використовувану при стисненні множини просторових компонентів.

39. Пристрій за п. 34, у якому один або більше процесорів виконані з можливістю генерації бітового потоку для включення коду Хаффмана для представлення ідентифікатора категорії, що ідентифікує категорію стиснення, якій відповідає просторовий компонент.

40. Пристрій за п. 34, у якому один або більше процесорів виконані з можливістю генерації бітового потоку для включення знакового біта, що ідентифікує, чи має просторовий компонент позитивне значення або негативне значення.

41. Пристрій за п. 34, у якому один або більше процесорів виконані з можливістю генерації бітового потоку для включення коду Хаффмана для представлення залишкового значення просторового компонента.

42. Пристрій за п. 34, у якому розкладання включає розкладання на сингулярні значення.

43. Пристрій, який містить:

засіб для генерації бітового потоку, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента; і засіб для збереження бітового потоку.

44. Довгостроковий зчитуваний комп'ютером носій даних, що містить інструкції, які, при виконанні, призначають одному або більше процесорам:

генерувати бітовий потік, що містить стиснену версію просторового компонента звукового поля, причому просторовий компонент заданий в області сферичних гармонік і генерується шляхом здійснення розкладання відносно множини коефіцієнтів сферичних гармонік, у якому стиснена версія просторового компонента представлена в бітовому потоці з використанням, щонайменше частково, поля, що вказує режим прогнозування, використовуваний при стисненні просторового компонента.

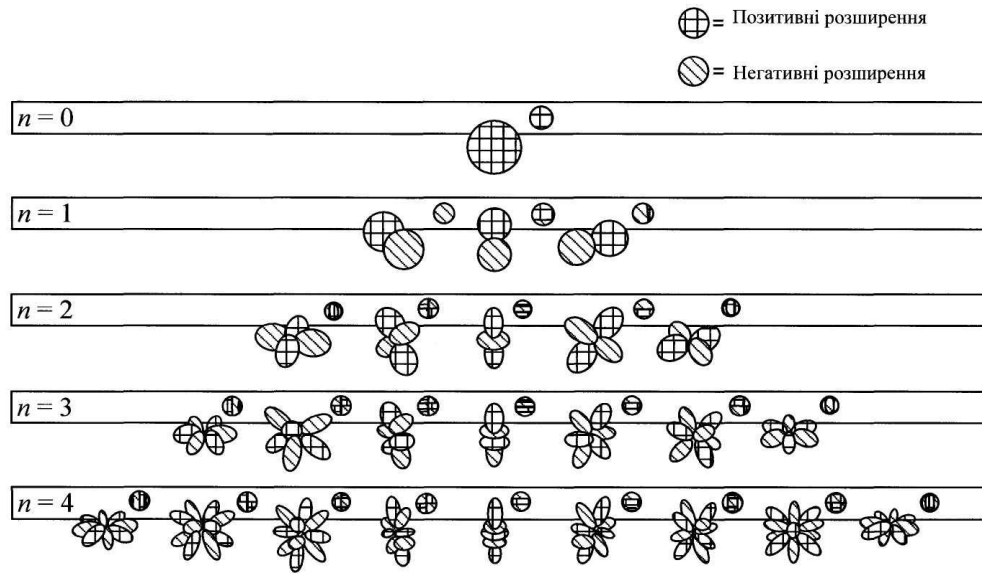


Fig. 1

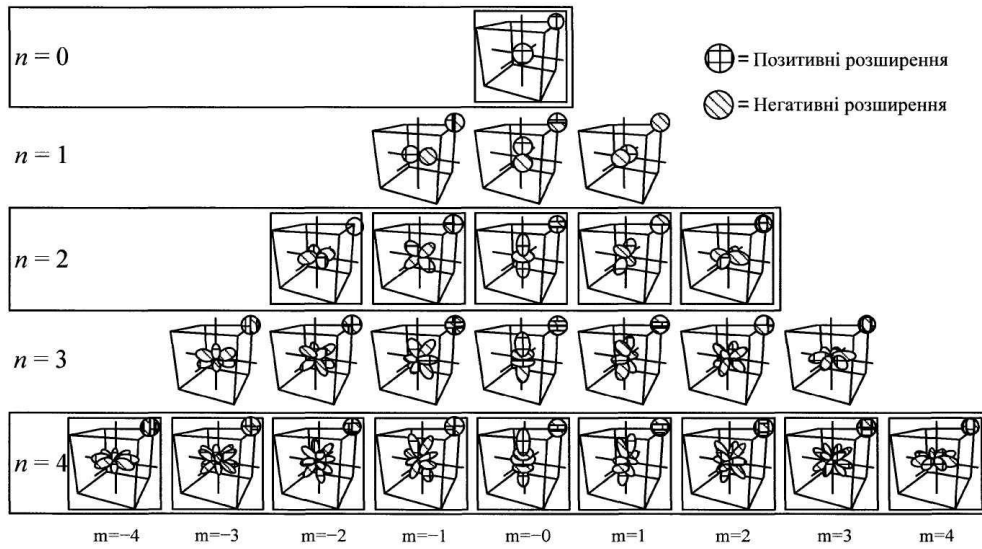


Fig. 2

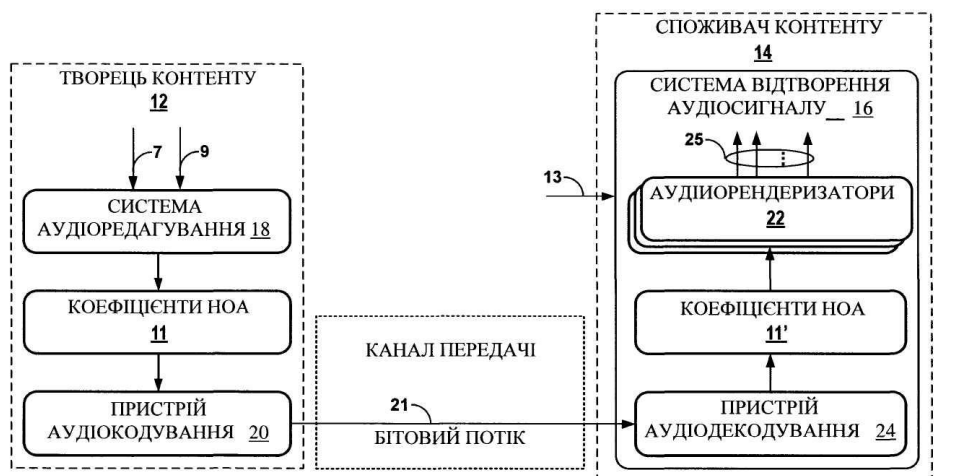


Fig. 3

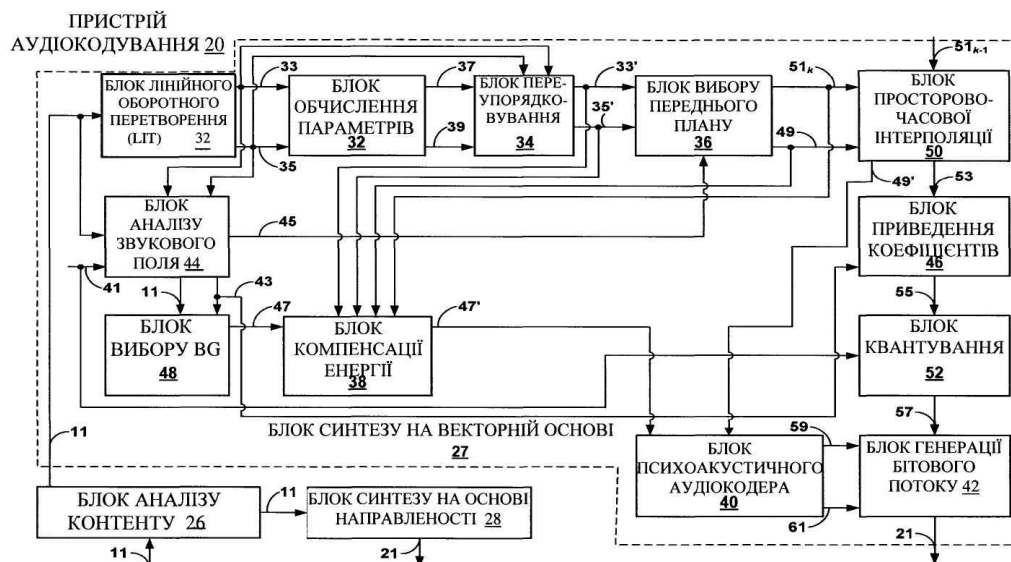


Fig. 4

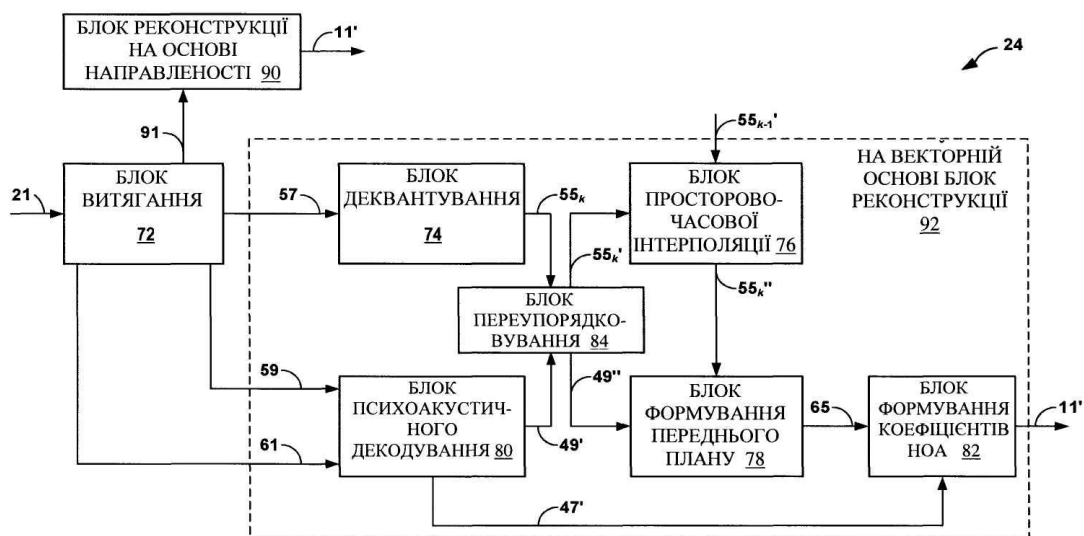
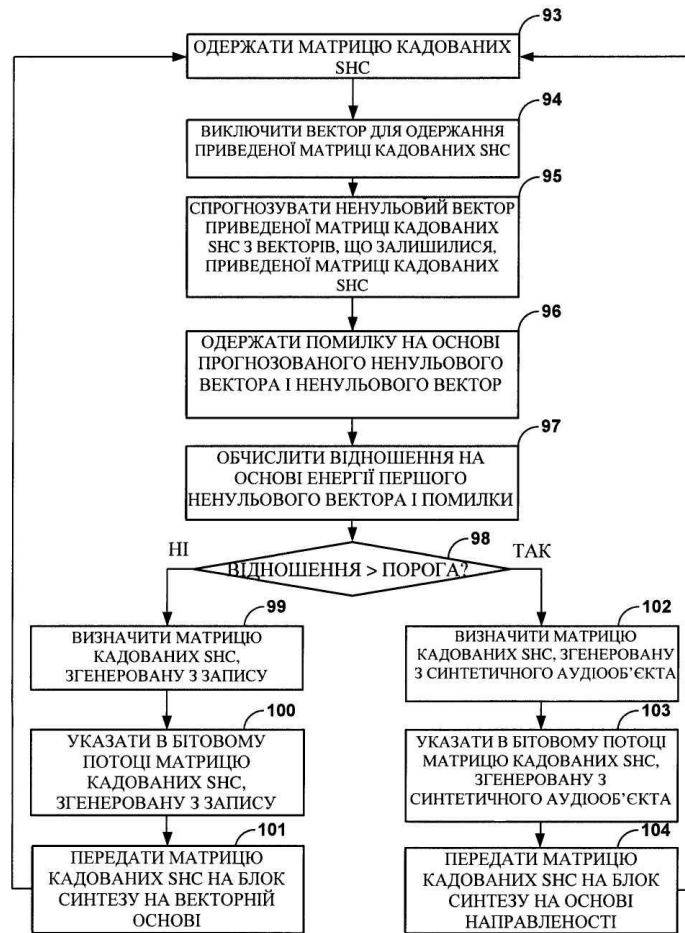
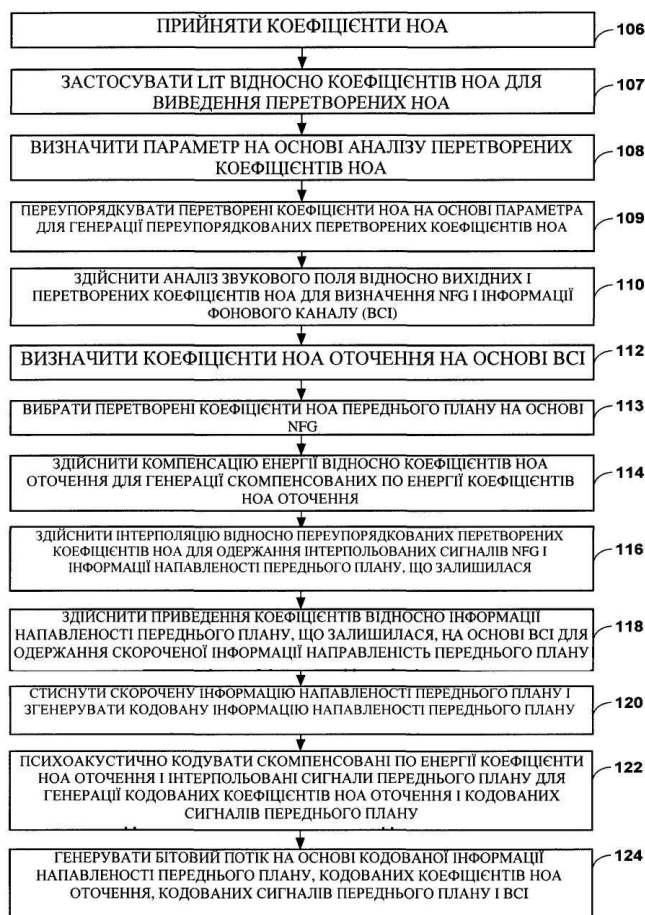


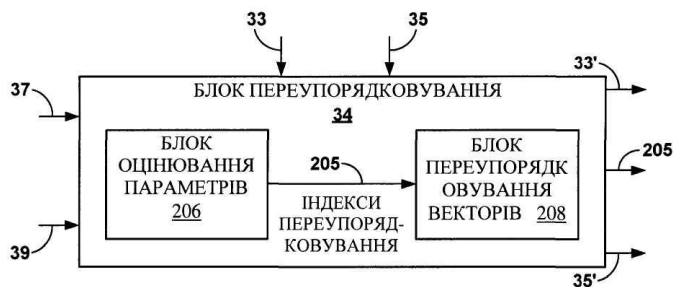
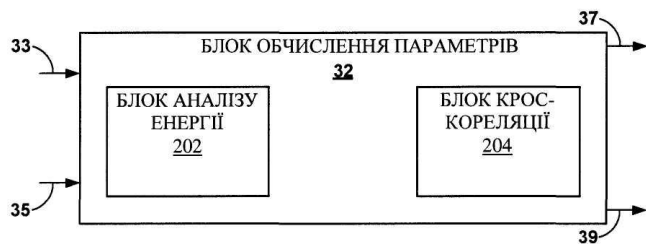
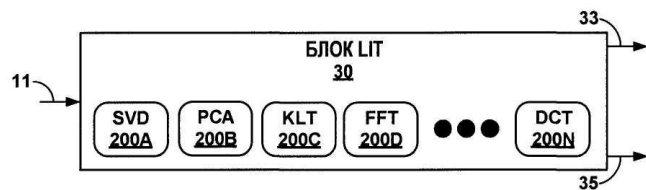
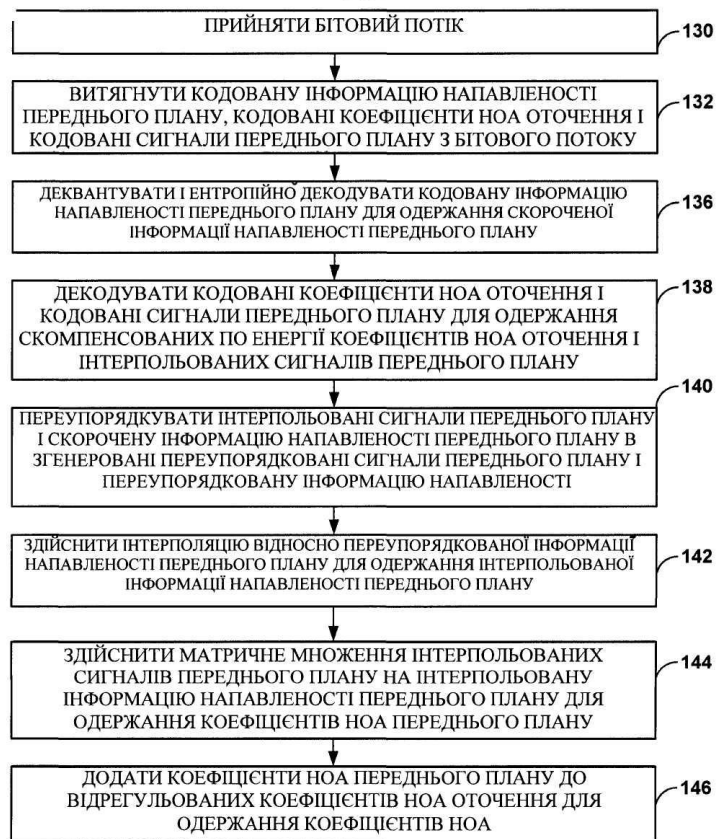
Fig. 5



Фиг. 6



Фіг. 7



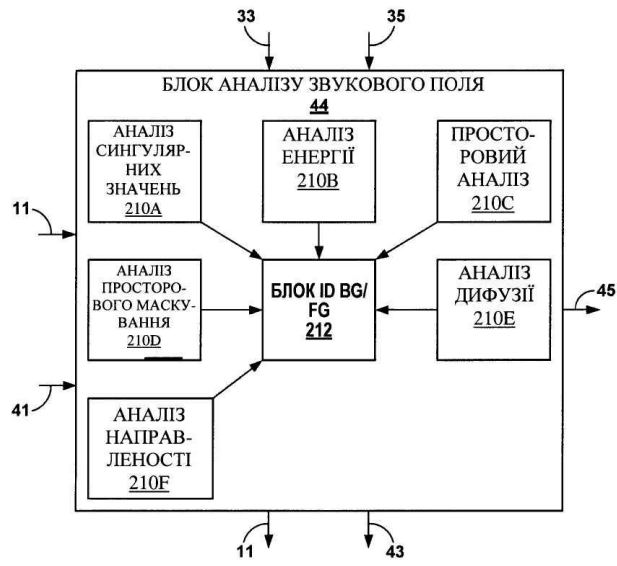


Fig. 9D

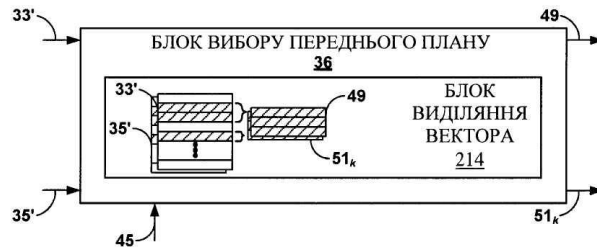
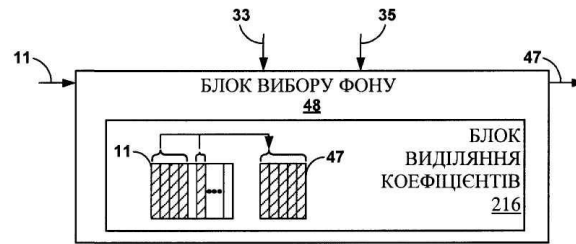
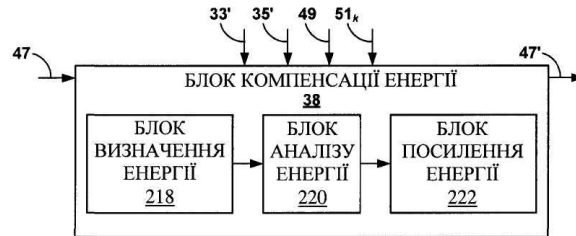


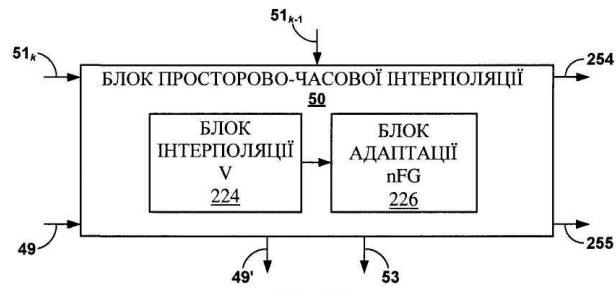
Fig. 9E



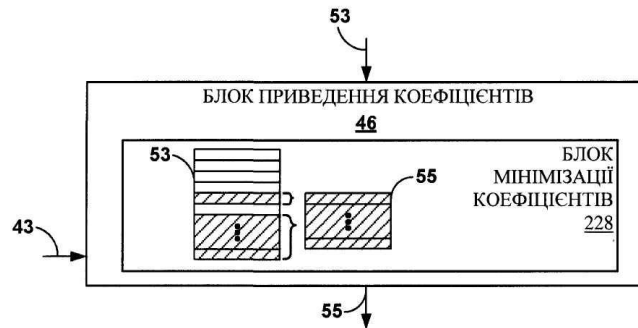
Фіг. 9F



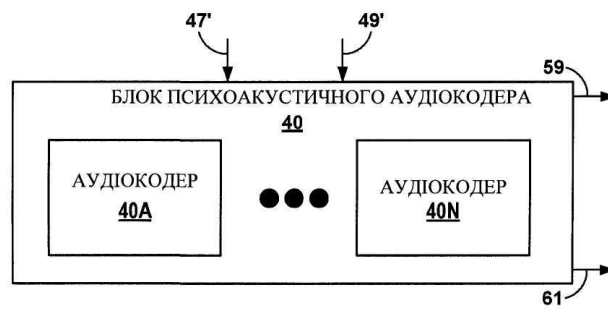
Фіг. 9G



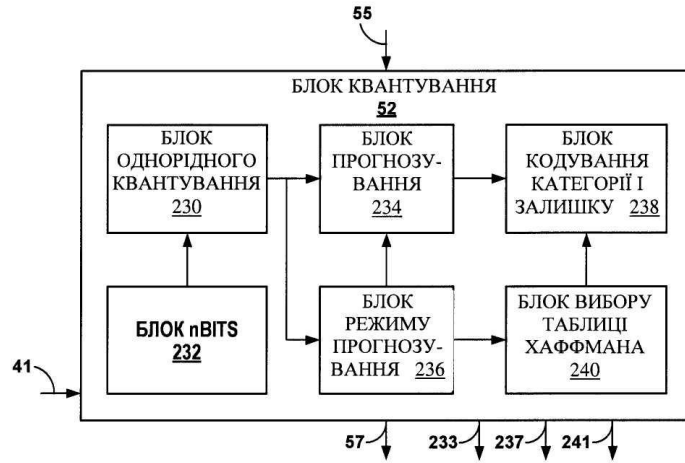
Фіг. 9H



Фіг. 9I



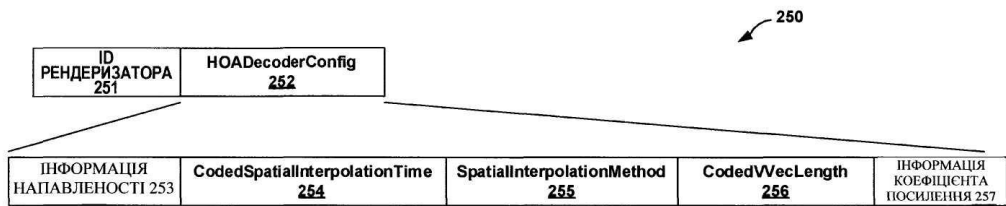
Фіг. 9J



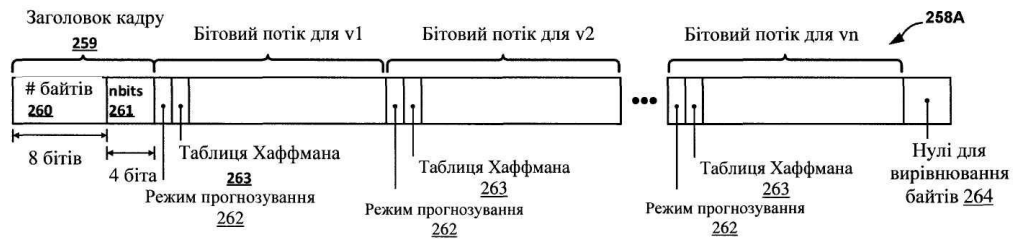
Фиг. 9К



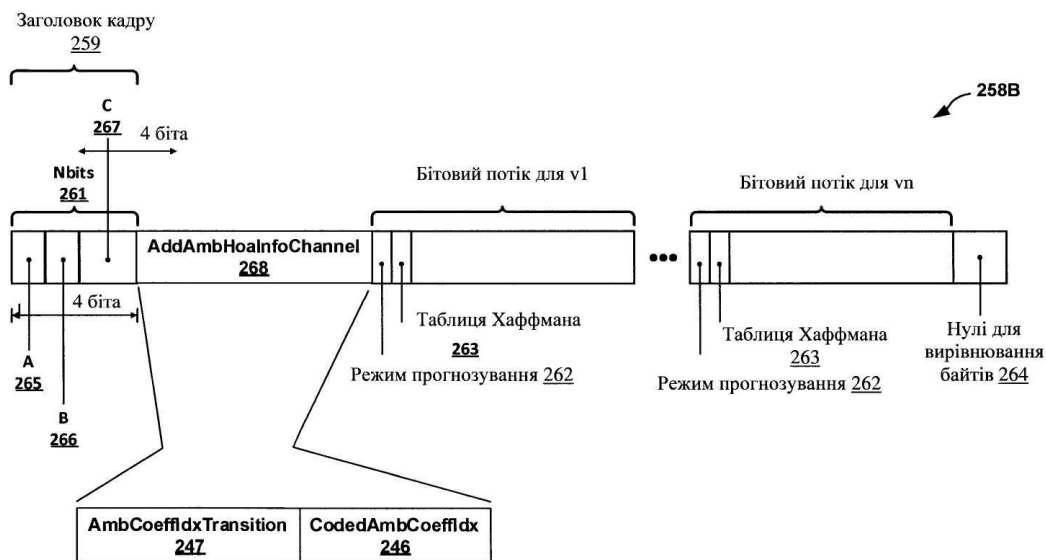
Фиг. 9L



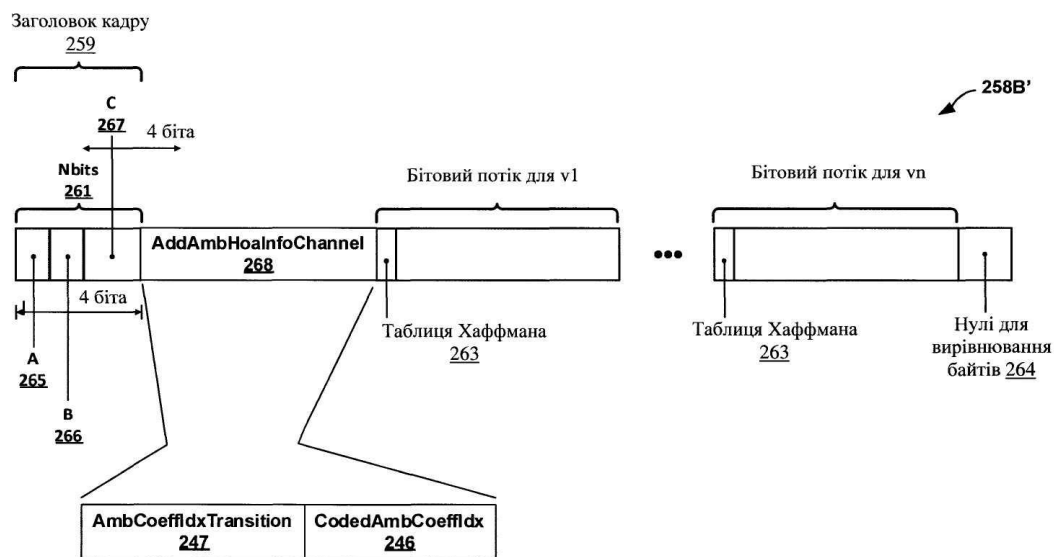
Фиг. 10A



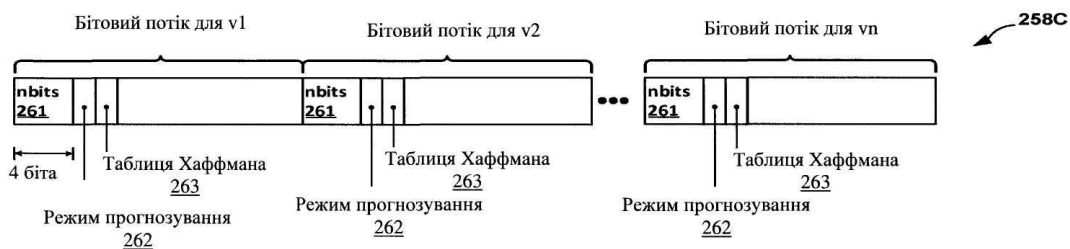
Фиг. 10B



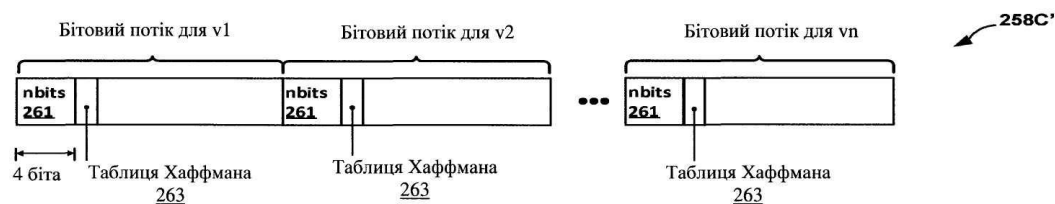
Фіг. 10C



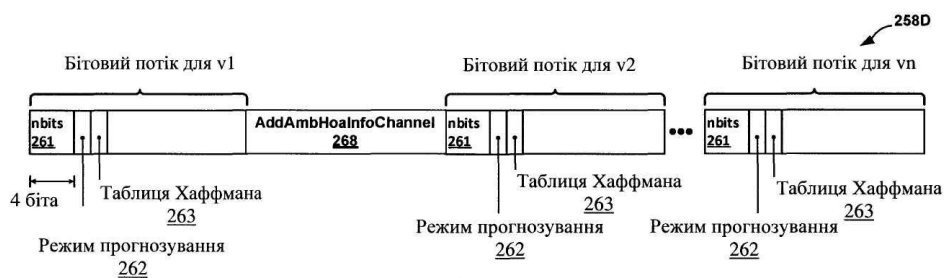
Фіг. 10C(i)



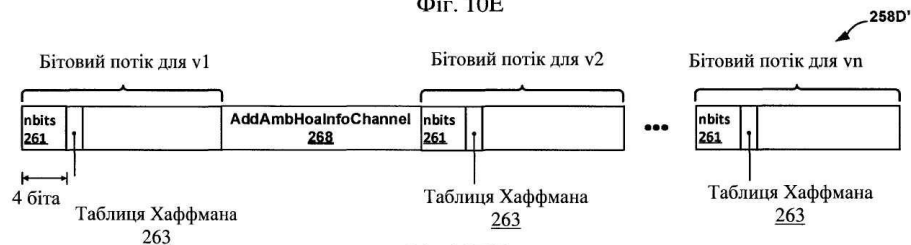
Фіг. 10D



Фіг. 10D(i)



Фіг. 10E

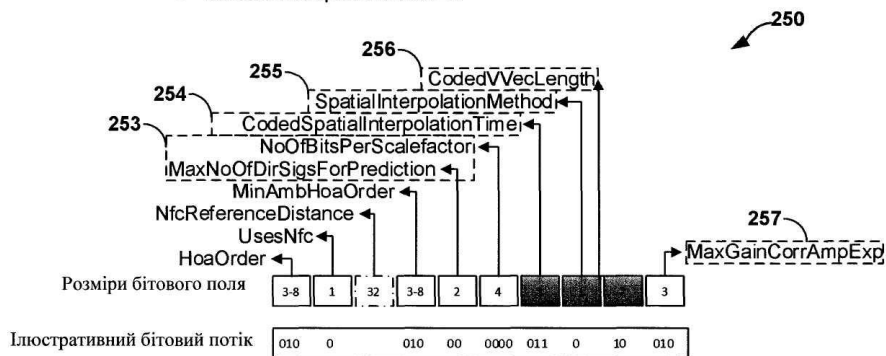


Фіг. 10E(i)

Бітовий потік HOAConfig

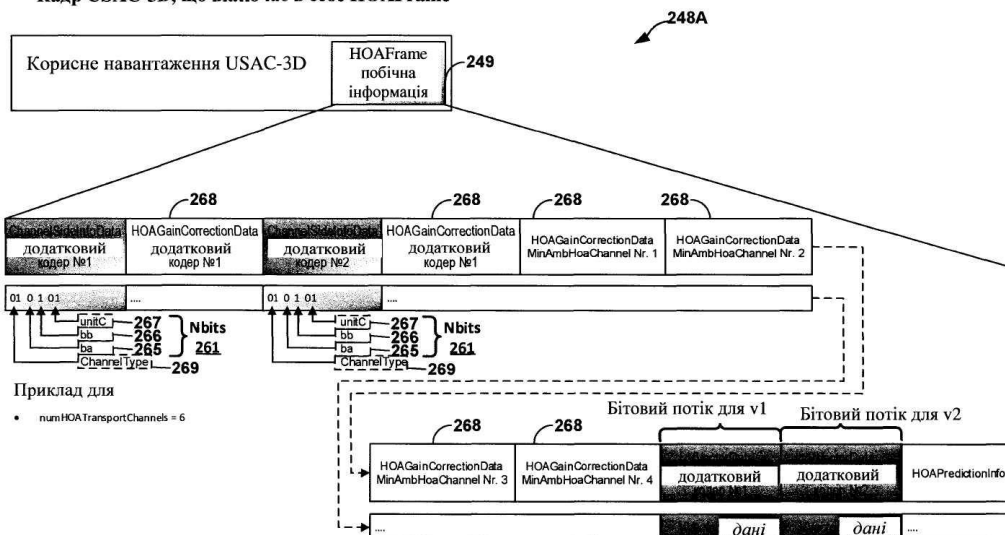
Приклад для

- numHOATransportChannels = 6

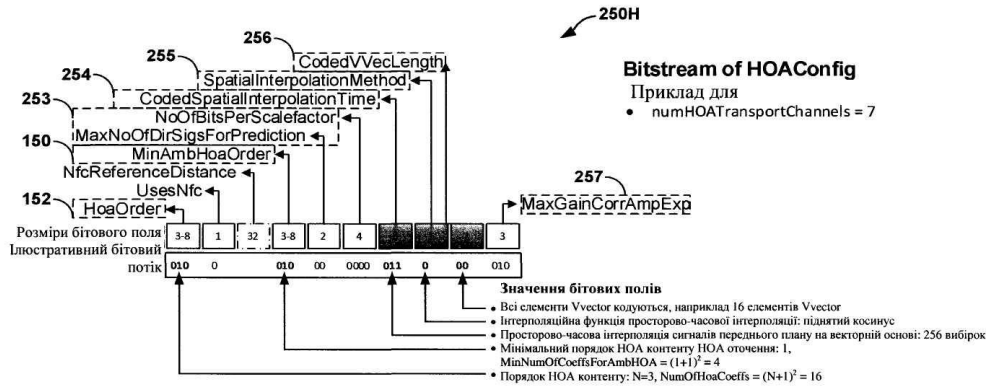


Фіг. 10F

Кадр USAC-3D, що включає в себе HOAFrame

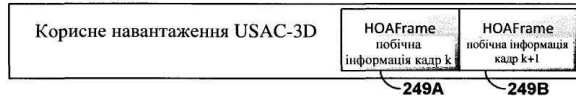


Фіг. 10G



Кадр USAC-3D, що включає в себе HOAFrame

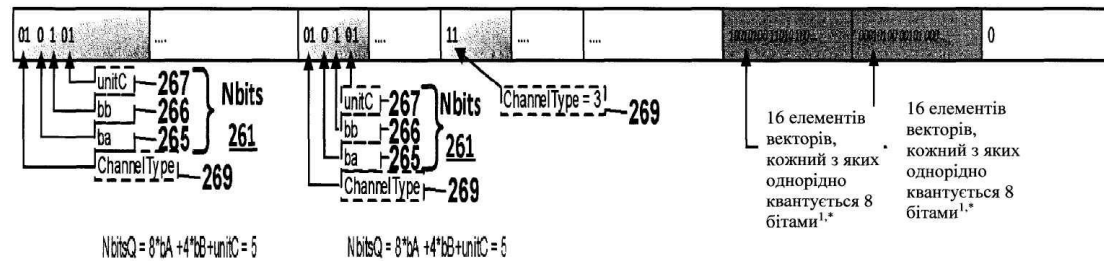
2 послідовні HOAFrame зберігаються в розширенні USAC корисного навантаження, оскільки два кадри аудіо зберігаються в одному кадрі USAC-3D, коли SBR дозволене. Кількість адаптивних транспортних каналів = numHOATransportChannels - MinNumOfCoeffsForAmbHOA = 7 - 4 = 3 адаптивних транспортних канали



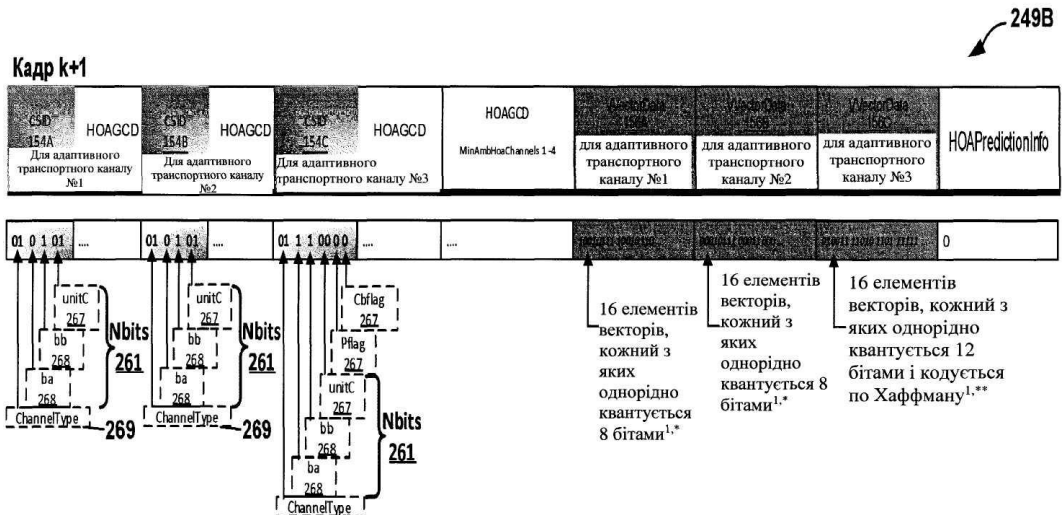
Фіг. 10H(i)

Кадр k

У цьому прикладі адаптивні транспортні канали містять два канали переднього плану на векторній основі в кадрі k і три канали переднього плану на векторній основі у другому кадрі k+1



Фіг. 10H(ii)

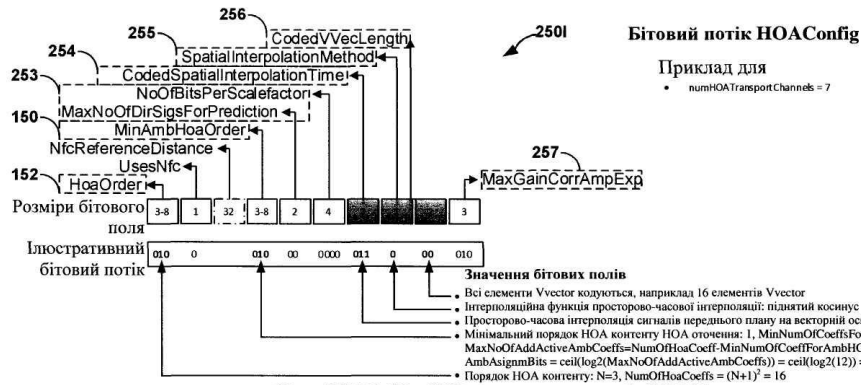


¹Кількість і індекси кодованих елементів VvectorData, вказаних параметром CodedVVecLength=0

*Схема кодування, сигналізована за допомогою NbBitsQ=5 в ChannelSideInfo для цього конкретного транспортного каналу

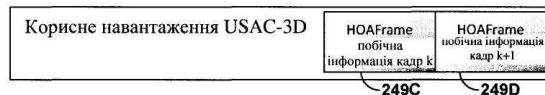
**Схема кодування по Хаффману, сигналізована за допомогою NbBitsQ=12, CbFlag=0 і PFlag=0 в ChannelSideInfo для цього конкретного транспортного каналу

Фіг. 10H(ii)



Кадр USAC-3D з SBR, що включає в себе HOAFrame

2 послідовні HOAFrame зберігаються в розширенні USAC корисного навантаження, оскільки два кадри аудіо зберігаються в одному кадрі USAC-3D, коли SBR дозволене. Кількість адаптивних транспортних каналів $= \text{numHOATransportChannels} - \text{MinNumOfCoeffsForAmbHOA} = 7 - 4 = 3$ адаптивних транспортних каналів

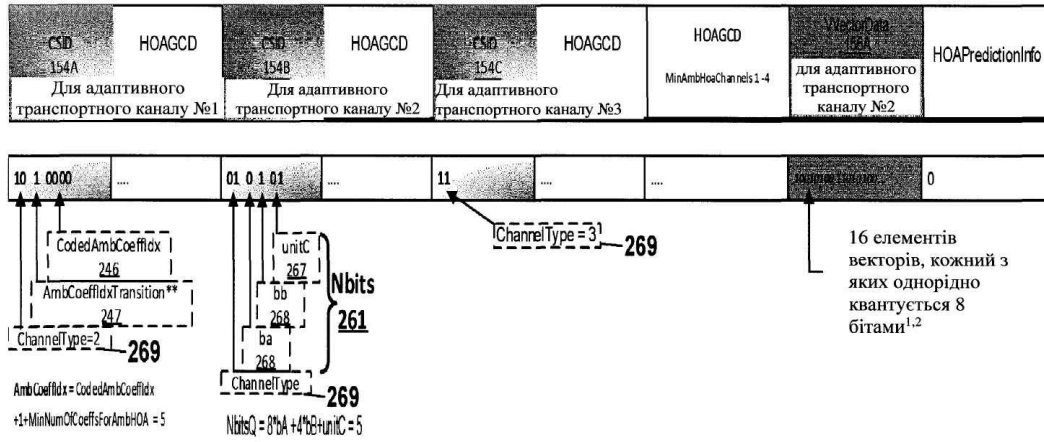


Фіг. 10I(i)

Кадр k

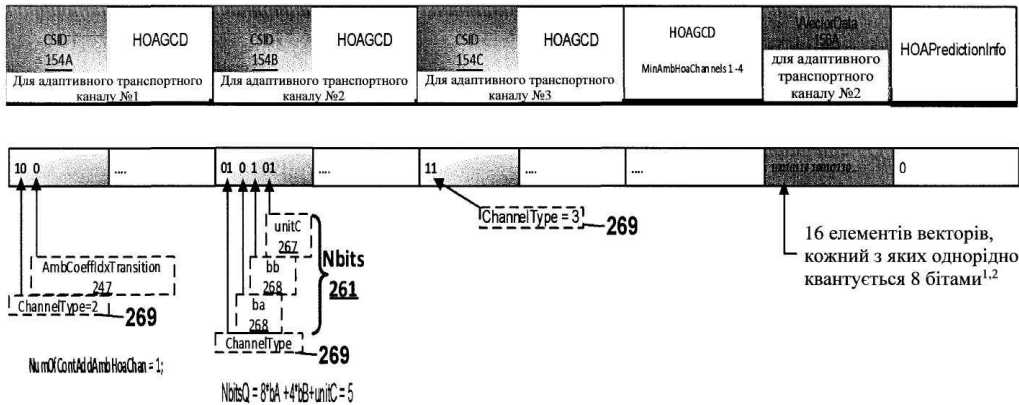
У цьому прикладі адаптивні транспортні канали містять один додатковий фоновий канал HOA і один канал переднього плану на векторній основі

249C



Фіг. 10I(ii)

Кадр k+1

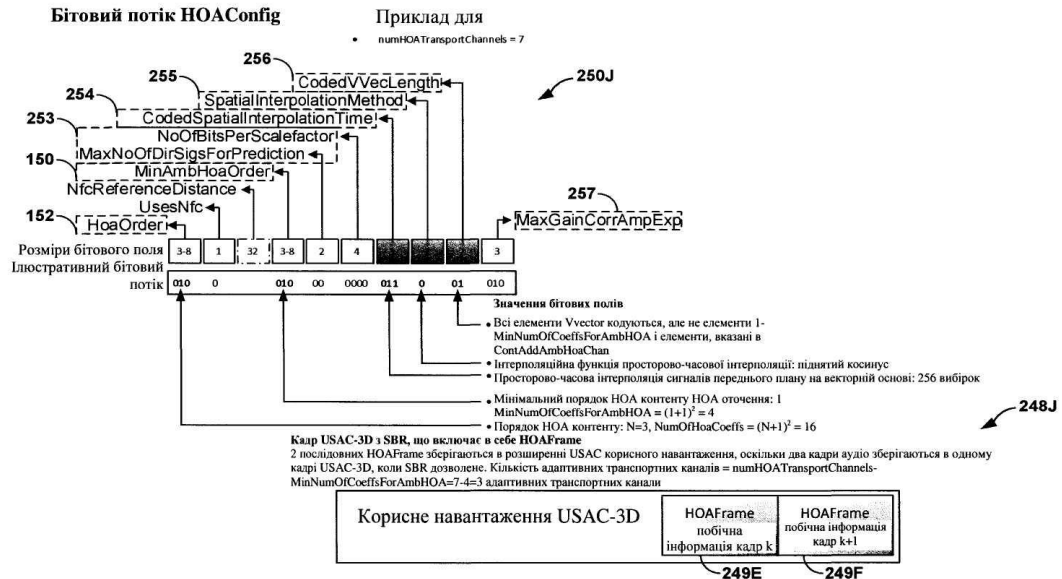


¹Кількість і індекси кодованих елементів VvectorData, вказаних параметром CodedVVecLength=0

²Схема кодування, сигналізована за допомогою NbitsQ=5 в ChannelSideInfo для адаптивного транспортного каналу № 2

**Для адаптивного транспортного каналу №1 внутрішній стан декодера тут передбачається рівним AmbCoeffIdxTransitionState=2. Тому сигналізується бітове поле CodedAmbCoeffIdx

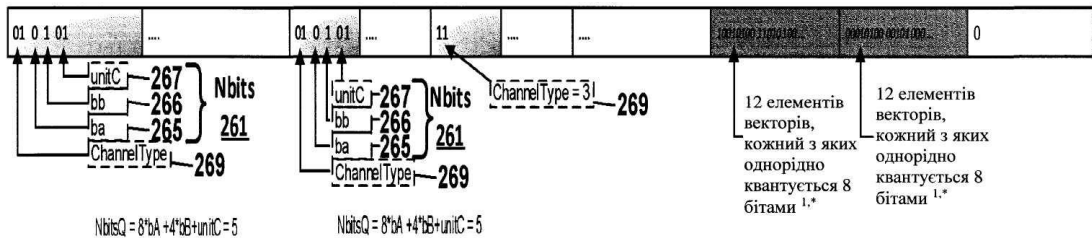
Фіг. 10I(ii)



Фіг. 10J(i)

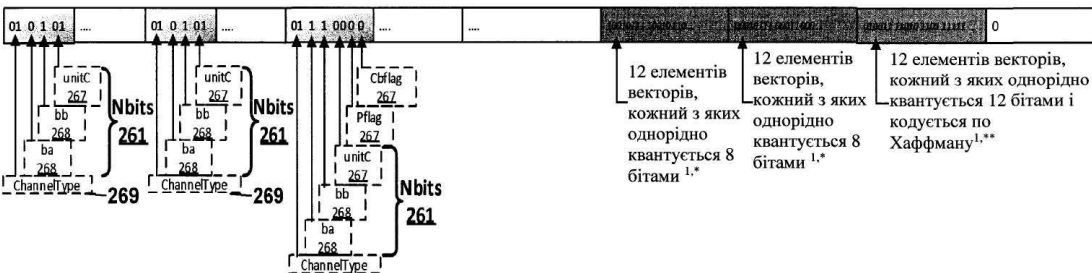
Кадр k

У цьому прикладі адаптивні транспортні канали містять два канали переднього плану на векторній основі в кадрі k і три канали переднього плану на векторній основі у другому кадрі k+1



Фіг. 10J(ii)

Кадр k+1

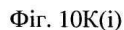


¹Кількість і індекси кодованих елементів VvectorData, вказаних параметром CodedVVecLength, MinAmbHoaOrder, HoaOrder, ніякі додаткові канали HOA не сигналізуються в цих кадрах

*Схема кодування, сигналізована за допомогою NbitsQ=5 в ChannelSideInfo для цього конкретного транспортного каналу

**Схема кодування по Хаффману, сигналізована за допомогою NbitsQ=12, CbFlag=0 і Pflag=0 в ChannelSideInfo для цього конкретного транспортного каналу

Фіг. 10J(ii)



У цьому прикладі адаптивні транспортні канали містять один додатковий фоновий канал НОА і один канал переднього плану на векторній основі

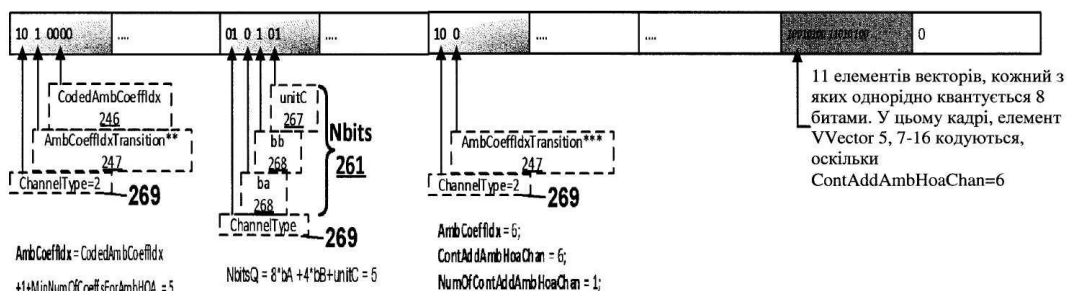
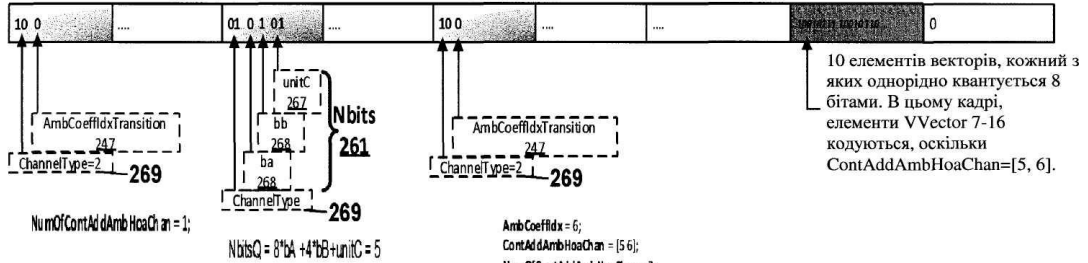
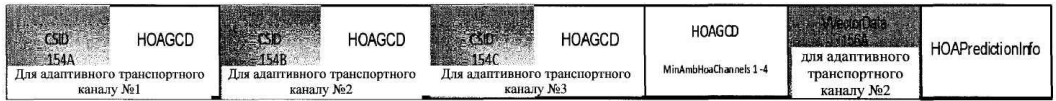


Fig. 10K(ii)

Кадр k+1



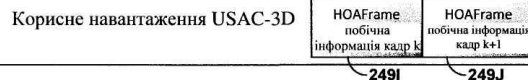
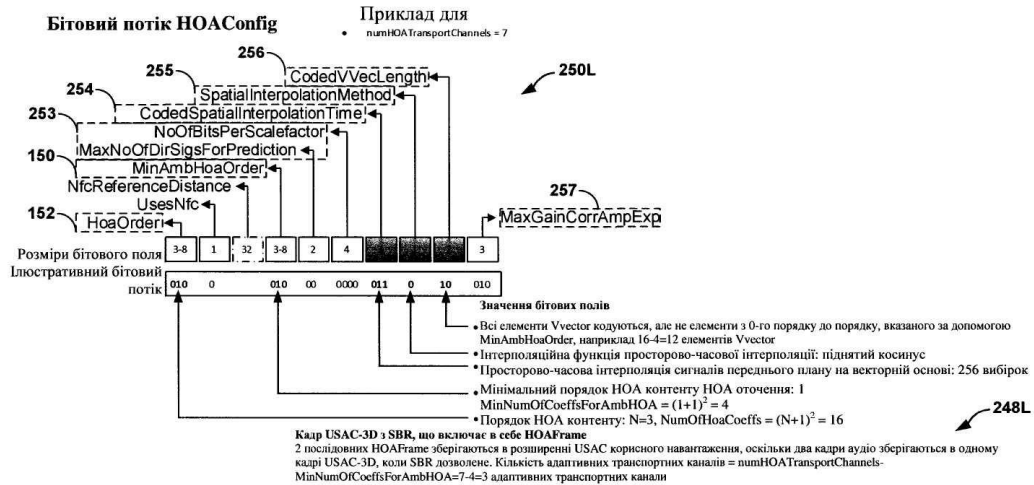
¹Кількість і індекси кодованих елементів VvectorData, вказаних параметром CodedVVecLength, MinNumOfCoeffsForAmbHOA, NumOfHoaCoeff

²Схема кодування, сигналізована за допомогою NbitsQ=5 в ChannelSideInfo для адаптивного транспортного каналу №2

**Для адаптивного транспортного каналу №1 внутрішній стан декодера тут передбачається рівним AmbCoeffIdxTransitionState=2. Тому сигналізується бітове поле CodedAmbCoeffIdx

***Адаптивний транспортний канал №3 заданий в попередньому кадрі для перенесення додаткового коефіцієнта HOA 6. Тому бітове поле CodedAmbCoeffIdx не сигналізується

Фіг. 10K(ii)

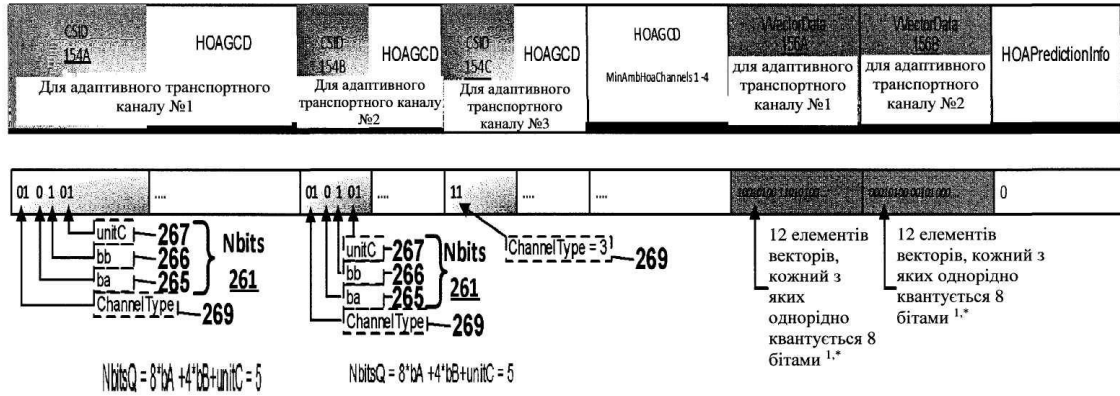


Фіг. 10L(i)

Кадр k

У цьому прикладі адаптивні транспортні канали містять два канали переднього плану на векторній основі в кадри k і три канали переднього плану на векторній основі у другому кадри k+1

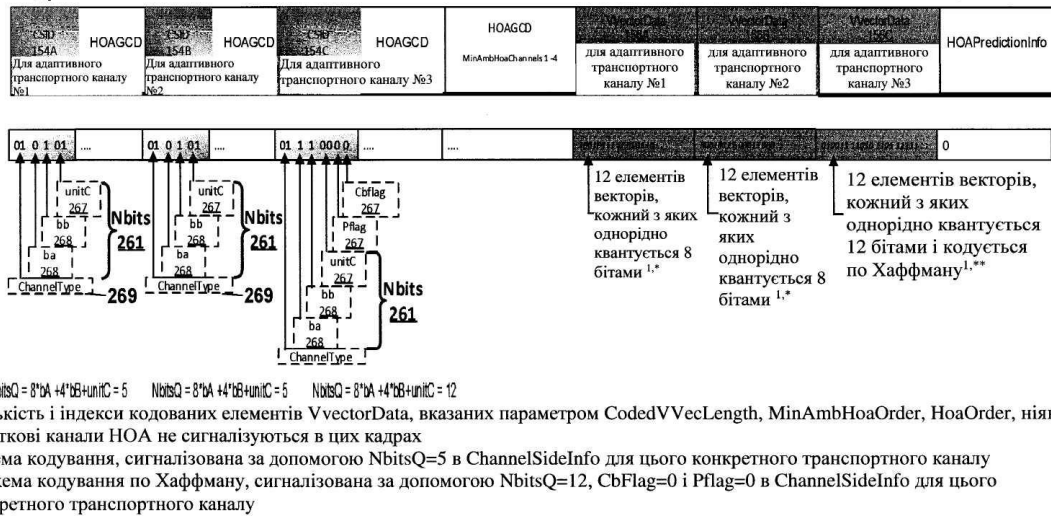
249I



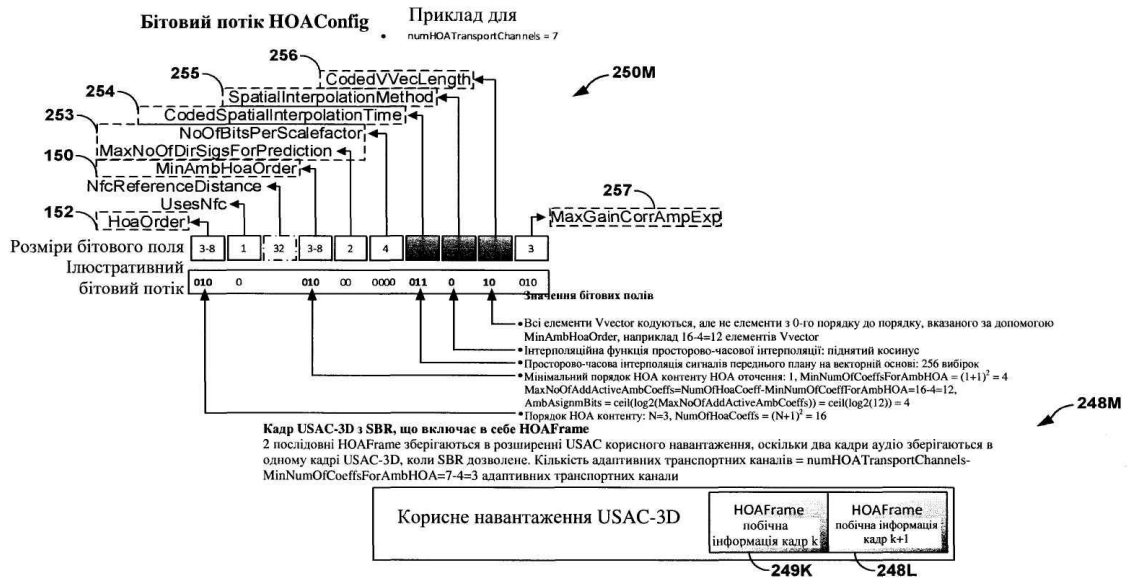
Фіг. 10L(ii)

Кадр k+1

249J



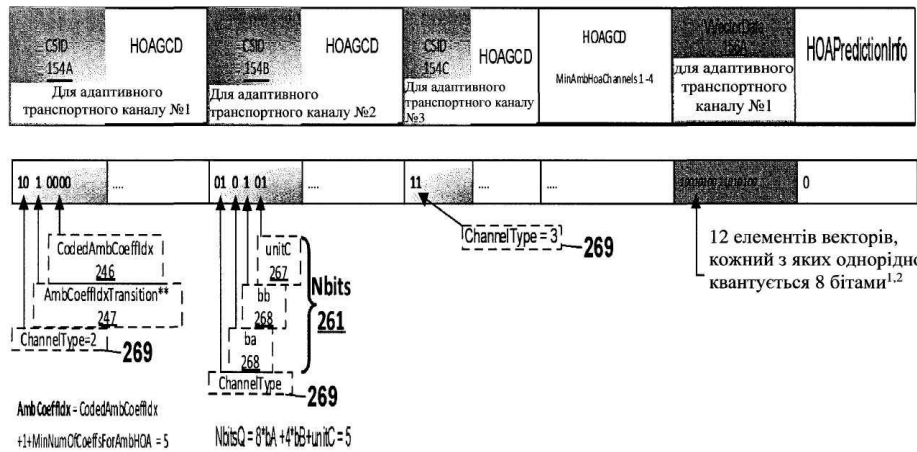
Фіг. 10L(ii)



Фіг. 10M(i)

Кадр k

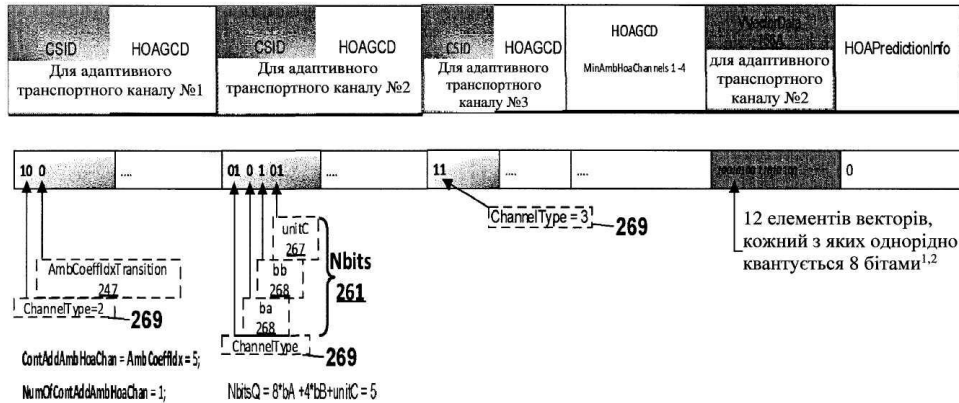
У цьому прикладі адаптивні транспортні канали містять один додатковий фоновий канал HOA і один канал переднього плану на векторній основі



Фіг. 10M(ii)

249L

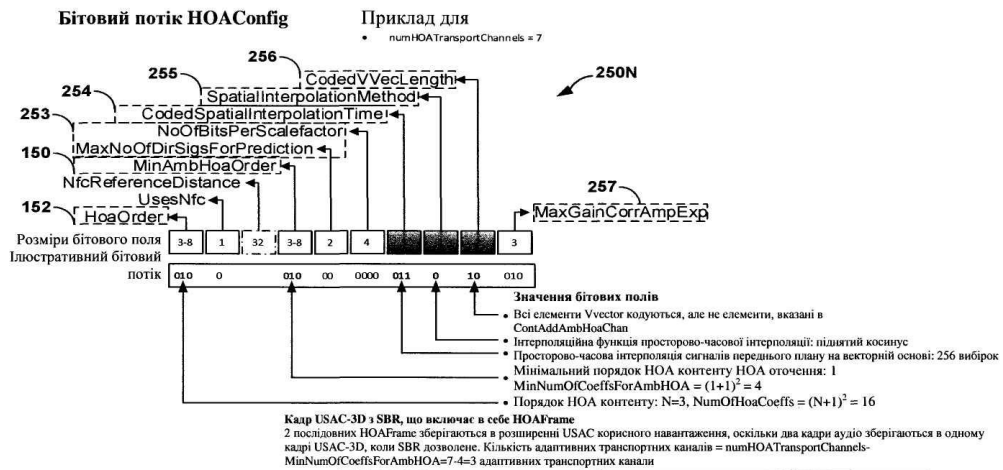
Кадр k+1


¹Кількість і індекси кодованих елементів VvectorData, вказаних параметром CodedVVecLength, MinAmbHoaOrder, HoaOrder

²Схема кодування, сигналізована за допомогою NbitsQ=5 в ChannelSideInfo для адаптивного транспортного каналу № 2

**Для адаптивного транспортного каналу №1 внутрішній стан декодера тут передбачається рівним AmbCoeffIdxTransitionState=2. Тому сигналізується бітове поле CodedAmbCoeffIdx.

Фіг. 10M(ii)



Корисне навантаження USAC-3D

HOAFrame

побічна інформація кадр k

HOAFrame

побічна інформація кадр k+1

249M

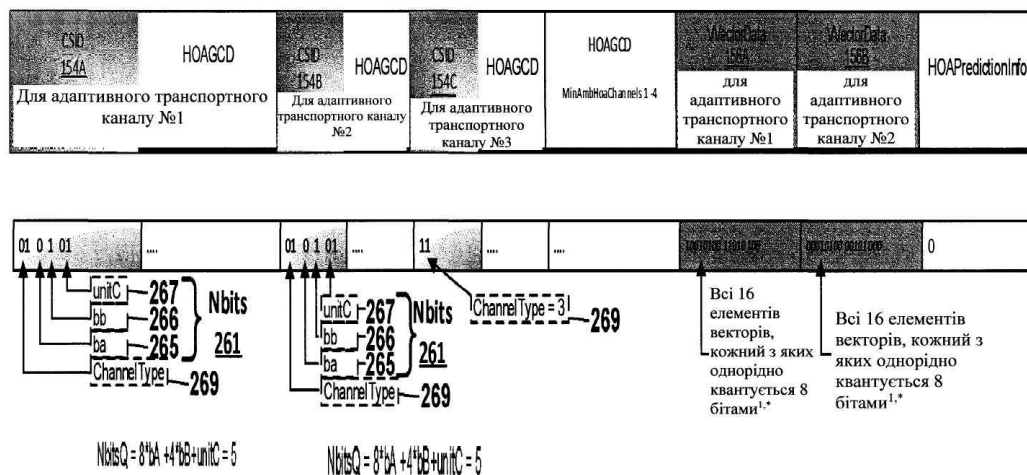
249N

Фіг. 10N(i)

Кадр k

У цьому прикладі адаптивні транспортні канали містять два канали переднього плану на векторній основі в кадрі k і три канали переднього плану на векторній основі у другому кадрі $k+1$

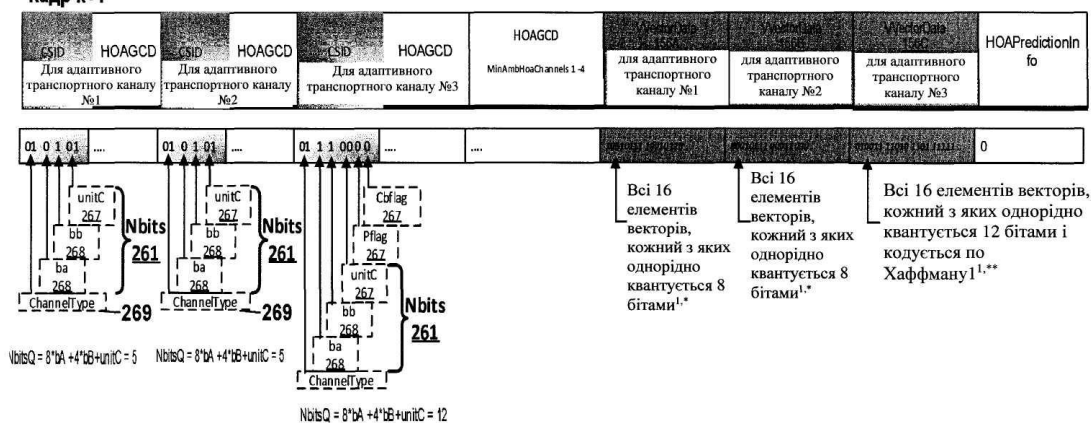
249M



Фиг. 10N(ii)

Кадр $k+1$

249N

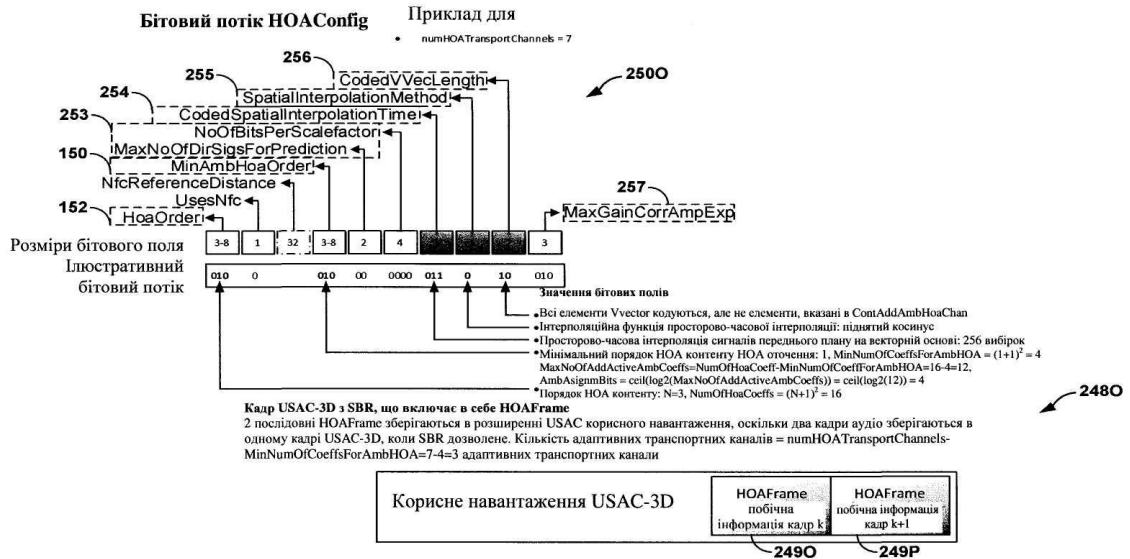


¹Кількість і індекси кодованих елементів VectorData, вказаних параметром CodedVVecLength, MinAmbHoaOrder, HoaOrder, ніякі додаткові канали НОА не сигналізуються в цих кадрах

*Схема кодування, сигналізована за допомогою NbitsQ=5 в ChannelSideInfo для цього конкретного транспортного каналу

**Схема кодирования по Хаффману, сигналізована за допомогою NbitsQ=12, CbFlag=0 і Pflag=0 в ChannelSideInfo для цього конкретного транспортного каналу

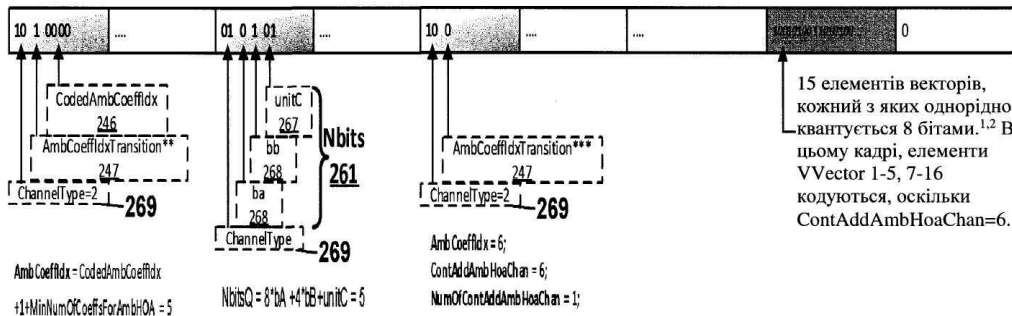
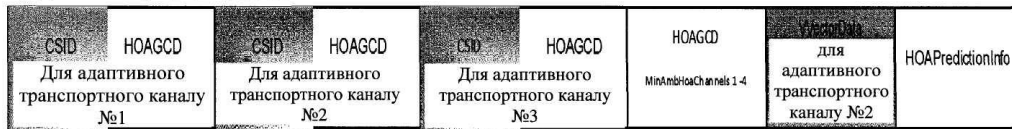
Фиг. 10N(ii)



Фиг. 100(i)

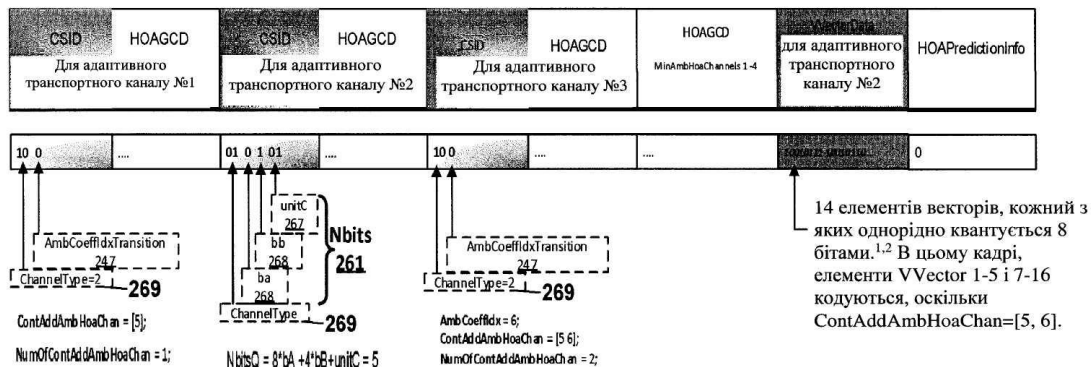
Кадр k

У цьому прикладі адаптивні транспортні канали містять один додатковий фоновий канал HOA і один канал переднього плану на векторній основі



Фиг. 100(ii)

Кадр k+1



¹Кількість і індекси кодованих елементів VvectorData, вказаних параметром CodedVVecLength, MinNumOfCoeffsForAmbHOA, NumOfHoaCoeff
²Схема кодування, сигналізована за допомогою NbitsQ=5 в ChannelSideInfo для адаптивного транспортного каналу № 2
 **Для адаптивного транспортного каналу №1 внутрішній стан декодера тут передбачається рівним AmbCoeffIdxTransitionState=2. Тому сигналізується бітове поле CodedAmbCoeffIdx
 ***Адаптивний транспортний канал № 3 заданий в попередньому кадрі для перенесення додаткового коефіцієнта HOA 6. Тому бітове поле CodedAmbCoeffIdx не сигналізується

Фиг. 100(ii)



Фиг. 11A



Фиг. 11B



Фиг. 11C



Fig. 11D

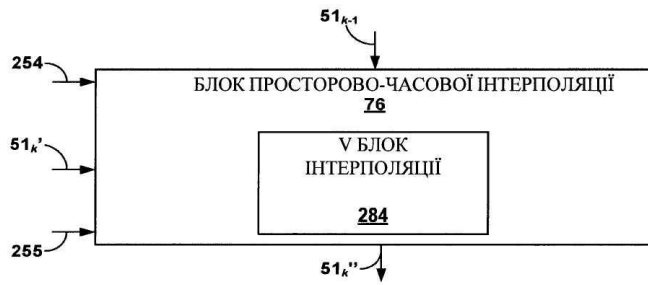


Fig. 11E

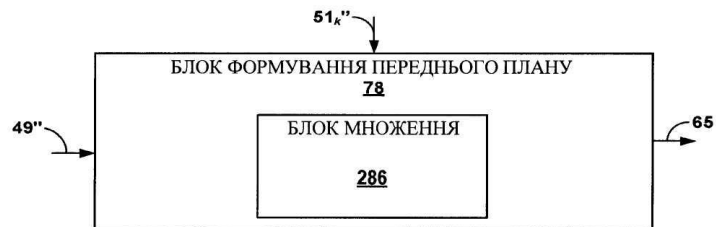


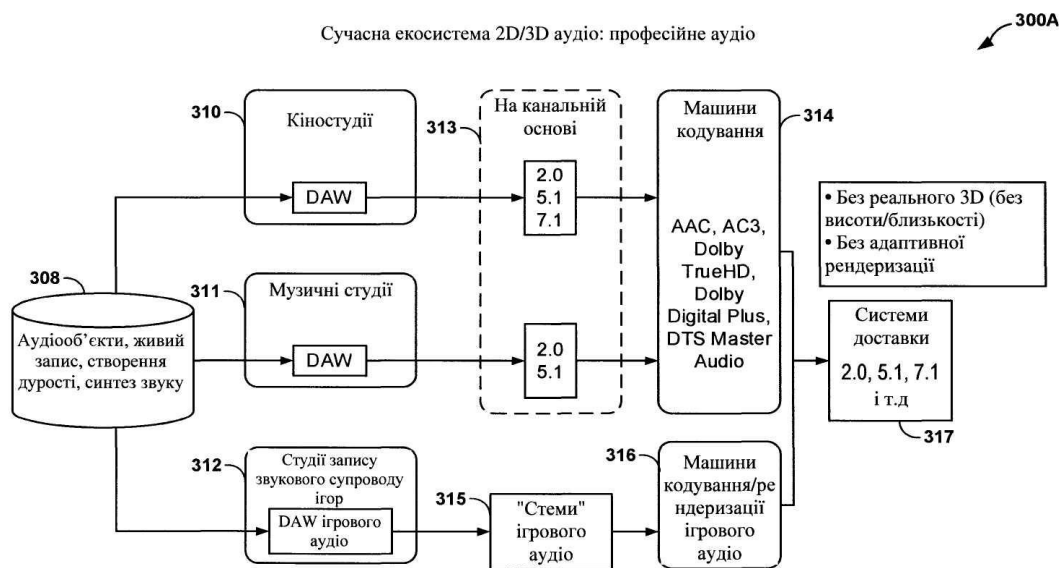
Fig. 11F



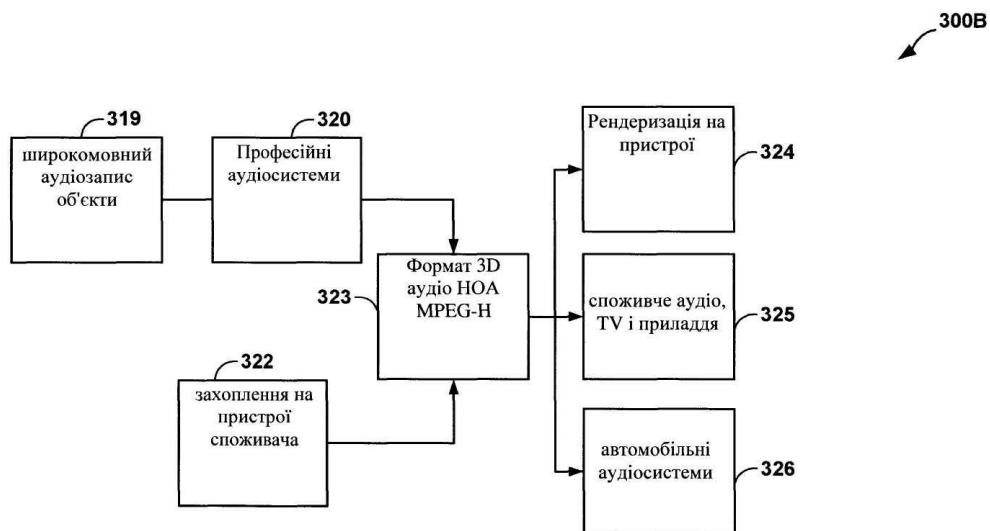
Fig. 11G



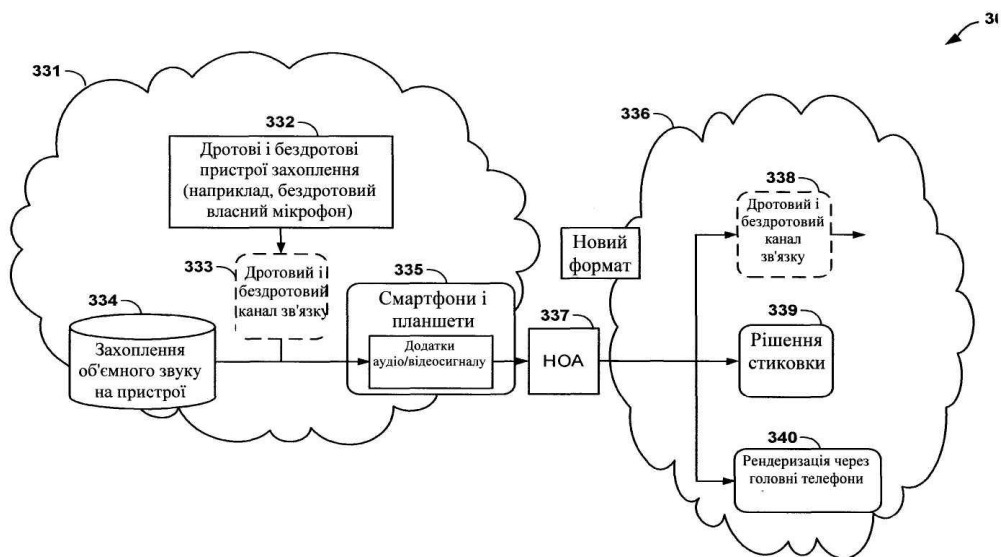
Фіг. 12



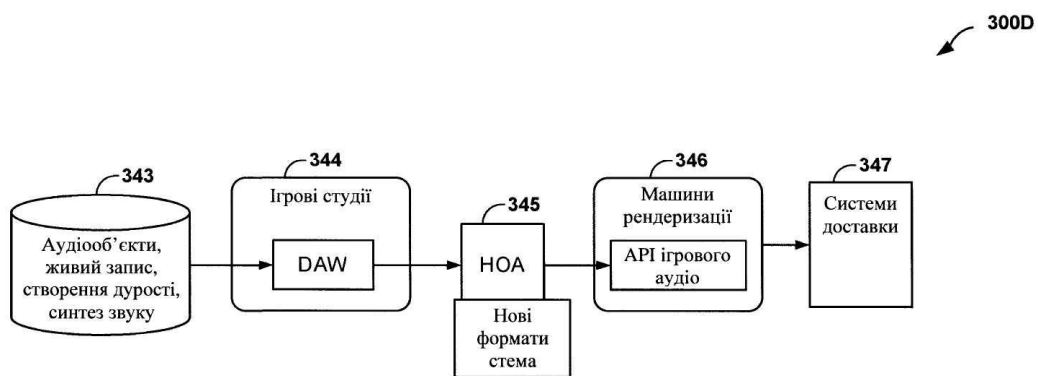
Фіг. 13



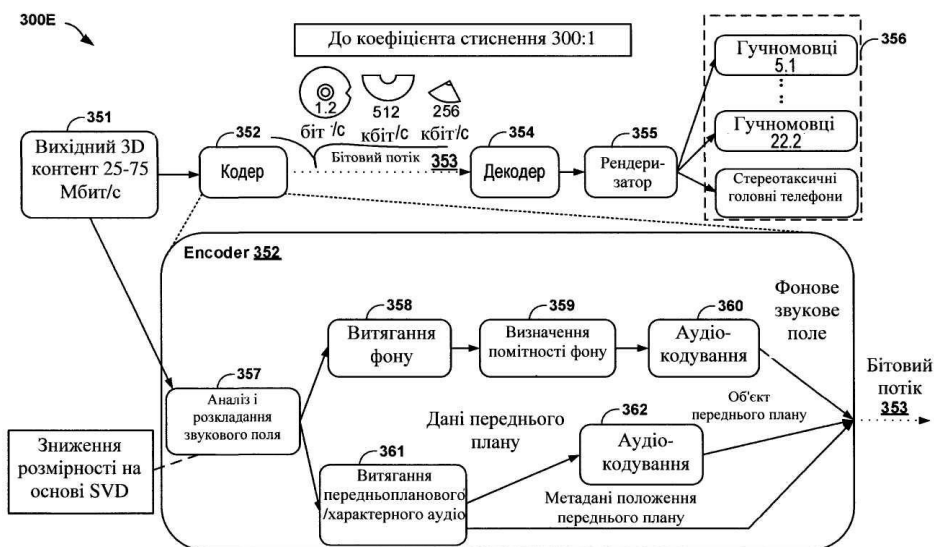
Фіг. 14



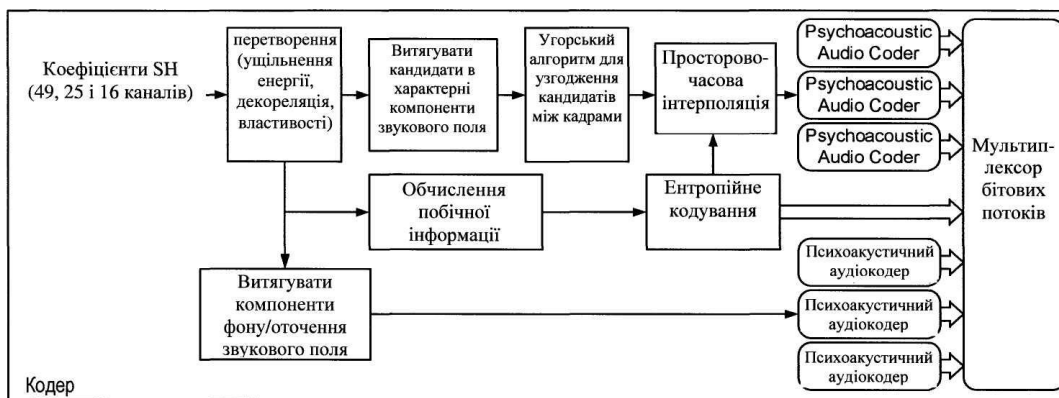
Фіг. 15A



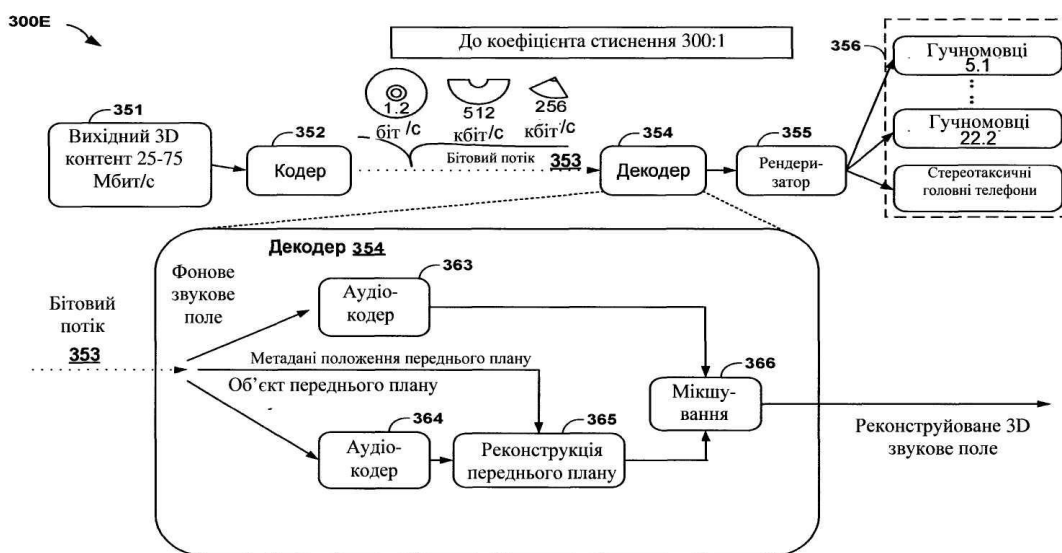
Фіг. 15B



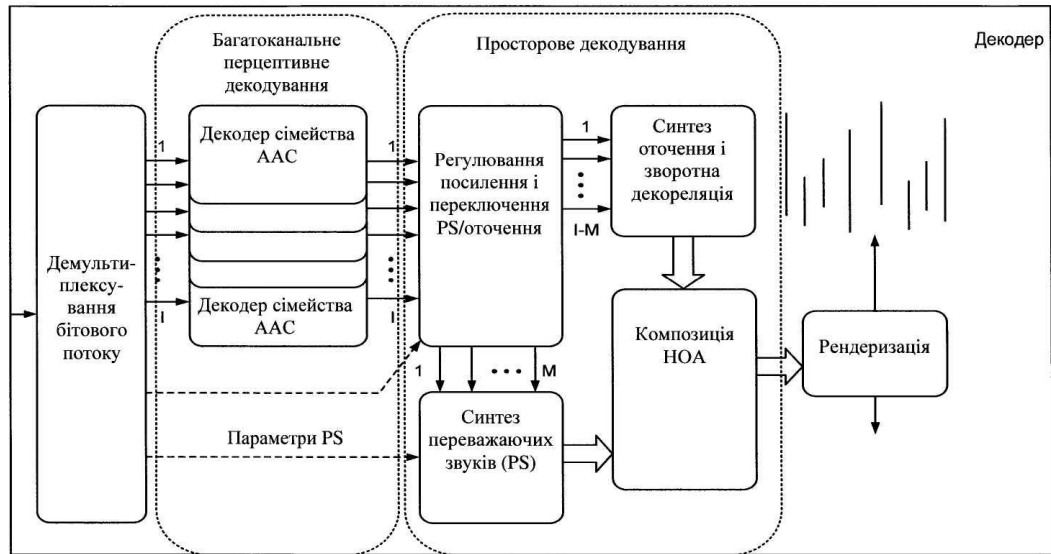
Фіг. 16



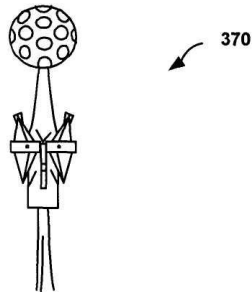
Фіг. 17



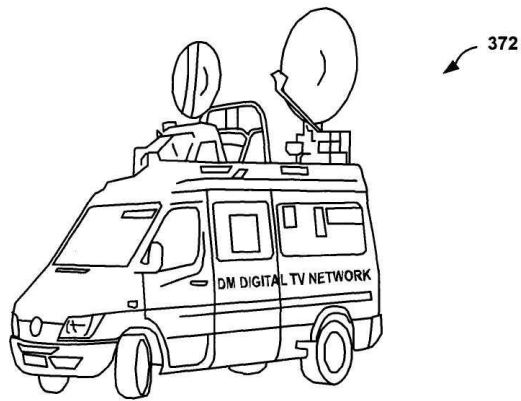
Фіг. 18



Фіг. 19



Фіг. 20A



Фіг. 20B

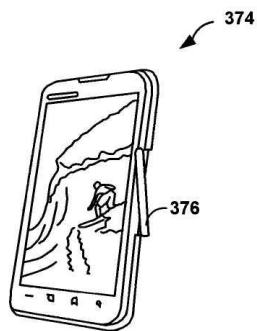


Fig. 20C

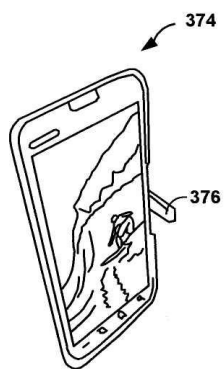


Fig. 20D

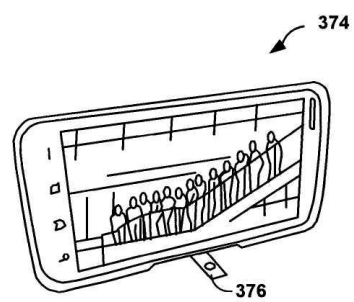


Fig. 20E

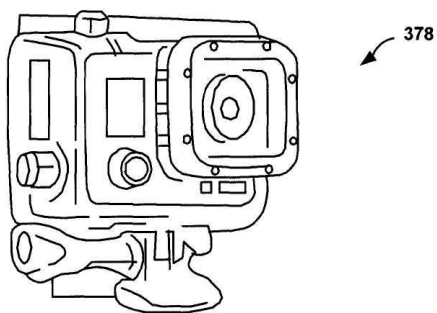


Fig. 20F

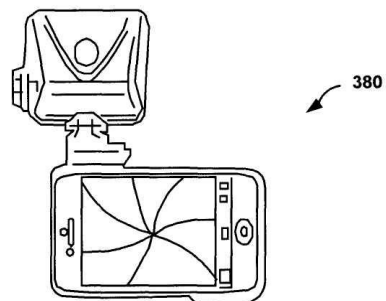


Fig. 20G

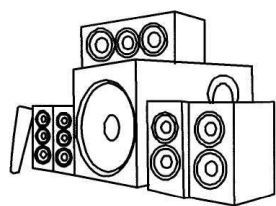


Fig. 21A

382



386A

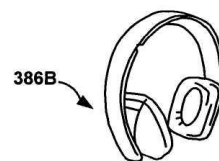


Fig. 21D

386B

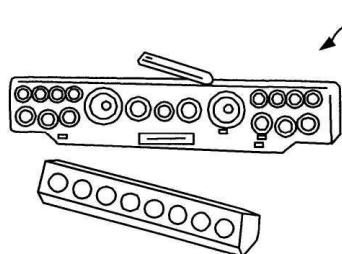


Fig. 21C

384

386C



Fig. 21E

Fig. 21D

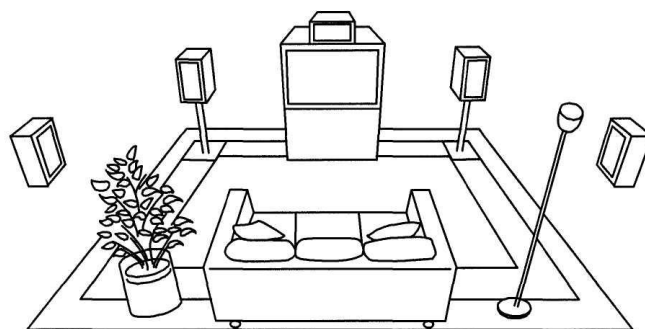


Fig. 22A

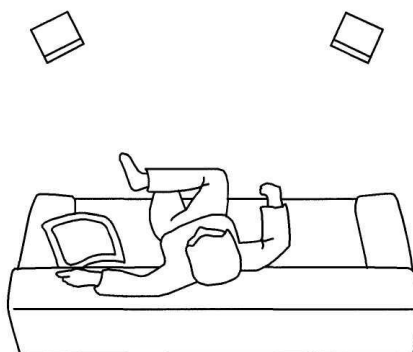
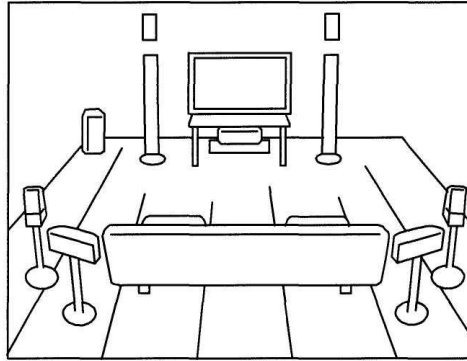
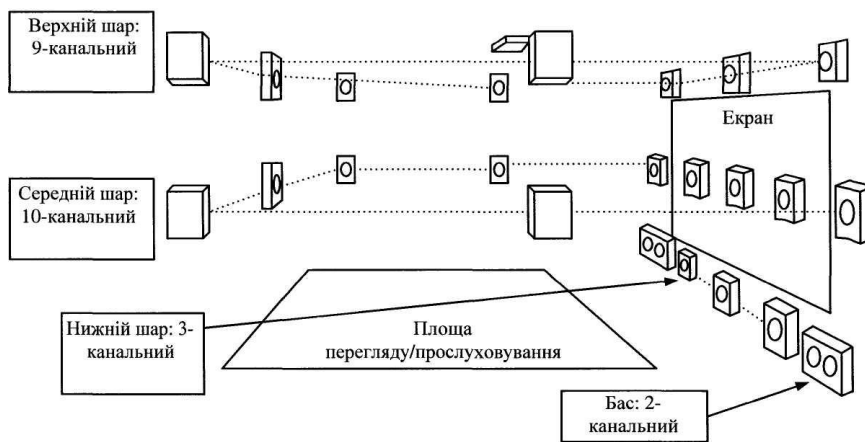


Fig. 22B



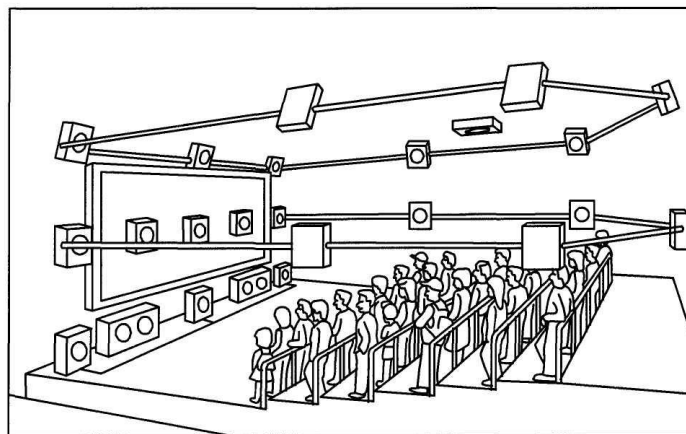
9.1 з повною висотою

Фіг. 22С

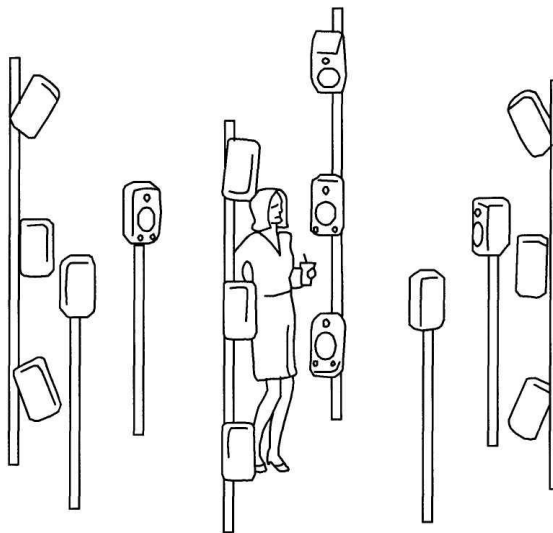


22.2

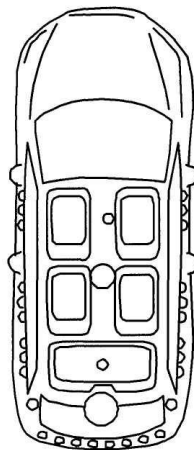
Фіг. 22D



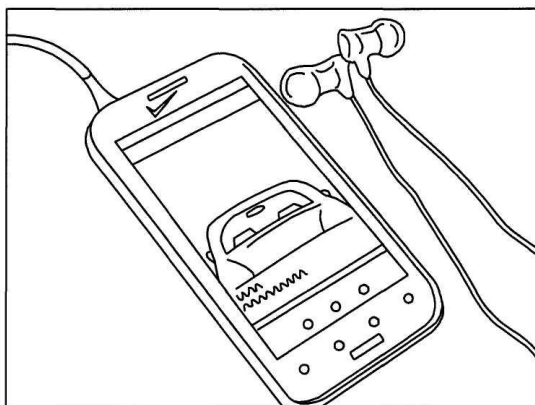
Фіг. 22Е



Φir. 22F



Φir. 22G



Φir. 22H

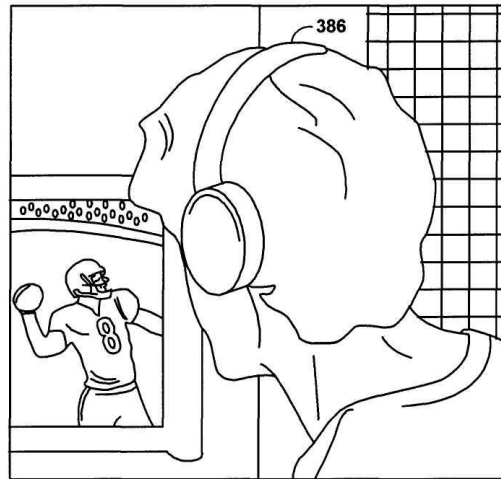


Fig. 23

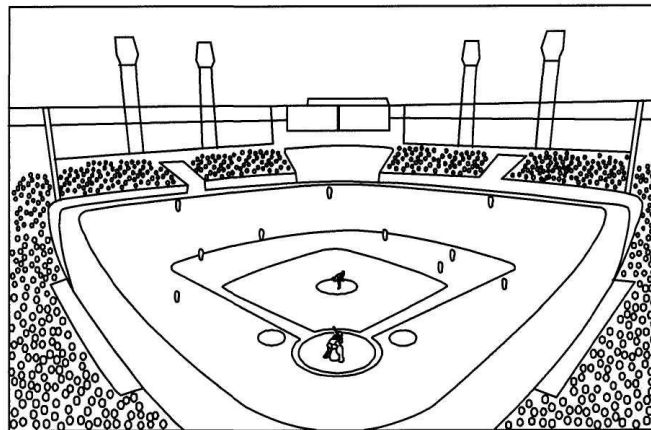
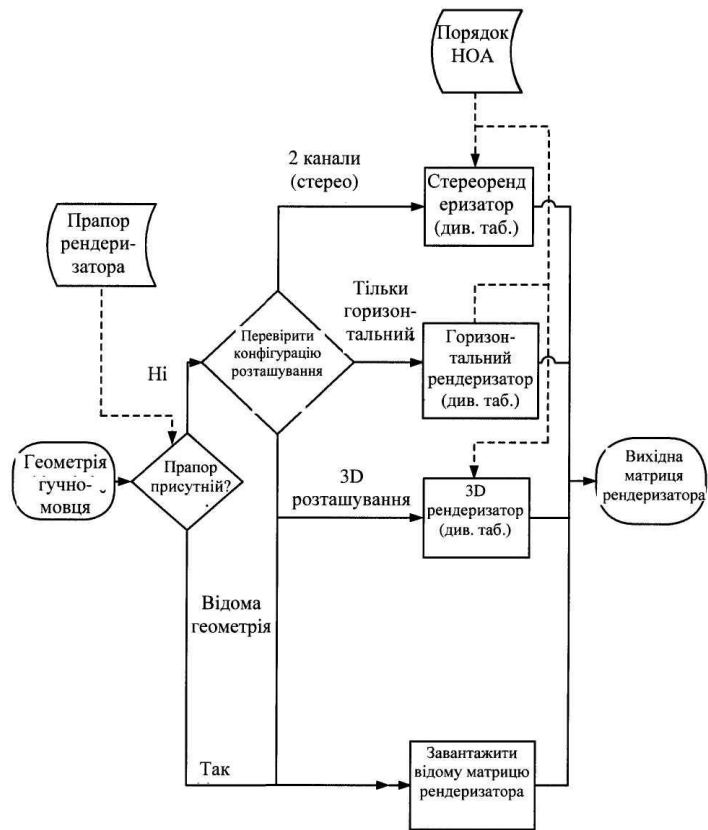
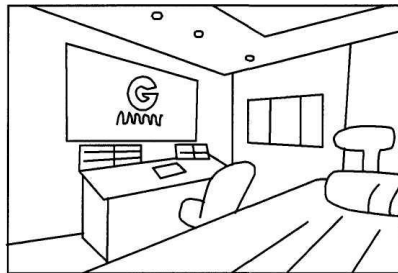


Fig. 24



Фіг. 25



Фіг. 26



Фіг. 27

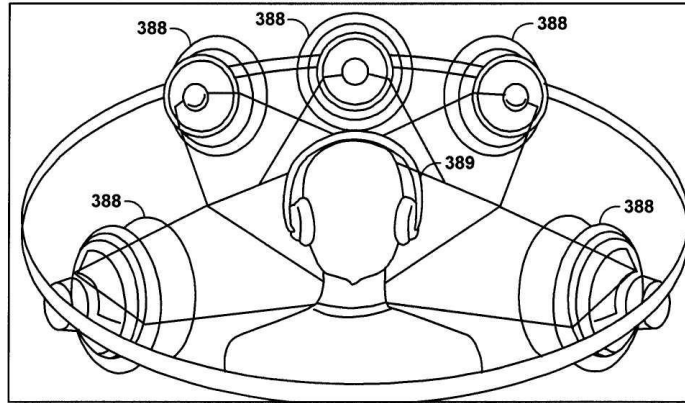


Fig. 28

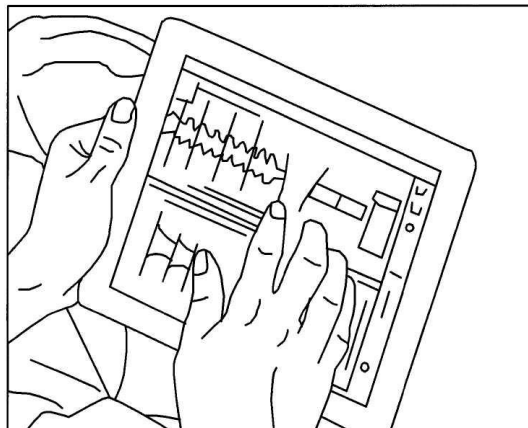
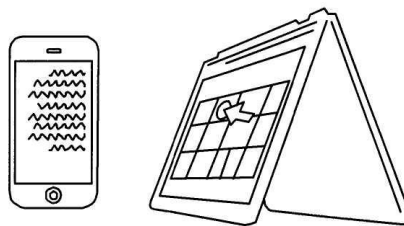


Fig. 29

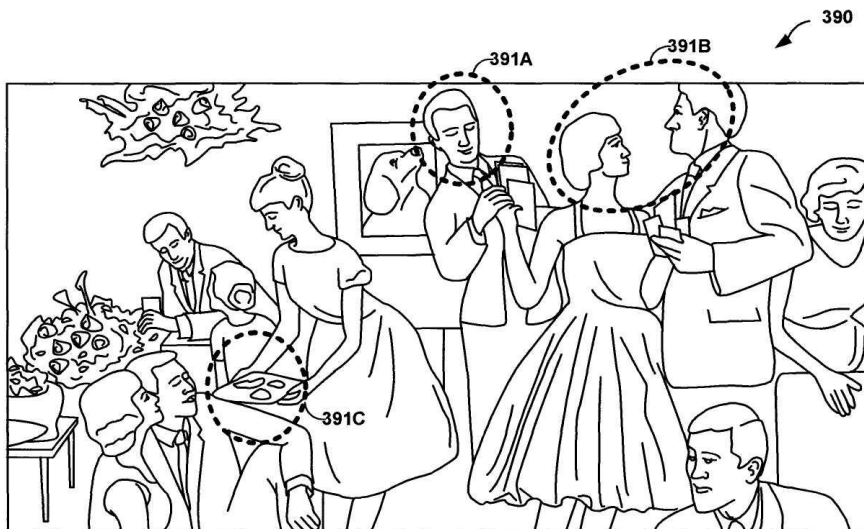


Fig. 30

Результат автоматичного покадрового класифікатора:

DATA \geq THR(30) \rightarrow RMO Enc/Dec
DATA $<$ THR(30) \rightarrow CE Enc/Dec

ΓΡΑΦΙΚ 1

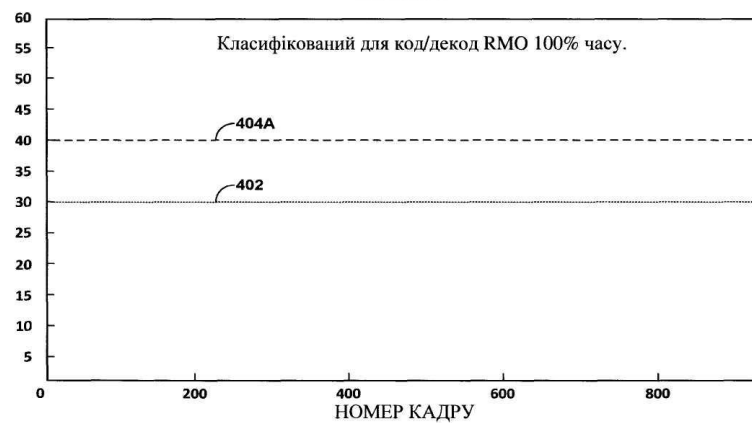


Fig. 31A

Результат автоматичного покадрового класифікатора:

DATA \geq THR(30) -> RMO Enc/Dec
DATA < THR(30) -> CE Enc/Dec

ΓΡΑΦΙΚ 2

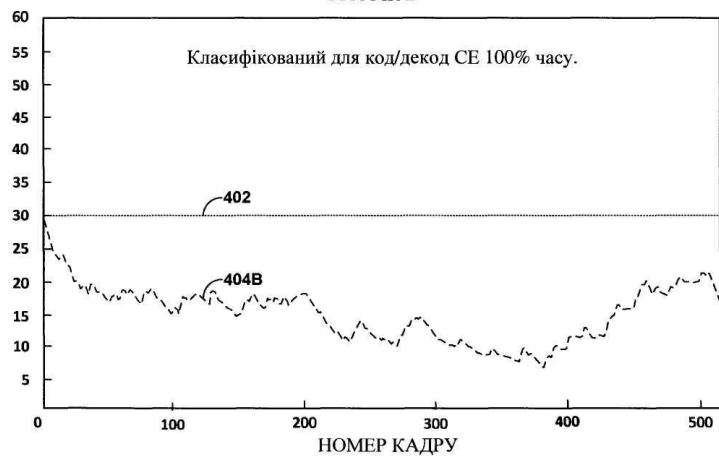


Fig. 31B

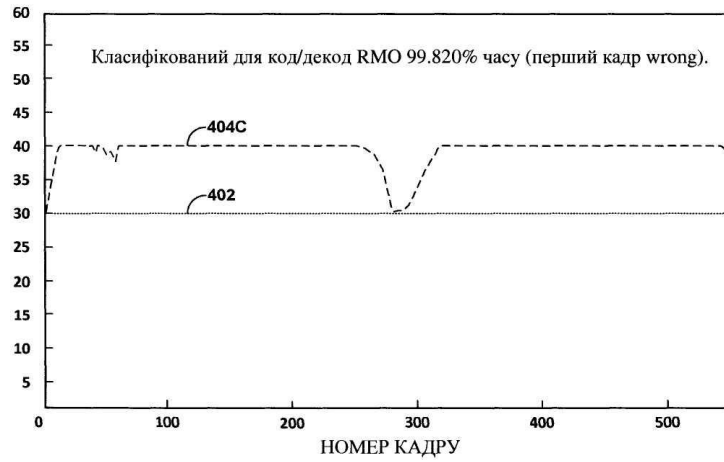
Результат автоматичного покадрового класифікатора:

DATA \geq THR(30) -> RMO Enc/Dec

DATA < THR(30) -> CE Enc/Dec

400C

ГРАФІК 3



Фіг. 31C

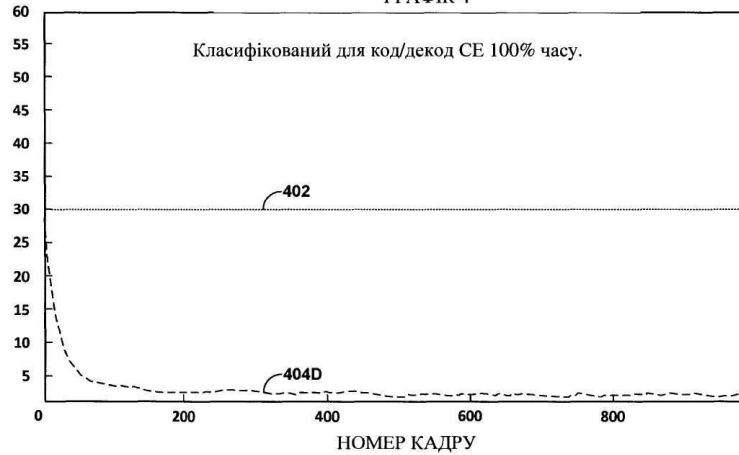
Результат автоматичного покадрового класифікатора:

DATA \geq THR(30) -> RMO Enc/Dec

DATA < THR(30) -> CE Enc/Dec

400D

ГРАФІК 4



Фіг. 31D

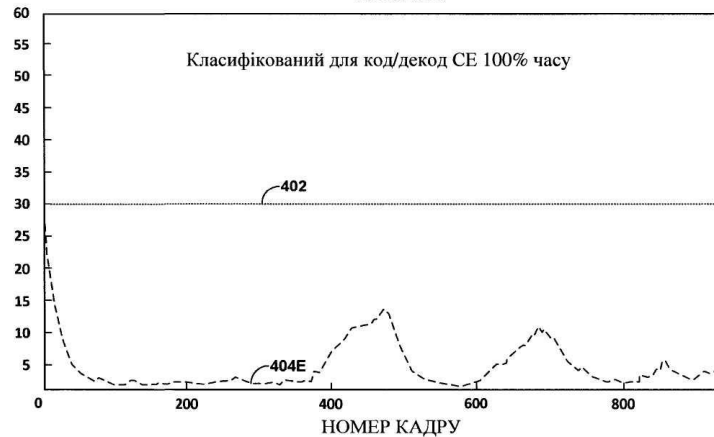
Результат автоматичного покадрового класифікатора:

DATA \geq THR(30) -> RMO Enc/Dec

DATA < THR(30) -> CE Enc/Dec

400E

ГРАФІК 5



Фіг. 31E

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

400F

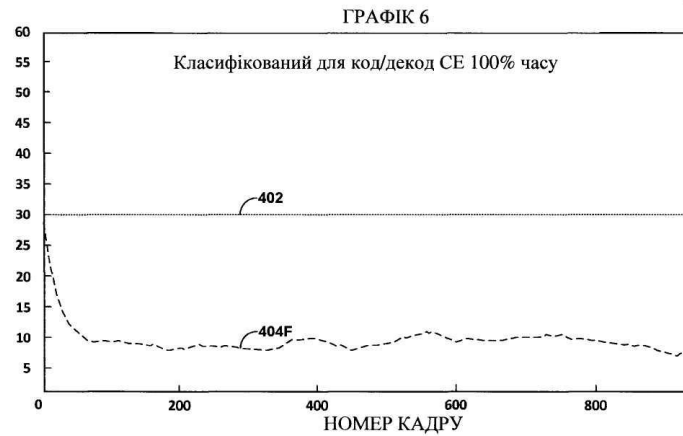


Fig. 31F

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

400G

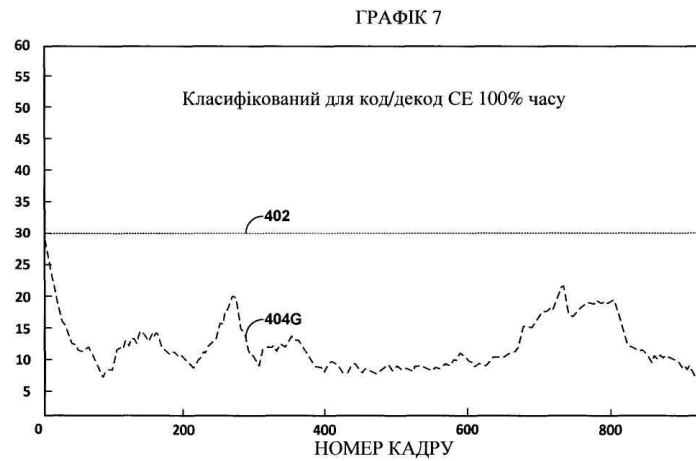


Fig. 31G

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

400H

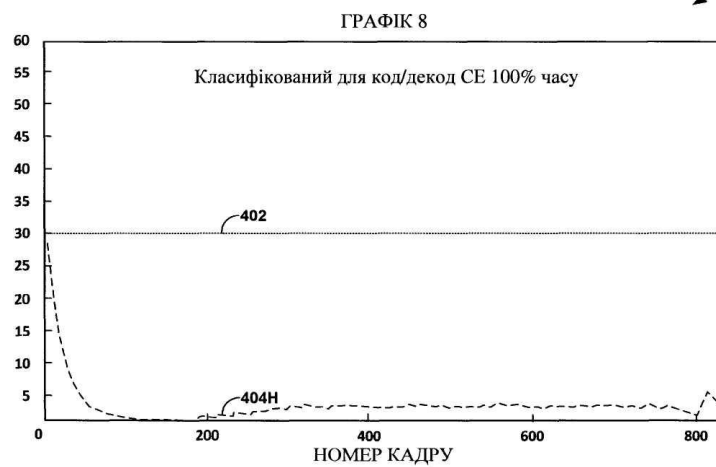


Fig. 31H

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

400I

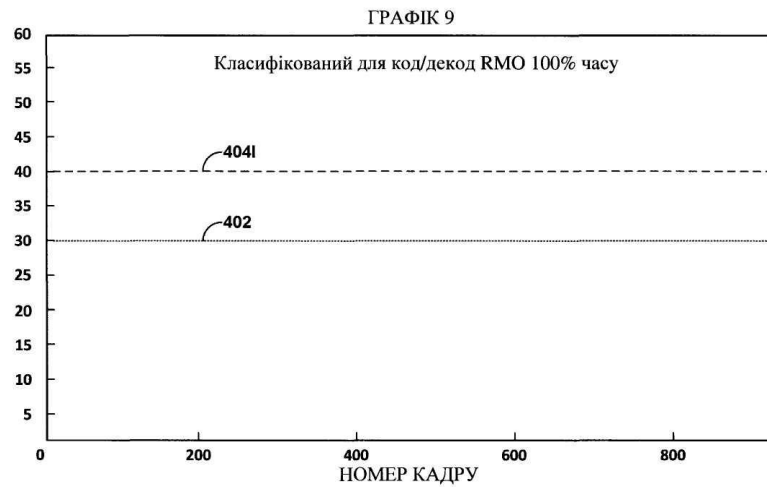


Fig. 31I

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

400J

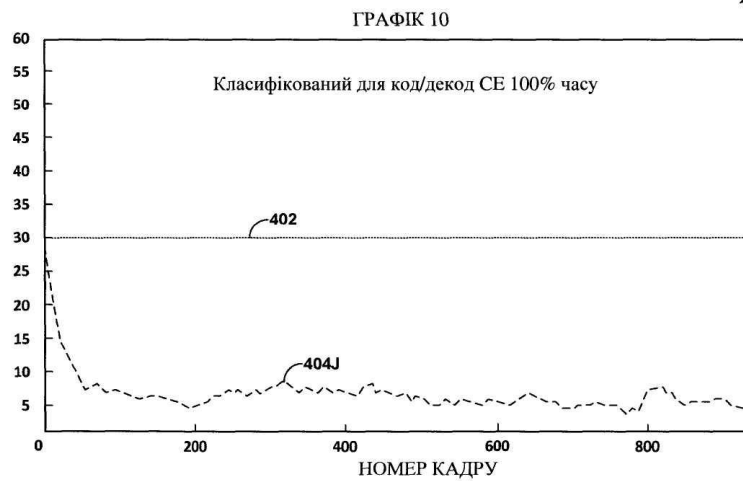


Fig. 31J

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

400K

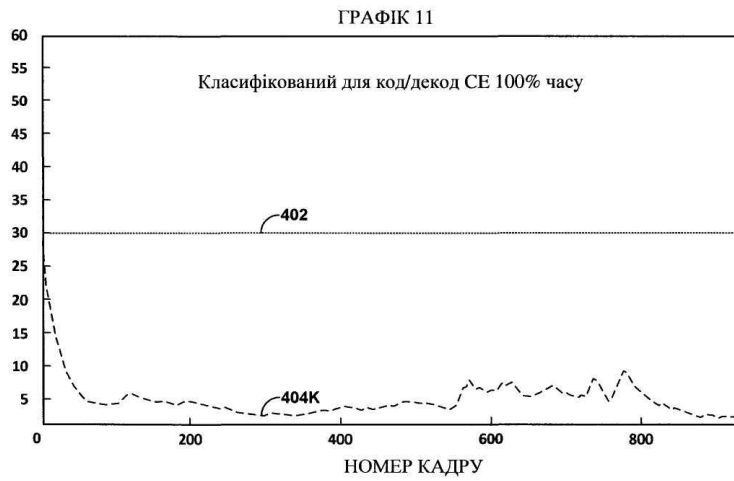
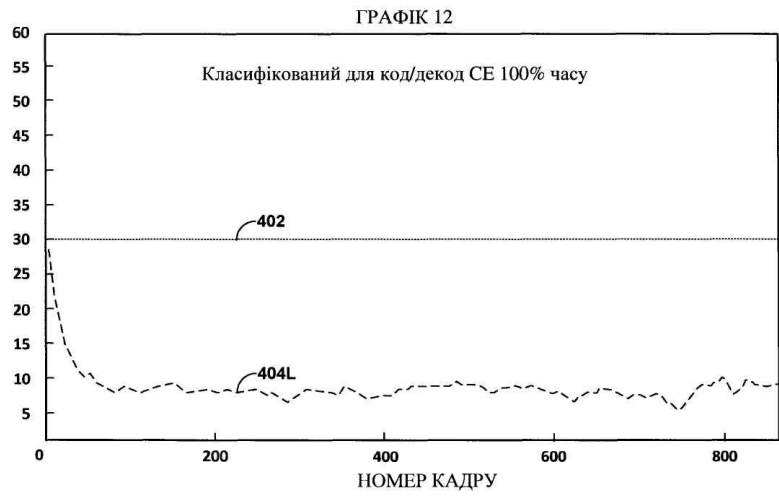


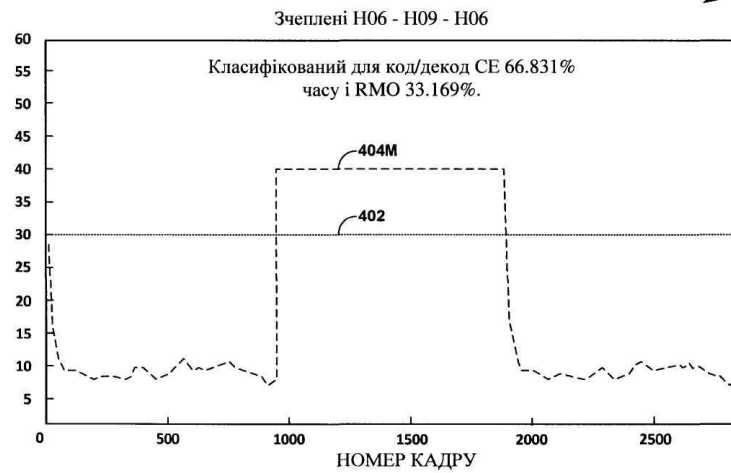
Fig. 31K

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

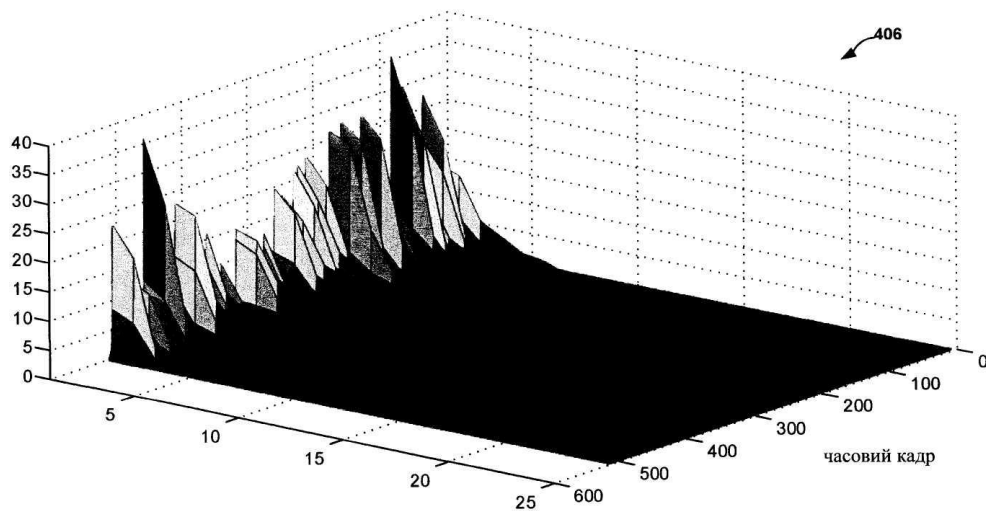


Фіг. 31L

Результат автоматичного покадрового класифікатора:
 DATA \geq THR(30) -> RMO Enc/Dec
 DATA < THR(30) -> CE Enc/Dec

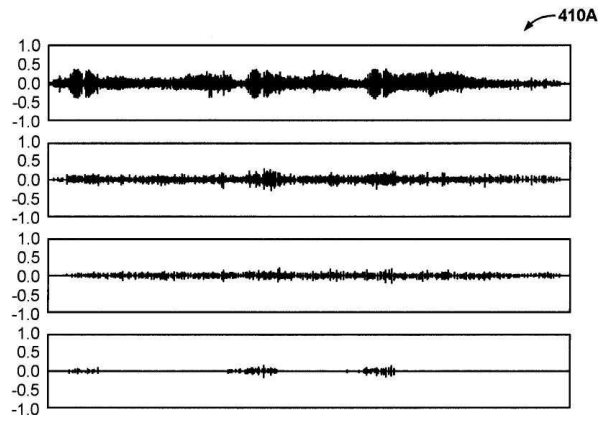


Фіг. 31M

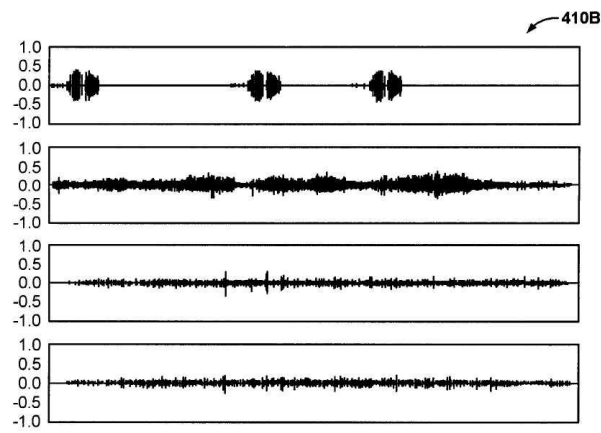


ненульові сингулярні значення S (знайдені на діагональних елементах)

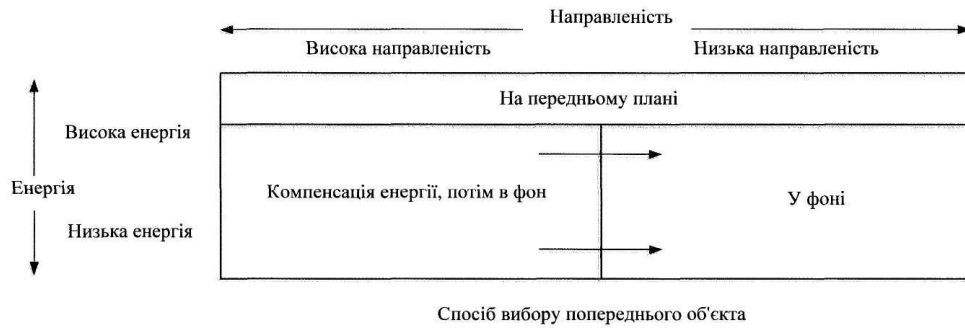
Фіг. 32



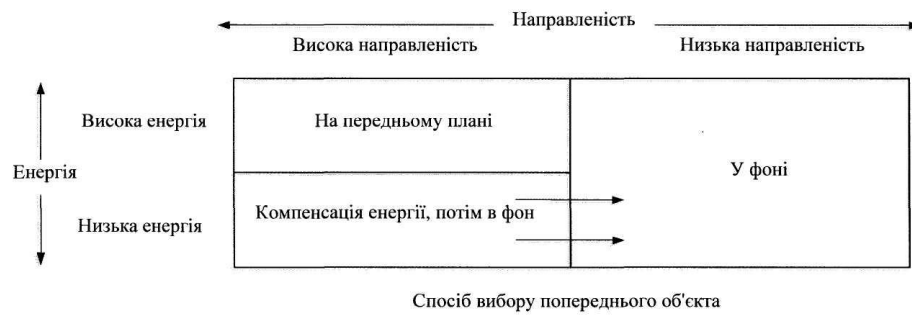
Фіг. 33А



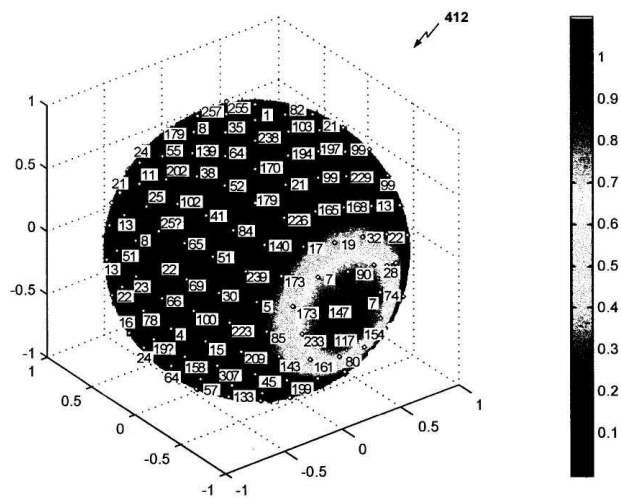
Фіг. 33В



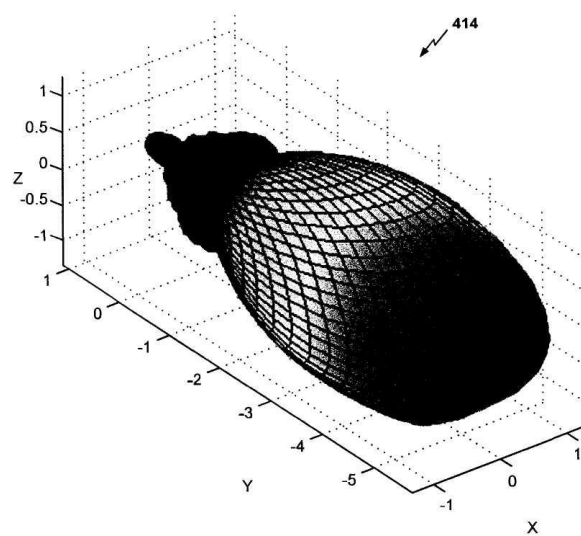
Фіг. 34



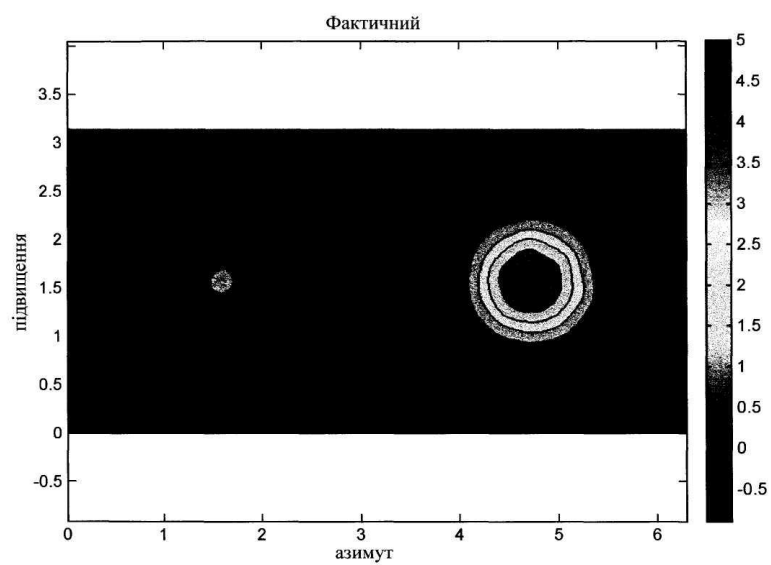
Фіг. 35



Фиг. 36А



Фиг. 36В



Фиг. 36С

Звук бджоли надходить з азимута = 0° і підвищення = 45°

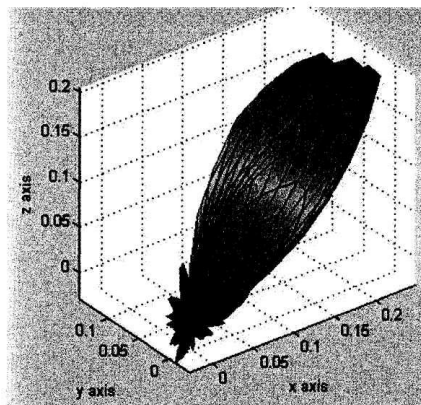


Fig. 36D

Звук вертольота надходить з неба

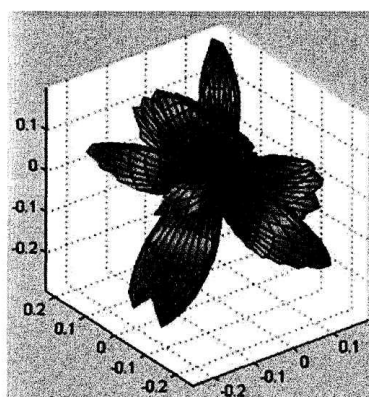


Fig. 36E

Сучасна електронна музика надходить з різних напрямів

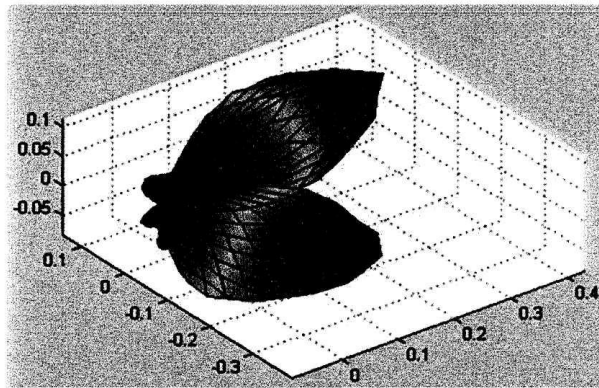


Fig. 36F

Люди кричать на стадіоні

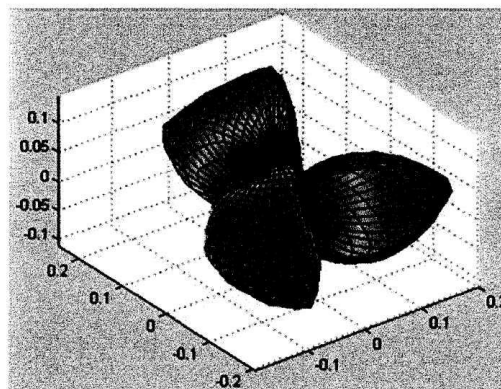


Fig. 36G

$$V_x = \frac{d_2}{d_1 + d_2} V_1 + \frac{d_1}{d_1 + d_2} V_2$$

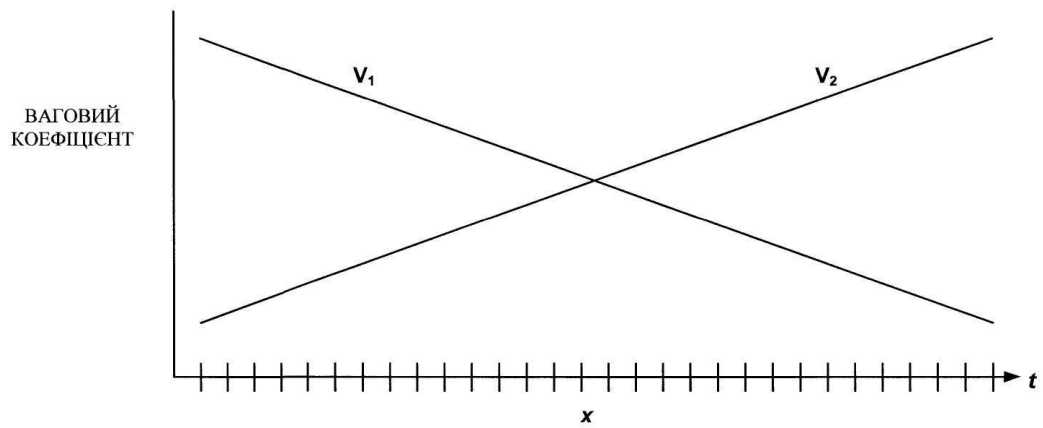


Fig. 37

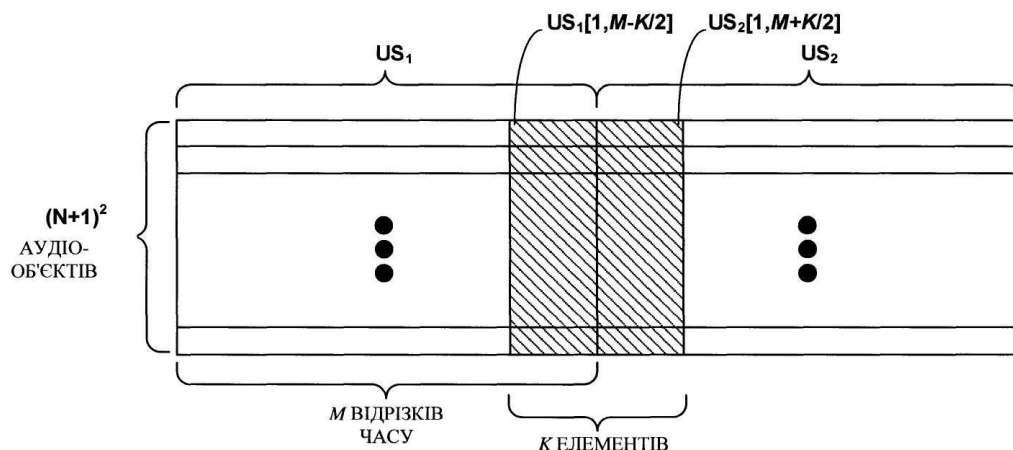


Fig. 38

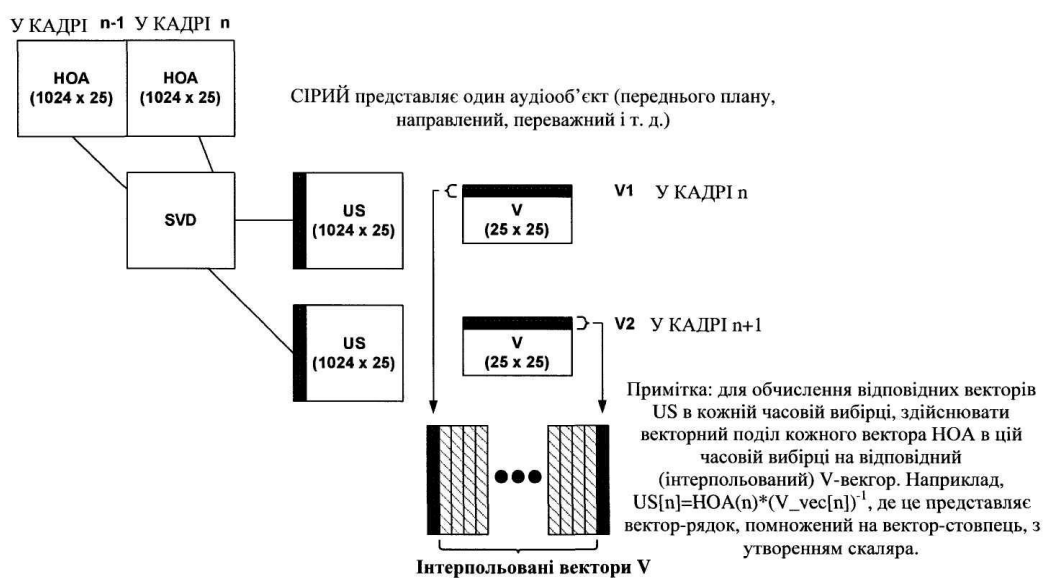


Fig. 39

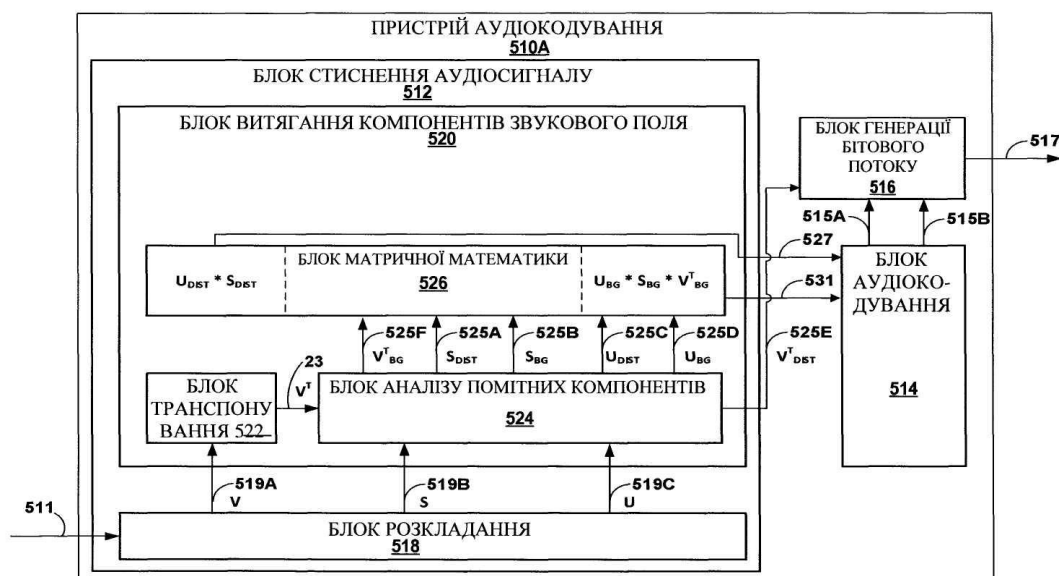
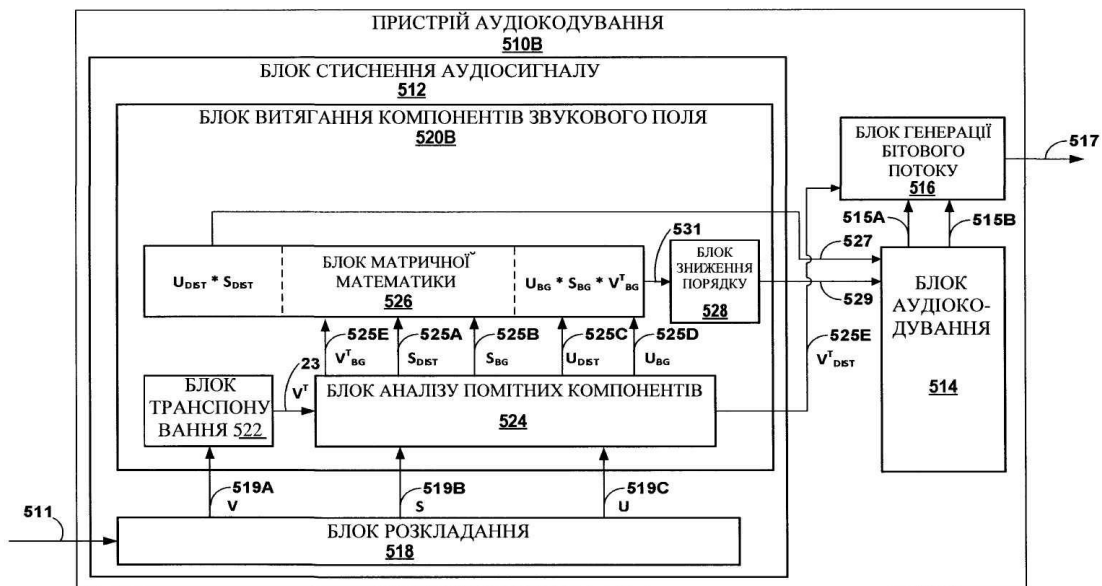
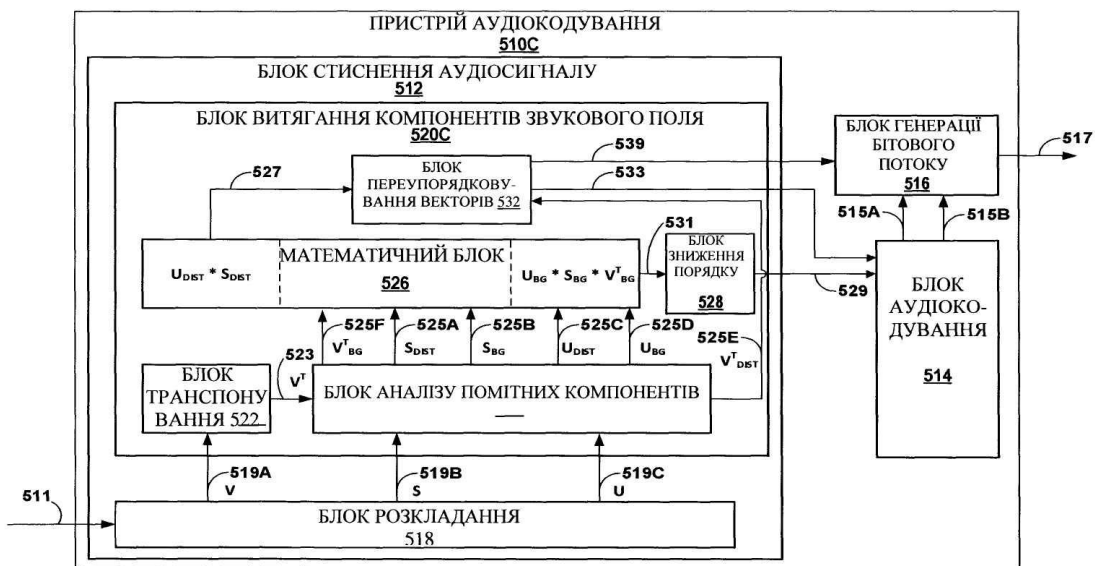


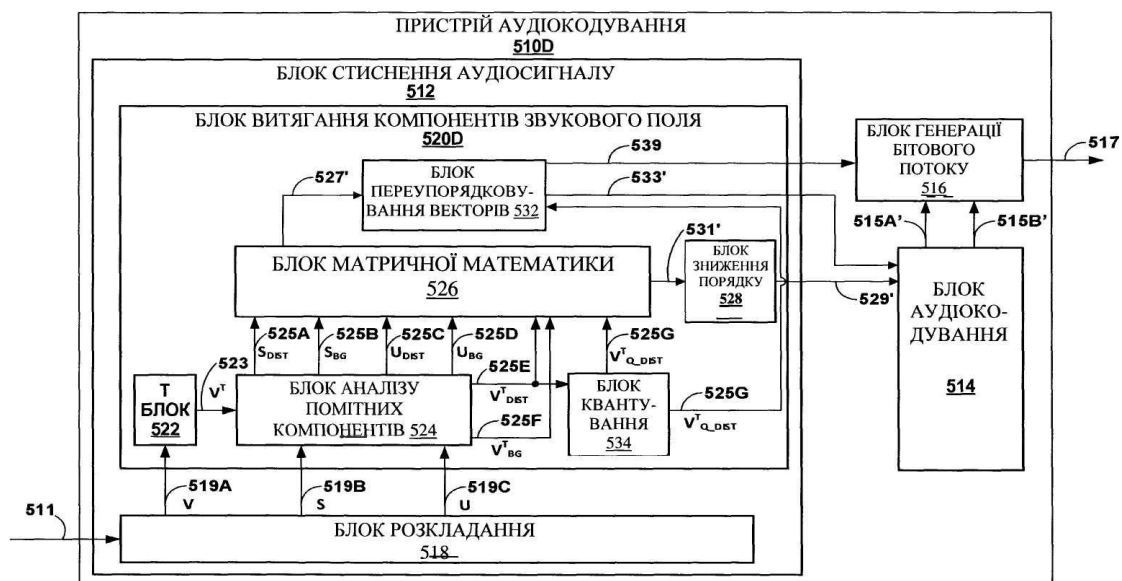
Fig. 40A



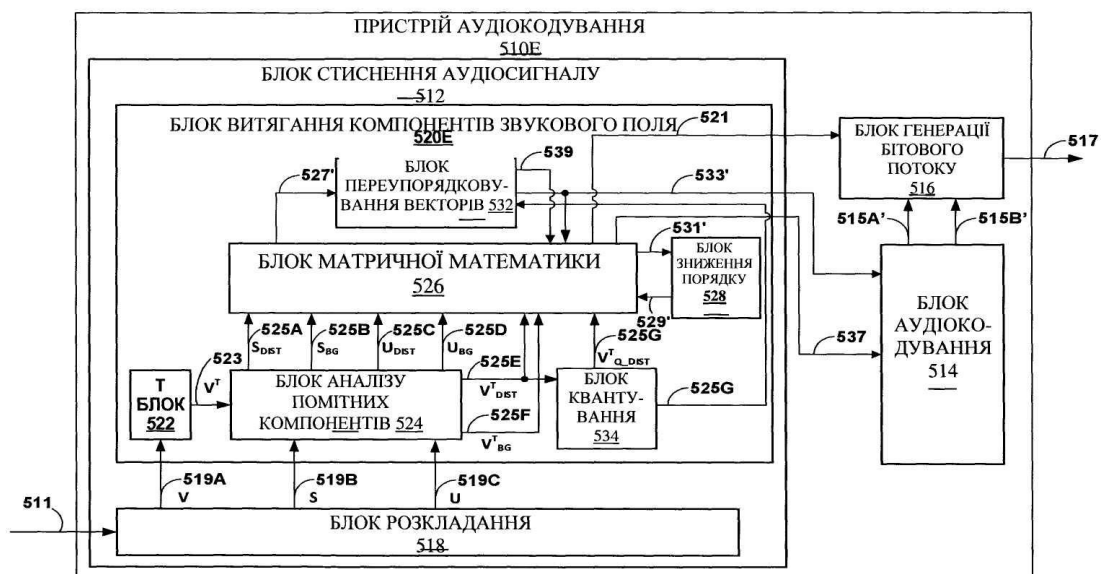
Фіг. 40B



Фіг. 40C



Фіг. 40D



Фіг. 40E

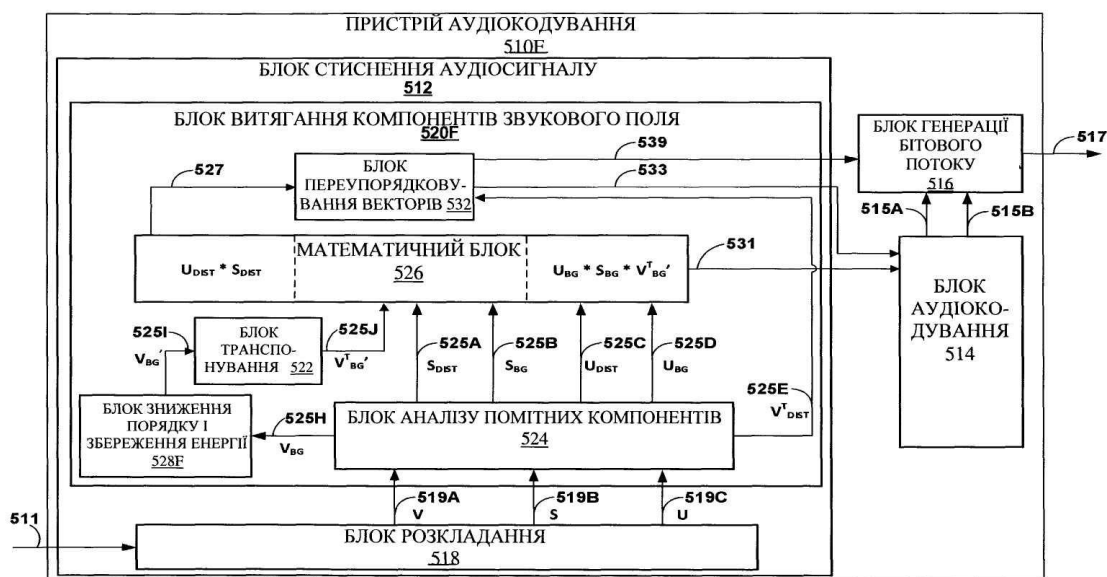


Fig. 40F

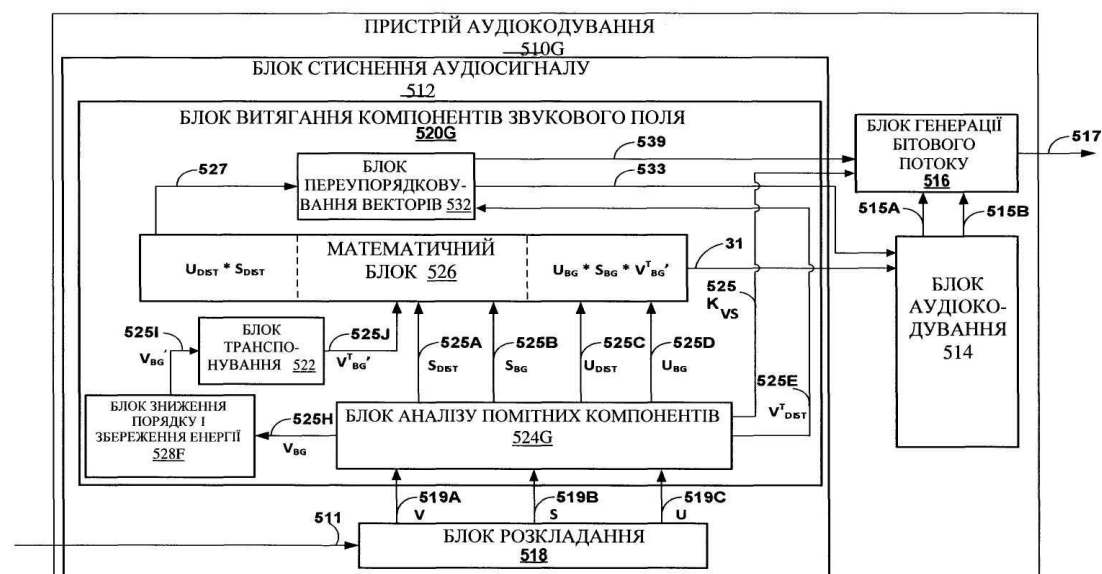


Fig. 40G

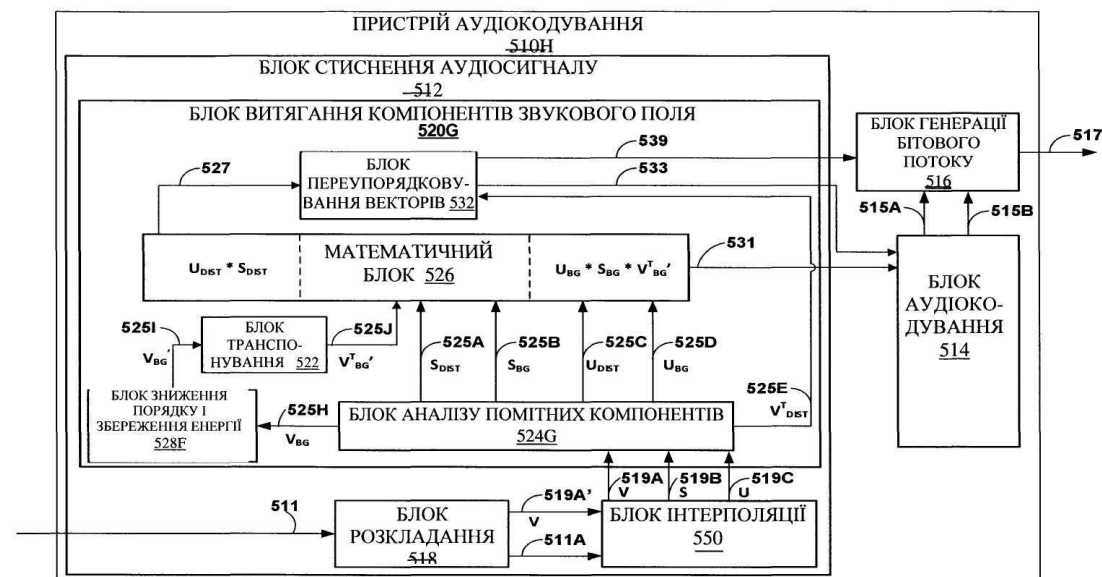


Fig. 40H

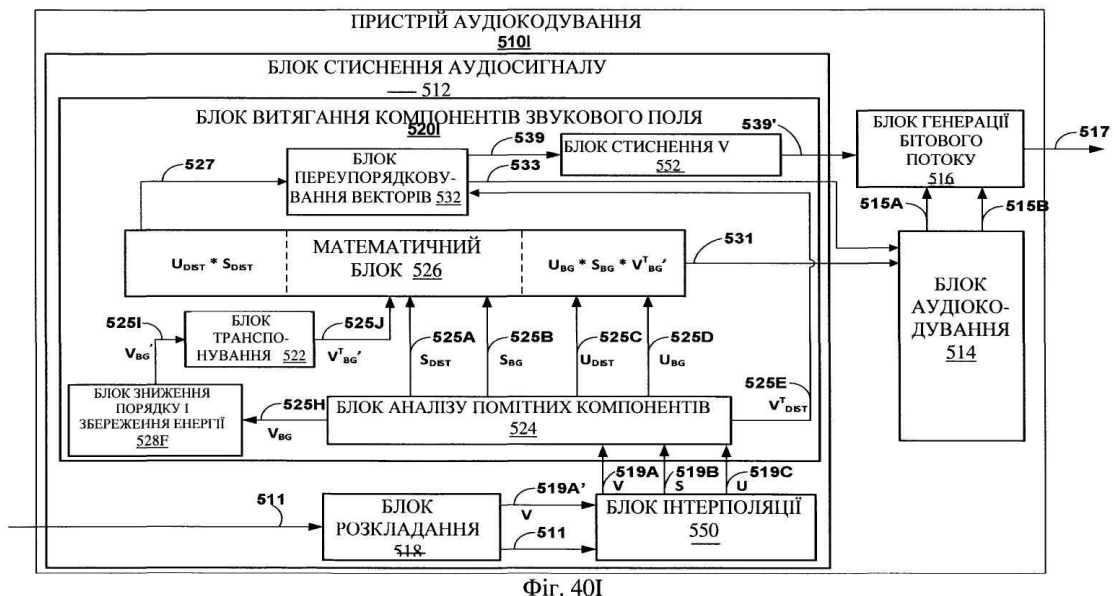
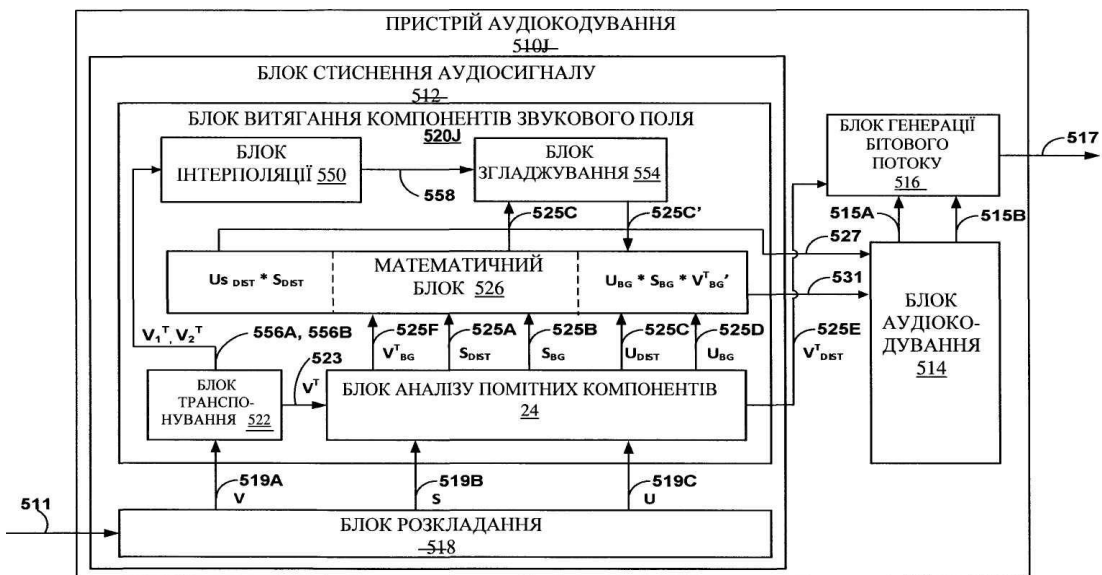


Fig. 40I



Фиг. 40J

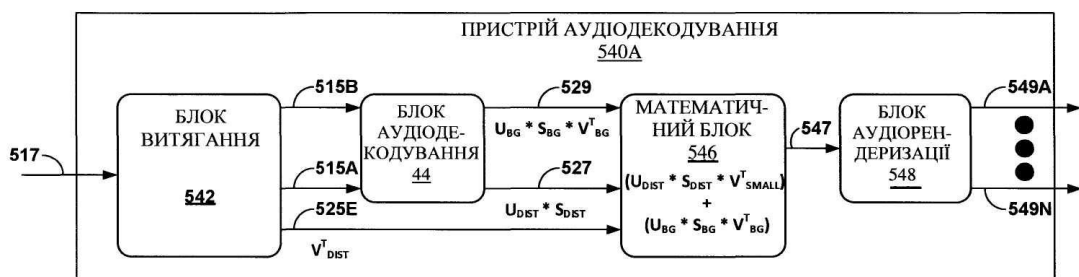
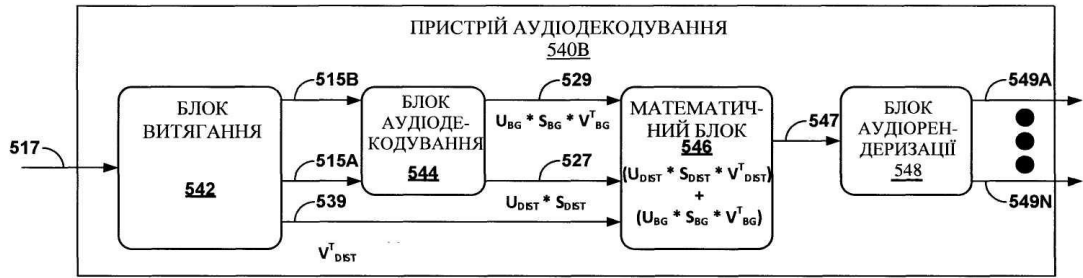
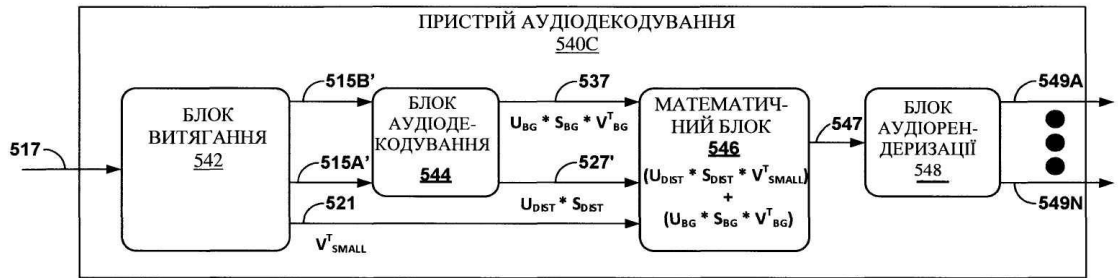


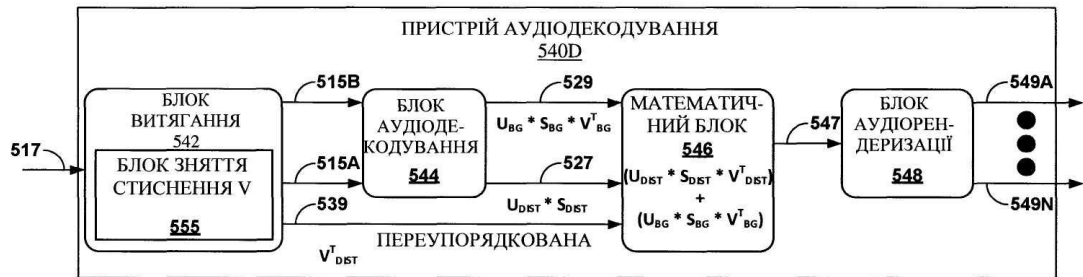
FIG. 41A



Фіг. 41B



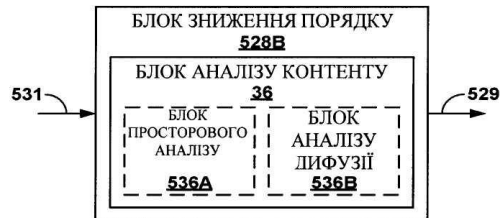
Фіг. 41C



Фіг. 41D



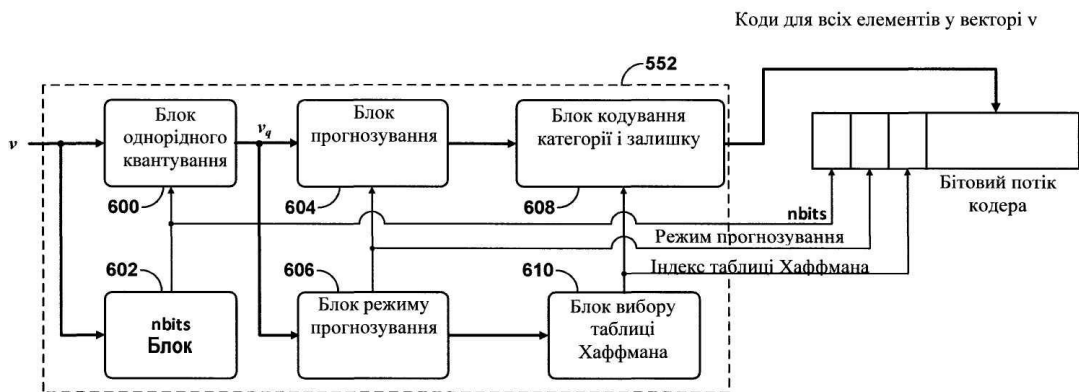
Фіг. 42А



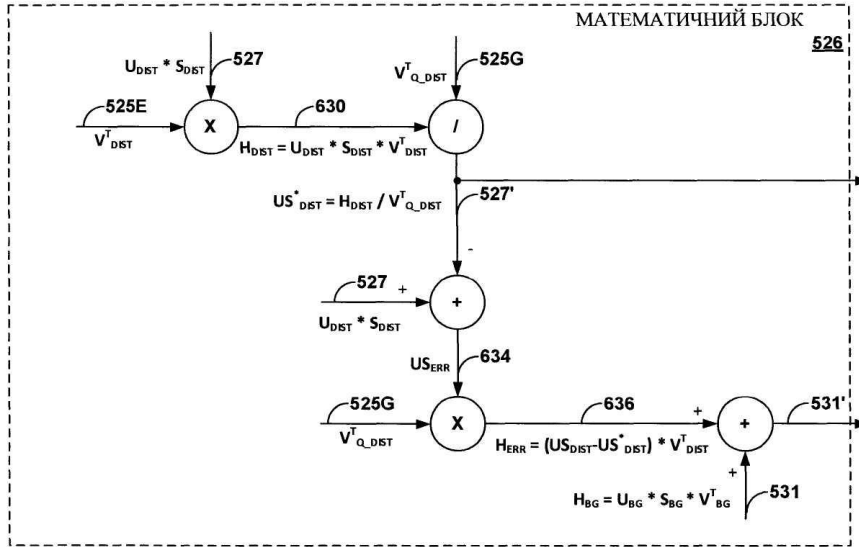
Фіг. 42В



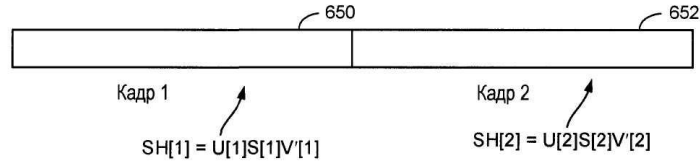
Фіг. 42С



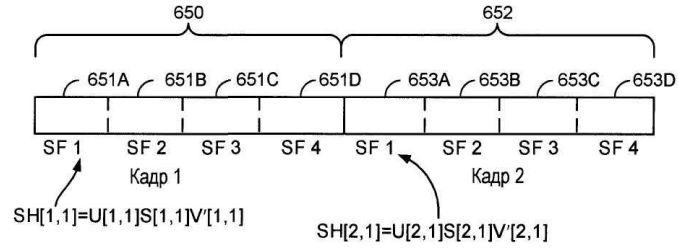
Фіг. 43



Фиг. 44

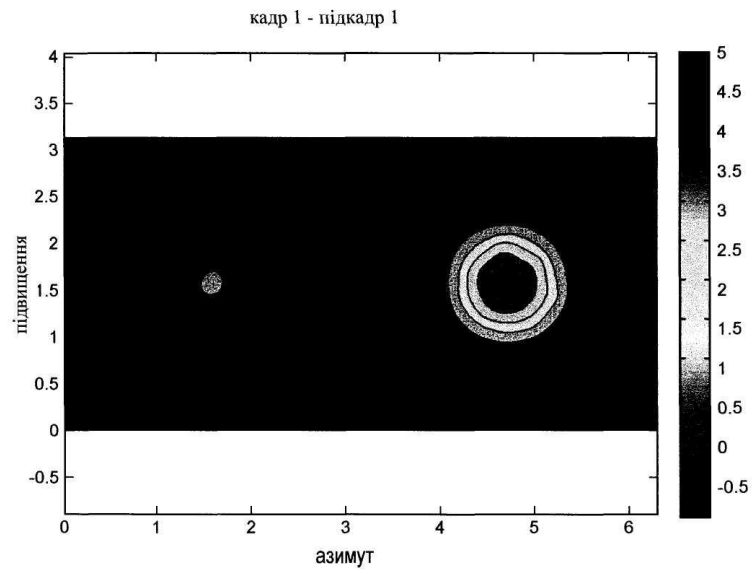


Фиг. 45A

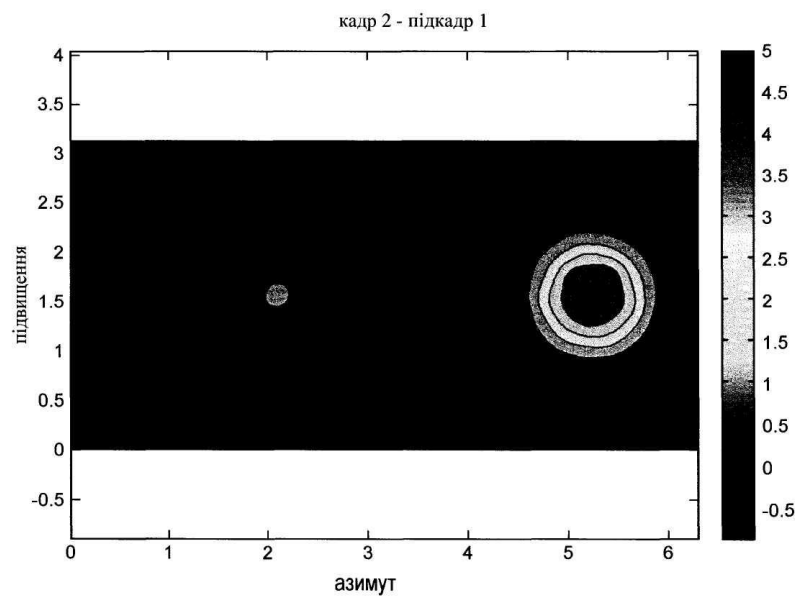


$$\begin{aligned}
 V[1,2] &= \text{interpolation}(V[1,1], V[2,1]) \rightarrow U[1,2]S[1,2] = SH[1,2](V[1,2])^{-1} \\
 V[1,3] &= \text{interpolation}(V[1,1], V[2,1]) \rightarrow U[1,3]S[1,3] = SH[1,3](V[1,3])^{-1} \\
 V[1,4] &= \text{interpolation}(V[1,1], V[2,1]) \rightarrow U[1,3]S[1,3] = SH[1,3](V[1,3])^{-1}
 \end{aligned}$$

Фиг. 45B



Фіг. 46А



Фіг. 46В

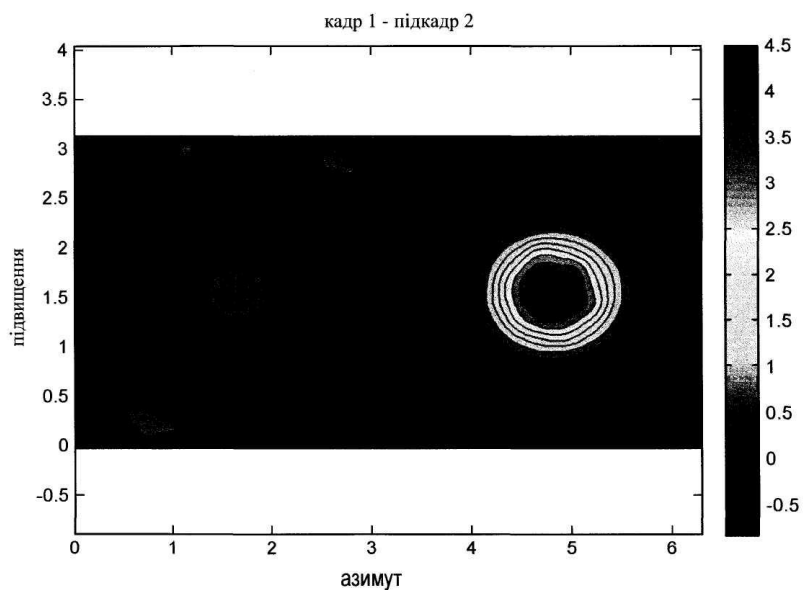


Fig. 46C

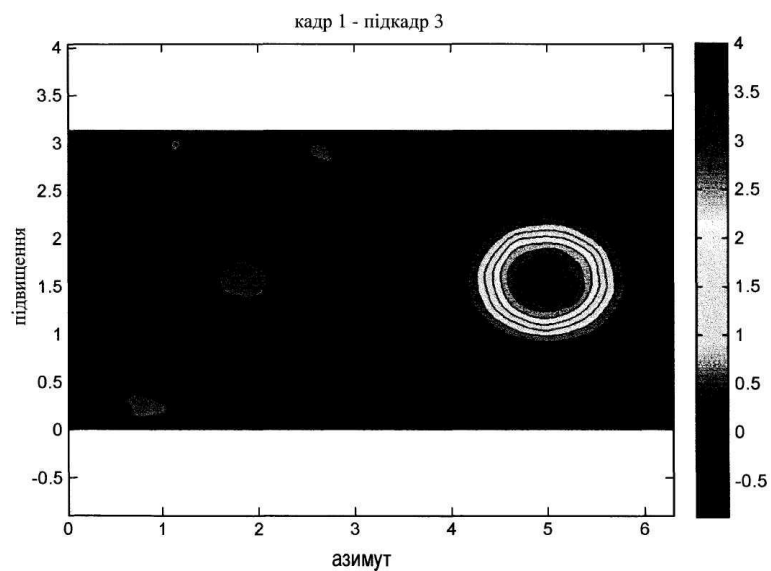


Fig. 46D

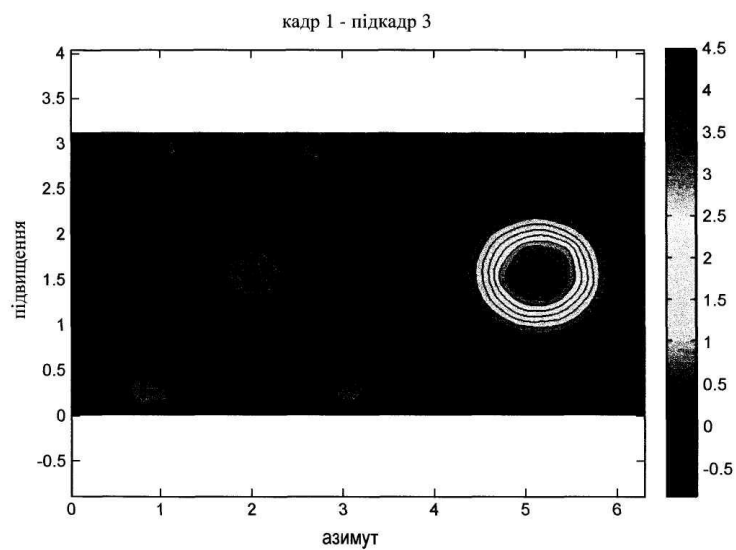
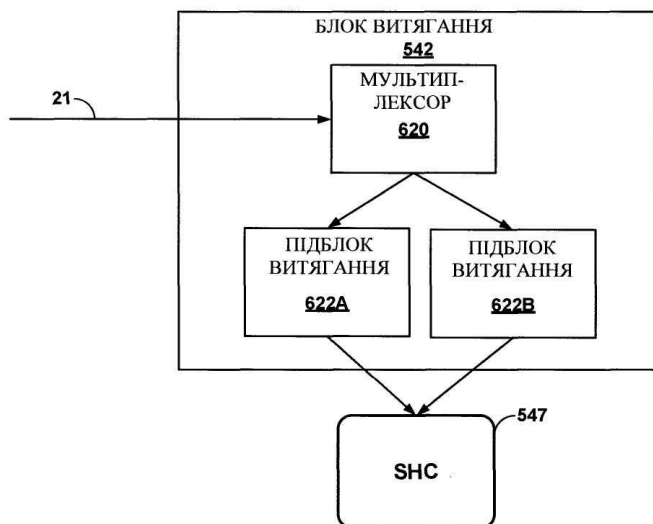


Fig. 46E



Фиг. 47



Фиг. 48

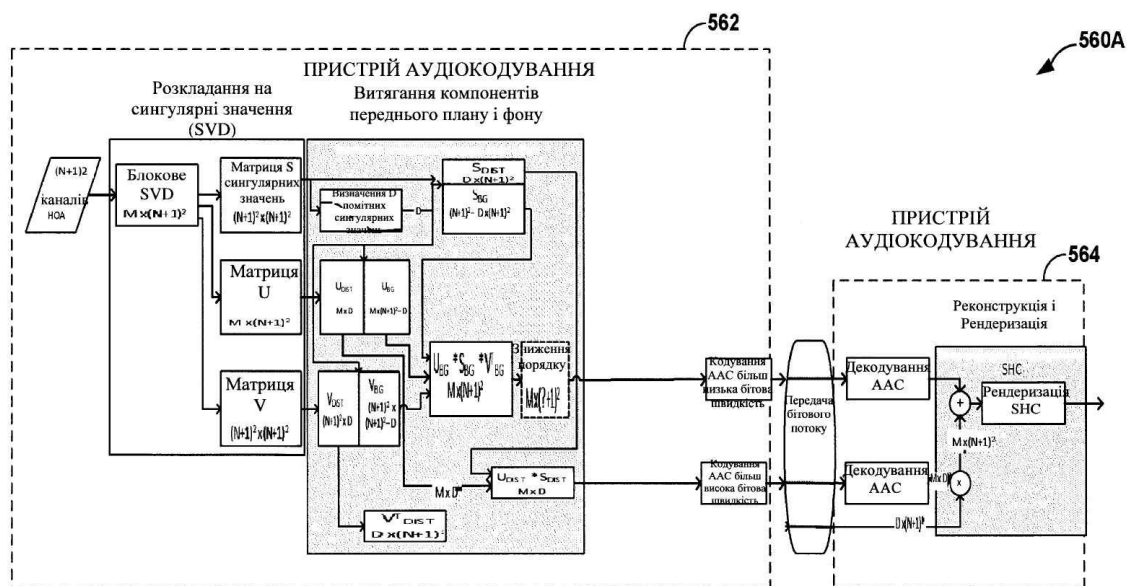
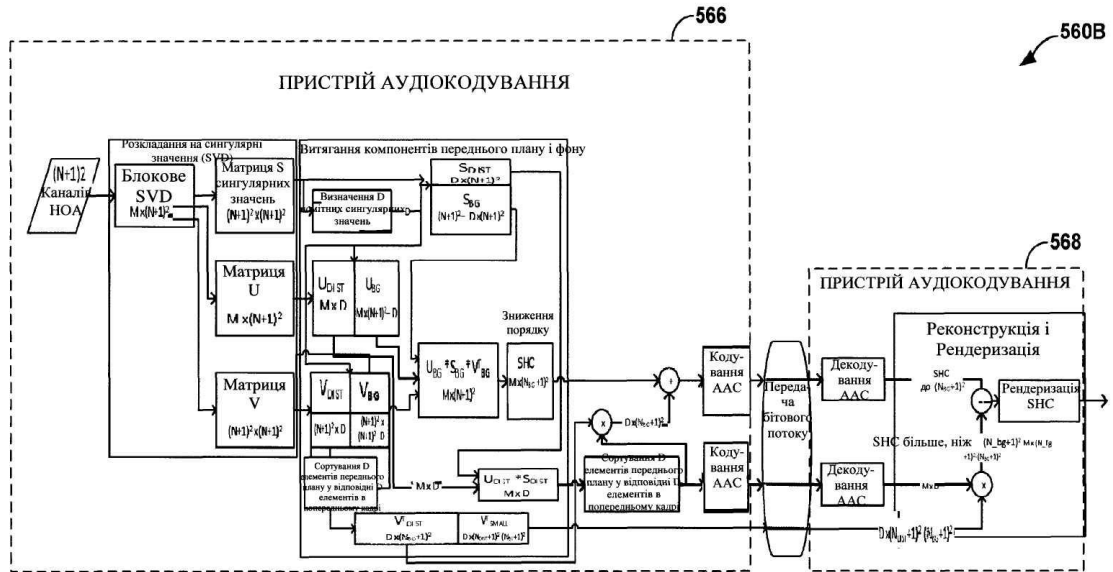
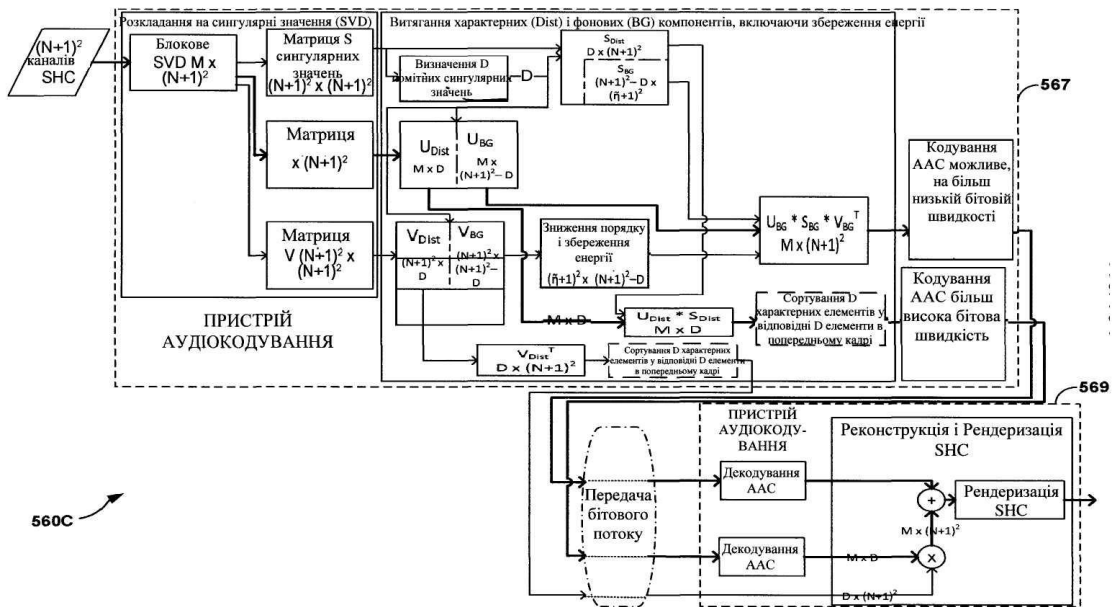


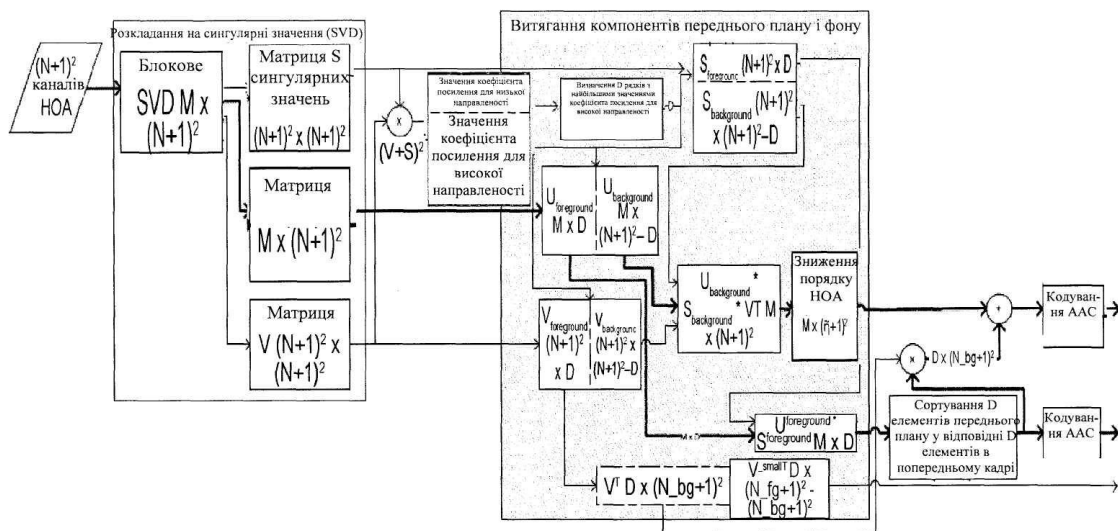
Fig. 49A



Фіг. 49B

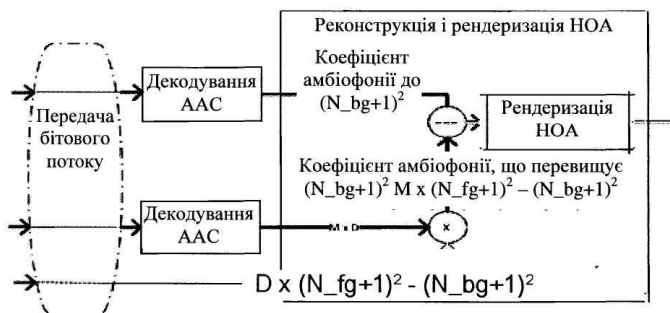


Фіг. 49C



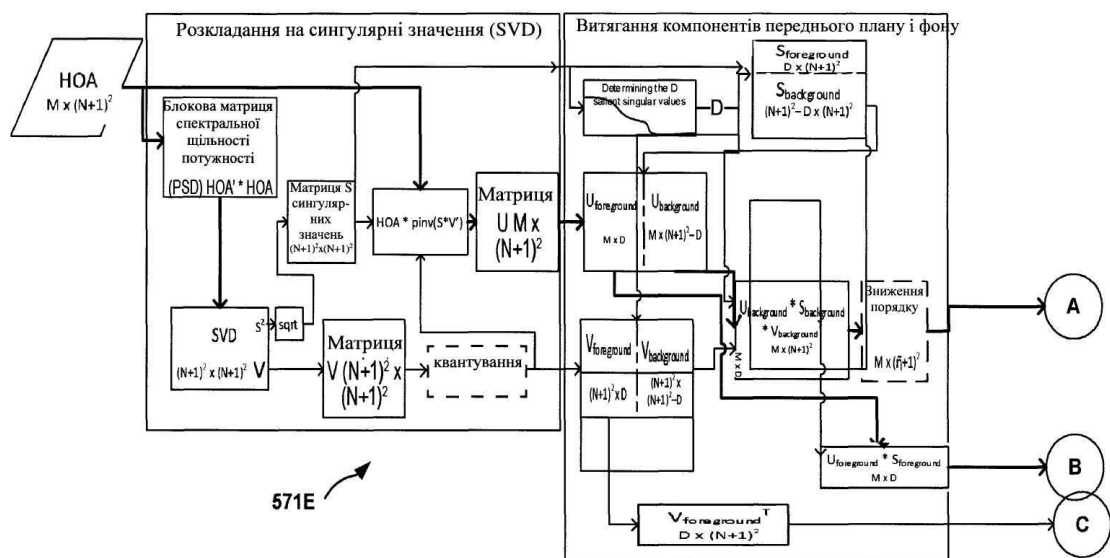
567D

Фіг. 49D(i)

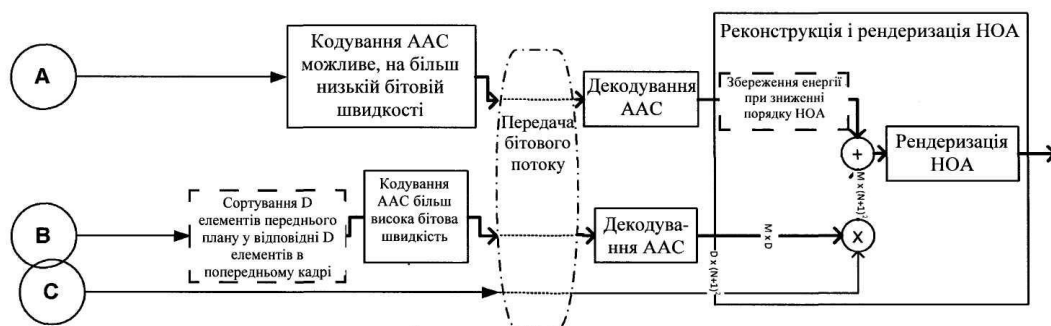


569D

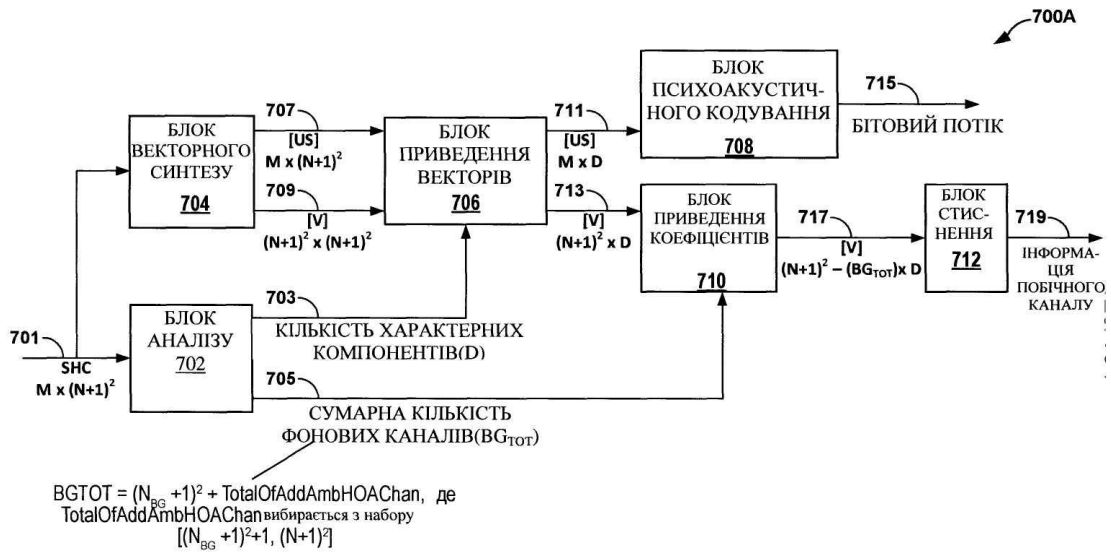
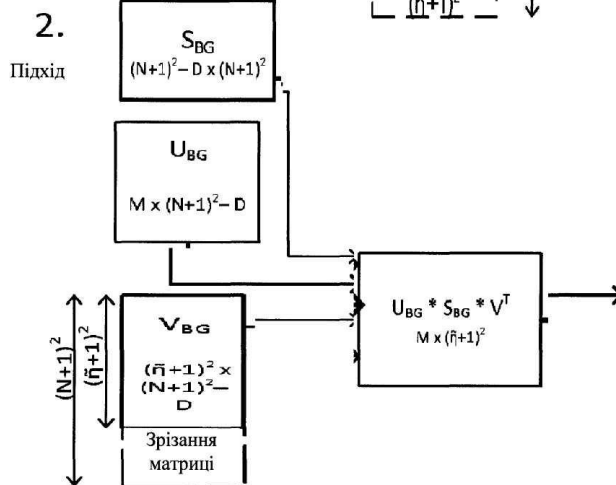
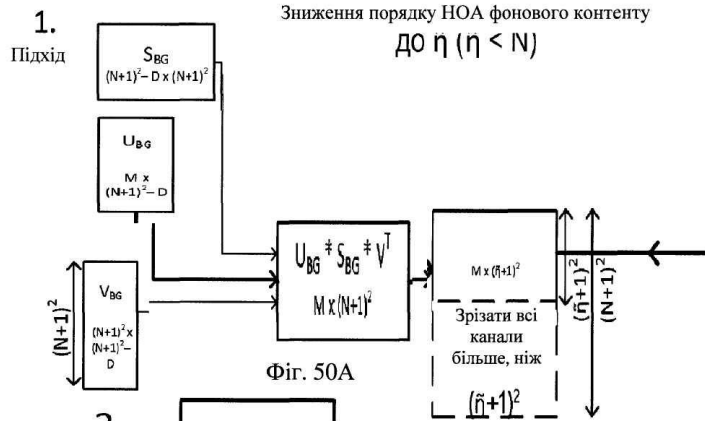
Фіг. 49D(ii)



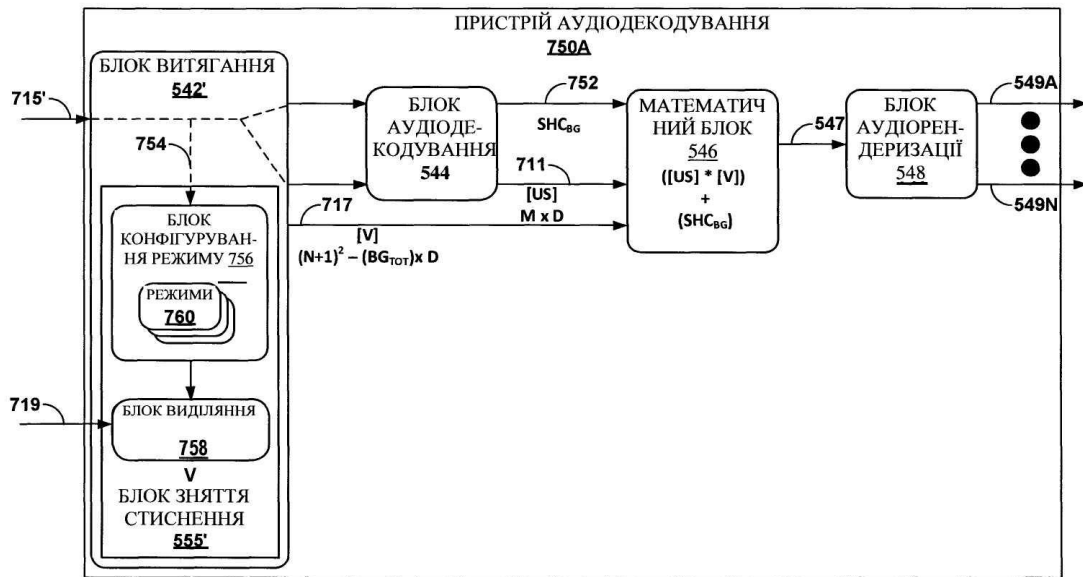
Фіг. 49E(i)



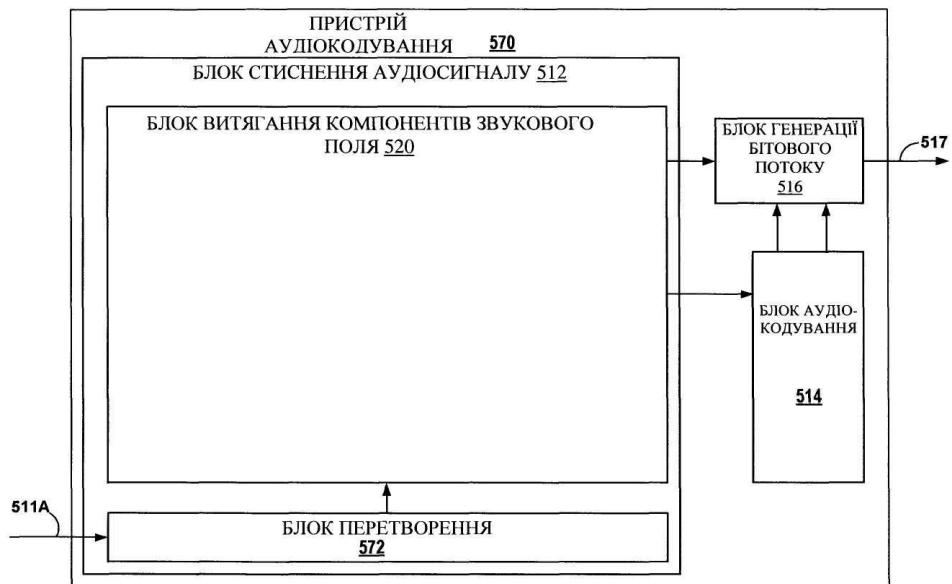
Фіг. 49E(ii)



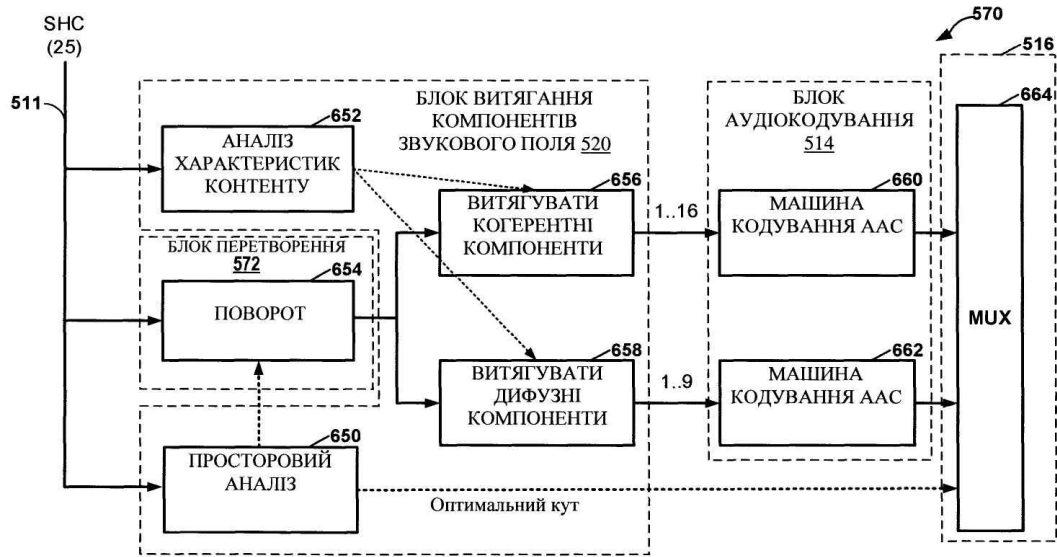
Фіг. 51



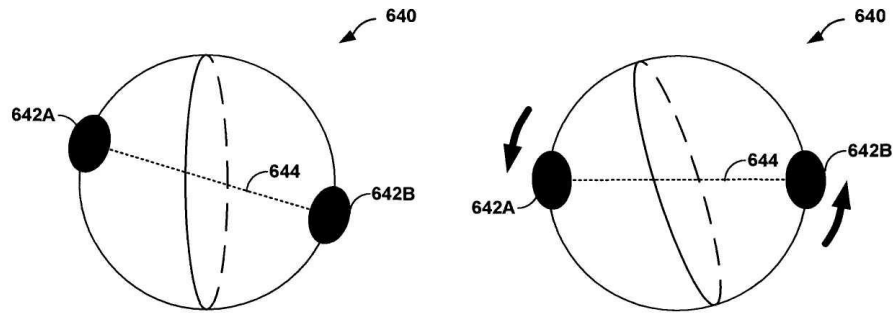
Фіг. 52



Фіг. 53

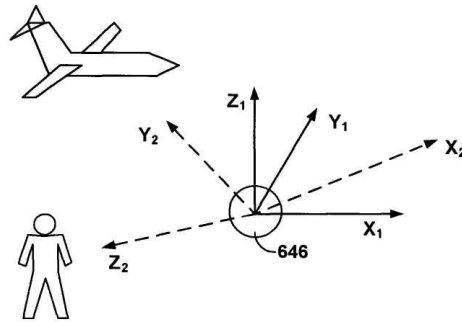


Фіг. 54



Фіг. 55A

Фіг. 55B



Фіг. 56

ПОЛЕ ПРИСУТНОСТІ SHC 670	SHC 511'
-----------------------------	-------------

517A

Фіг. 57A

ІНФОРМАЦІЯ ПЕРЕТВОРЕННЯ 672	SHC 511'
--------------------------------	-------------

517B

Фіг. 57B

ІНФОРМАЦІЯ ПЕРЕТВОРЕННЯ 672	ПОЛЕ ПРИСУТНОСТІ SHC 670	SHC 511'
--------------------------------	-----------------------------	-------------

517C

Фіг. 57C

0	7	32	33	34	44	53	53+X
ПОРЯДОК 674	ПОЛЕ ПРИСУТНОСТІ SHC 670	AZF 676	ELF 678	АЗИМУТ 680	ПІДВИЩЕННЯ 682	SHC 511'	

517D

Фіг. 57D

0	7	32	52	53+X
ПОРЯДОК 674	ПОЛЕ ПРИСУТНОСТІ SHC 670	ПОЛЕ ІНДЕКСУ ПОВОРОТУ 684	SHC 511'	

517E

Фіг. 57E

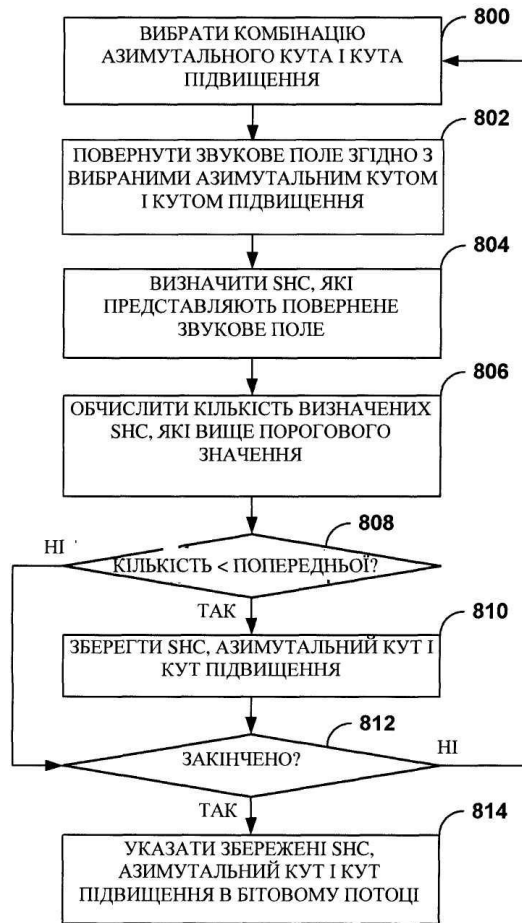


Fig. 58

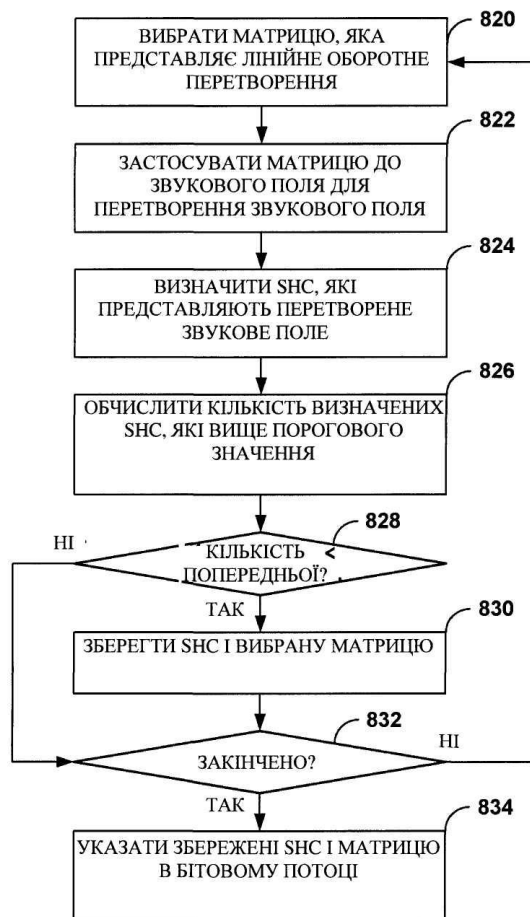


Fig. 59