



УКРАЇНА

(19) UA (11) 48220 (13) U
(51) МПК (2009)
G10L 15/00МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИДЕРЖАВНИЙ ДЕПАРТАМЕНТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІОПИС
ДО ПАТЕНТУ
НА КОРИСНУ МОДЕЛЬвидається під
відповідальність
власника
патенту

(54) ГОЛОСОВИЙ ЕЛЕКТРОННИЙ ДОВІДНИК

1

(21) u200909607

(22) 21.09.2009

(24) 10.03.2010

(46) 10.03.2010, Бюл.№ 5, 2010 р.

(72) ВІНЦЮК ТАРАС КЛИМОВИЧ, ГРИЦЕНКО ВО-
ЛОДИМИР ІЛЛІЧ, ПАВЛОВ ОЛЕГ ІГОРЕВИЧ,
СТАСЕВИЧ ПЕТРО АНАТОЛІЙОВИЧ, ТЕРТИЧ-
НИЙ ГРИГОРІЙ МИКОЛАЙОВИЧ(73) МІЖНАРОДНИЙ НАУКОВО-НАВЧАЛЬНИЙ
ЦЕНТР ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА СИС-
ТЕМ(57) Голосовий електронний довідник, що містить
мікрофон, з'єднаний з аналого-цифровим перет-
ворювачем, цифро-аналоговий перетворювач,

2

з'єднаний з динаміком, мікроконтролер, входи яко-
го підключені до кнопки керування пошуком та
кнопки керування іншими функціями, до входів і
виходів якого також підключена своїми виходами і
входами енергонезалежна пам'ять, акумулятор,
порт приєднання персонального комп'ютера, який
відрізняється тим, що він містить процесор циф-
рової обробки сигналів для розпізнавання та по-
шуку мовленнєвого фрагмента, одним із входів
підключений до аналого-цифрового перетворюва-
ча, другим входом - до мікроконтролера, одним із
виходів підключений до цифро-аналогового перет-
ворювача, а другим виходом підключений до одно-
го із входів мікроконтролера.

Модель відноситься до галузі розпізнавання
мовленнєвих сигналів і призначений для автома-
тичного пошуку інформації в масивах аудіозаписів
за ключовими словами або словосполученнями.

До аналогів належать пристрої, в яких способи
пошуку інформації в аудіофайлах виконуються
завдяки індексним позначкам. На відміну від тек-
стових файлів, пошук в звукових файлах значно
складніший і вимагає відтворення запису аж поки
бажаний фрагмент не буде знайдений. Відомі спо-
соби пошуку на основі індексних позначок, за яки-
ми пошук всередині звукозапису виконується за
попередньо визначеними індексами (адресами).
Але, незважаючи на значне прискорення пошуку, в
цьому способі залишаються ряд недоліків. По-
перше, пошук обмежений кількістю позначок в
середині аудіозапису, тому завжди залишаються
не позначені фрагменти; по-друге, при великій
кількості позначок зростає й час на вибір потрібної
позначки; по-третє, метод індексних позначок
вимагає певного часу на попередню розмітку за-
пису.

За прототип моделі обраний диктофон OLYM-
PUS DS-330 (www.olympus.com). OLYMPUS DS-
330, який дозволяє записувати звукові фрагменти
у вигляді файлів та встановлювати до 16
індексних позначок (index marks) в середині
файлів.

Недоліками прототипу є: обмеження пошуку
лише позначеними звуковими фрагментами; об-

межена кількість індексних позначок; необхідність
попередньої розмітки аудіозапису.

В основу моделі покладено задачу вдоскона-
лення процесу пошуку мовленнєвого фрагмента в
масиві аудіозапису, шляхом реалізації так званої
технології автоматичного пошуку та розпізнавання
ключового слова або словосполучення, яка
полягає в виконанні параметричного перетворення
звукового сигналу на послідовність наборів
параметрів обраної моделі мовотворення, її
стискання з втратами за допомогою процедури
динамічного програмування, її використанні для
розпізнавання схожих за обраною мірою
відмінності аналогічних послідовностей з засто-
суванням динамічного програмування, пошуку мож-
ливих початків таких схожих послідовностей, їх
впорядкування за інтегральною мірою відмінності
та озвученні у відповідному порядку, що дозволяє
значно спростити та пришвидшити пошук
потрібного фрагмента запису.

Модель пропонує новий спосіб пошуку звуко-
вого фрагмента. Так, спосіб пошуку певного звуко-
вого фрагменту в пам'яті диктофона OLYMPUS
DS-330 залежить, чи позначений даний звуковий
фрагмент. Якщо фрагмент має індексну позначку
користувачеві достатньо вибрати її з переліку по-
значок. Якщо ж фрагмент не позначений, користу-
вач має уважно прослухати звукозапис, аж поки
потрібний фрагмент не буде відтворений.
Натомість, у способі, який пропонується винахо-
дом, користувач має промовити будь-яке слово

(19) UA (11) 48220 (13) U

або словосполучення що є у звуковому фрагменті, який шукається, і відтворення автоматично розпочнеться з фрагмента, який містить сказане слово або словосполучення.

Спосіб автоматичного пошуку мовленнєвого фрагмента в суцільному масиві звукозапису базується на представленні сказаного ключового слова або словосполучення у вигляді послідовності наборів параметрів обраної моделі мовотворення, кожен з яких розглядається як багатомірний вектор параметрів, здійсненні процедури самосегментації отриманої послідовності векторів параметрів за допомогою процедури динамічного програмування, стискання початкової послідовності векторів параметрів шляхом апроксимації кожного з отриманих сегментів (які містять певну кількість початкових векторів, мінімальне і максимальне значення якої визначається обраним правилом самосегментації) за допомогою одного для всього сегменту апроксимуючого вектору параметрів, порівнянні з використанням процедури динамічного викривлення часу отриманої стиснутої послідовності апроксимуючих векторів параметрів ключового слова або словосполучення з кожним можливим початком подібної часткової послідовності в не стиснутій послідовності векторів параметрів, яка утворюється під час подібного перетворення суцільного масиву звукозапису з метою розпізнавання ключового слова або словосполучення, підрахунку інтегральної міри відстані для кожного з таких порівнянь, перевірки перевищення інтегральною мірою відстані встановленого порогу відмови від розпізнавання і, вразі його не перевищення фіксації інтегральної міри відстані разом з індексом початку такої часткової послідовності векторів параметрів суцільного масиву звукозапису. В залежності від режиму роботи після знаходження часткової послідовності, інтегральна міра відстані для якої від стиснутої послідовності апроксимуючих векторів параметрів ключового слова або словосполучення є нижчою за встановлений поріг відмови від розпізнавання може відбуватися озвучення суцільного масиву запису з відповідного місця, або продовження пошуку решти входжень стиснутої послідовності апроксимуючих векторів параметрів в не стиснутій послідовності векторів параметрів суцільного масиву звукозапису, після чого відбувається озвучення найкращого за інтегральною мірою відстані (тобто з найменшим її значенням) фрагменту фонограми. Якщо після озвучення фрагменту фонограми користувач натискає відповідну кнопку і подає команду продовжити пошук, то в залежності від режиму або відбувається пошук наступного за часом фрагменту в описаний вище спосіб, або озвучується той зі всіх знайдених фрагментів, який має найближчу вищу за значенням інтегральну міру відстані. Після того, як всі фрагменти озвучені, або якщо жодного фрагменту не знайдено користувач отримує відповідне повідомлення.

Пошук закінчується за певної кількості спроб, або якщо користувач вирішує переформулювати запит.

На фіг. 1 зображено послідовність дій по пошуку звукового фрагменту в суцільних масивах аудіозаписів. Звуковий сигнал ключового слова або словосполучення поступає на так званий аналізатор, в якому сигнал перетворюється на стиснуту послідовність апроксимуючих векторів параметрів. В блоці розпізнавання стиснуту послідовність апроксимуючих векторів параметрів порівнюється зі всіма можливими початками подібної послідовності в не стиснутих послідовностях векторів параметрів суцільних масивів аудіозаписів, які зберігаються разом з самими аудіозаписами в енергонезалежній пам'яті довідника. Створення таких не стиснутих послідовностей векторів параметрів відбувається в аналізаторі під час першого пошуку будь-якого ключового слова або словосполучення. Знайдений в результаті пошуку фрагмент озвучується з відповідного місця аудіозапису і до його кінця, або може бути перерваним, якщо результат пошуку не є задовільним. Користувач може продовжити пошук або продовжити прослуховування наступного знайденого фрагменту.

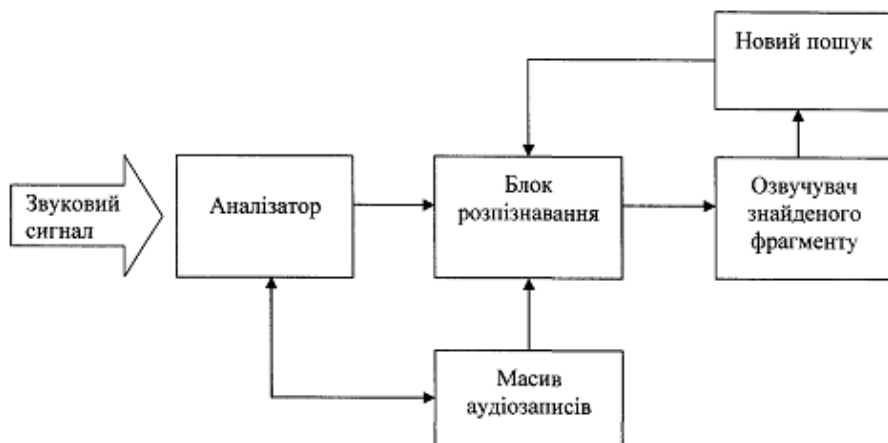
На фіг.2 зображена структурна схема голосового електронного довідника з голосовим пошуком. Довідник складається з наступних частин:

- 1 - мікрофона;
- 2 - аналого-цифрового перетворювача;
- 3 - мікроконтролера;
- 4 - процесора цифрової обробки сигналів;
- 5 - енергонезалежної пам'яті;
- 6 - цифро-аналогового перетворювача;
- 7 - кнопки керування пошуком;
- 8 - інших засобів керування та індикації;
- 9 - динаміка (або головних телефонів);
- 10 - виходу для з'єднання диктофона з комп'ютером;
- 11 - акумулятора.

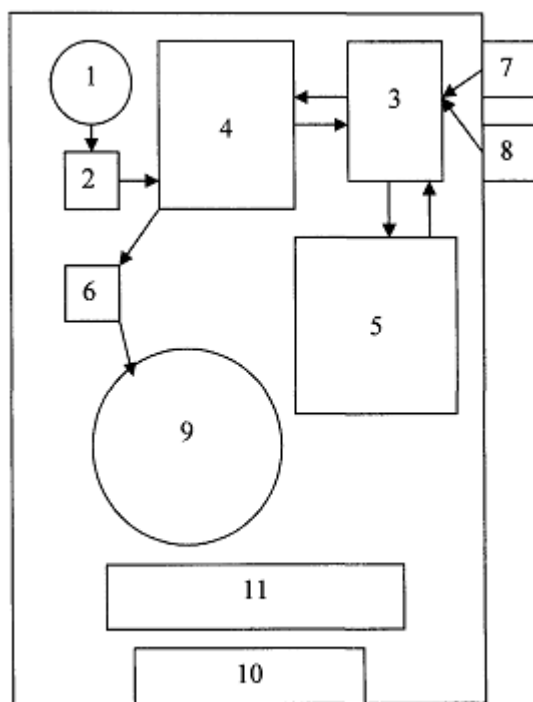
Мікрофон 1 приєднують до аналого-цифрового перетворювача 2. Аналого-цифровий перетворювач 2 підключають до процесора цифрової обробки сигналу 4. До процесора цифрової обробки сигналу 4 приєднують мікроконтролер 3 та цифро-аналоговий перетворювач 6, який має вихід на динамік або головні телефони 9. Мікроконтролер 3, окрім процесора, з'єднують з клавіатурою 7, засобами індикації 8 та зовнішньою пам'яттю 5. Живлення пристрою виконують від вбудованого акумулятора 11. Зв'язок з комп'ютером виконують через USB порт 10.

Звуковий сигнал надходить на вбудований мікрофон, потім поступає на аналого-цифровий перетворювач. Оцифрований звуковий сигнал надходить до процесора цифрової обробки сигналу й, в залежності від поточного режиму, або розпочинається пошук, або звуковий сигнал записується до зовнішньої пам'яті голосового довідника. Результат пошуку озвучується за допомогою динаміка або надходить до аудіо-виходу головних телефонів.

Для активації пошуку користувач має натиснути клавішу керування пошуком, сказати ключове слово або словосполучення та відпустити клавішу. Знайдений фрагмент буде озвучений.



Фіг. 1



Фіг. 2