



УКРАЇНА

(19) **UA** (11) **51137** (13) **U**
(51) МПК (2009)
G10L 15/00МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИДЕРЖАВНИЙ ДЕПАРТАМЕНТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ**ОПИС**
ДО ПАТЕНТУ
НА КОРИСНУ МОДЕЛЬвидається під
відповідальність
власника
патенту**(54) СПОСІБ ГОЛОСОВОГО УПРАВЛІННЯ ЦИФРОВИМ ДИКТОФОНОМ**

1

(21) u200910873

(22) 28.10.2009

(24) 12.07.2010

(46) 12.07.2010, Бюл.№ 13, 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 СИСТЕМ

2

(57) Спосіб голосового управління цифровим диктофоном, який базується на генерації сигналів керування функціями запису, відтворення, пошуку, перейменування, перемотування, повтору та видалення звукових файлів, який **відрізняється** тим, що мовленнєвий сигнал перетворюють в цифрову форму, виділяють параметри розпізнавання, порівнюють з еталонними сигналами команд керування, виділяють найбільш близький по мірі схожості еталонний сигнал та по результатах порівняння генерують сигнал керування функціями диктофона.

Модель відноситься до галузі розпізнавання мовленнєвих сигналів і призначена для створення портативного цифрового диктофона, що керується голосом.

До аналогів належать способи управління цифровим диктофоном, що ґрунтуються на використанні кнопок та екрану. Недоліком такого способу керування є складна структура активації функцій диктофона, що є результатом обмеженості засобів керування в портативних пристроях. Іншим недоліком є ускладненість пошуку потрібного аудіозапису в пам'яті диктофона. Також стандартні способи керування обмежують можливості використання диктофонів для людей з вадами зору.

Відомо багато видів цифрових диктофонів, але диктофонів із голосовим способом керування функціями не має. Тому за прототип обраний спосіб керування цифровим диктофоном OLYMPUS DS-330 (www.olympus.com), що керується за допомогою одинадцяти кнопок керування та жидкокристалічного екрану. Управління диктофоном відбувається за допомогою меню, яке в OLYMPUS DS-330 налічує тринадцять пунктів. Відповідно, в кожному з пунктів меню змінюються функції кнопок керування й призначення спеціальних позначок на екрані.

Недоліком прототипу є багаторівневий процес керування функціями диктофона. Зважаючи на значну кількість пунктів меню, станів екрану й різні функції кнопок за різних пунктів меню, керування диктофоном вимагає неабиякої початкової підготовки. Іншим недоліком є ускладненість пошуку

файлів, адже OLYMPUS DS-330 здатен зберігати у пам'яті до 995 файлів. Причиною цих недоліків є стала технологія екранного, візуального керування, натомість дана модель пропонує інше, голосове керування функціями диктофона.

В основу моделі покладено задачу вдосконалення процесу керування функціями диктофона шляхом реалізації технології розпізнавання голосових команд, що дозволяє керувати функціями диктофона природною мовою та організувати файлову систему диктофона на основі мовленнєвих імен файлів.

Модель пропонує принципово новий спосіб керування функціями диктофона. Так, для відтворення запису в OLYMPUS DS-330, користувач за допомогою кнопок та екрану має обрати певний файл й натиснути кнопку відтворення. У моделі для програвання файлу користувач називає голосову команду та ім'я файлу. Для іменування файлу в OLYMPUS DS-330, користувач має знайти відповідний пункт меню та за допомогою кнопок ввести бажане ім'я (коментар). У моделі ж ім'я файлу промовляється голосом. Загалом, у моделі за голосовими командами виконуються функції запису, відтворення, пошуку, перейменування, перемотування, повтору та видалення звукових файлів.

Найважливішою новацією моделі є спосіб іменування та зчитування файлів за допомогою голосових імен. Такий спосіб ідентифікації файлів гарантує зручність, швидкість та простоту при пошуку потрібного звукового фрагмента.

(13) **U**
(11) **51137**
(19) **UA**

Керування функціями диктофона відбувається за допомогою голосових команд. Для цього користувач має назвати команду та/або ім'я файлу. Результат дії команди, підказка або виявлені помилки озвучуються у вигляді голосових повідомлень. Іменування файлу робиться користувачем під час або після виконання запису. Для цього користувачеві достатньо вимовити ім'я файлу. За цим ім'ям файл зберігається в файловій системі, за цим же ім'ям відбуваються й всі інші звернення до файлу. Наприклад, відтворення файлу виглядає так: користувач вимовляє команду "ПРОЧИТАТИ" та ім'я файлу. Для видалення файлу користувач вимовляє "СТЕРТИ ФАЙЛ", диктофон запитує ім'я файлу, користувач називає ім'я. Стирання вимагає підтвердження, тому диктофон запитує: "Ви дійсно бажаєте стерти цей файл?". Якщо користувач відповідає позитивно, тобто "ТАК", файл видаляється, в іншому разі не видаляється. Інший приклад, після промовляння команд "ПРОЧИТАТИ", "ЗАПИСАТИ", "СТЕРТИ ФАЙЛ" диктофон підказує голосом "Назвіть ім'я файлу для прочитання", "Назвіть ім'я файлу для запису" або інформує: "Файл стерто".

На Фіг.1 зображено послідовність дії команди голосового управління. Звуковий сигнал поступає на аналізатор, в якому сигнал перетворюється до форми, прийнятної для розпізнавання. В блоці розпізнавання перетворений сигнал порівнюється з еталонами команд, що зберігаються в енергонезалежній пам'яті диктофона, в результаті чого або генерується сигнал керування функціями диктофона або відбувається відмова від розпізнавання. У випадку, коли голосова команда потребує подальшого уточнення або підтвердження дії, відбувається діалог між пристроєм та людиною.

Спосіб голосового управління функціями диктофона, реалізований у моделі, дозволив значно переробити зовнішній вигляд цифрового диктофона, зменшити кількість кнопок керування до двох та навіть відмовитись від екрану, замінивши екранні повідомлення звуковими.

На Фіг.2 зображено структурну схему цифрового диктофона з голосовим управлінням. Диктофон складається з наступних частин:

1 - двох мікрофонів (для голосового керування та запису);

2 - аналого-цифрового перетворювача;

3 - мікроконтролера;

4 - процесора цифрової обробки сигналів;

5 - енергонезалежної пам'яті;

6 - цифро-аналогового перетворювача;

7 - кнопки запису/живлення;

8 - кнопки введення голосової команди;

9 - двох світлових індикаторів;

10 - динаміка або аудіо-виходу головних телефонів;

11 - виходу для з'єднання диктофона з комп'ютером (USB-порт);

12 - акумулятора.

Мікрофони 1 приєднують до аналого-цифрового перетворювача 2, при цьому один з мікрофонів активують в режимі запису, а інший в режимі розпізнавання голосових команд. 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 10. Мікроконтролер 3, окрім процесора, з'єднують з клавіатурою 7, 8, світловими індикаторами 9 та зовнішньою пам'яттю 5. Живлення пристрою виконують від вбудованого акумулятора 12. Зв'язок з комп'ютером виконується через USB порт 11.

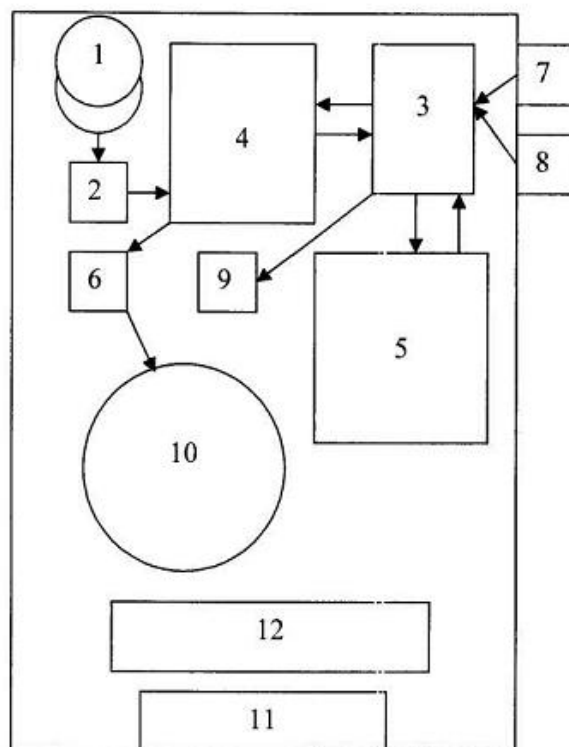
Звуковий сигнал надходить на вбудований мікрофон, потім поступає на аналого-цифровий перетворювач. Оцифрований звуковий сигнал надходить до процесора цифрової обробки сигналу А й, в залежності від поточного режиму, або розпізнається як голосова команда або записується до зовнішньої пам'яті диктофона. Записи та голосові повідомлення озвучуються за допомогою динаміка або надходять до аудіо-виходу головних телефонів.

Засоби керування диктофоном складаються з двох кнопок керування та світлових індикаторів. Початок або припинення роботи пристрою активується довгим натисненням кнопки живлення. Коротке натискання цієї кнопки активізує або припиняє режим запису. Друга кнопка призначена для вводу голосової команди. Світлові індикатори показують поточний режим.

Можливі різні модифікації моделі, зокрема збільшення кількості голосових команд або поєднання стандартних способів керування з елементами голосового управління. Це виглядає доцільним в складних системах з великим набором функцій.



Фіг. 1



Фіг. 2