



УКРАЇНА

(19) UA (11) 52350 (13) U
(51) МПК (2009)
G01L 15/00МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИДЕРЖАВНИЙ ДЕПАРТАМЕНТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІОПИС
ДО ПАТЕНТУ
НА КОРИСНУ МОДЕЛЬвидається під
відповідальність
власника
патенту

(54) ГОЛОСОВИЙ ФРАЗНИК-ПЕРЕКЛАДАЧ

1

2

(21) u201001607

(22) 16.02.2010

(24) 25.08.2010

(46) 25.08.2010, Бюл. № 16, 2010 р.

(72) ВІНЦЮК ТАРАС КЛИМОВИЧ, ГРИЦЕНКО ВО-
ЛОДИМИР ІЛЛІЧ(73) МІЖНАРОДНИЙ НАУКОВО-НАВЧАЛЬНИЙ
ЦЕНТР ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА СИСТЕМ

(57) Голосовий фразник-перекладач, що містить мікрофон, цифро-аналоговий та аналого-цифровий перетворювачі, динамік, процесор цифрової обробки сигналів, кнопки керування, індика-

тор та акумулятор, при цьому мікрофон з'єднаний через аналого-цифровий перетворювач з процесором цифрової обробки сигналу, який через цифро-аналоговий перетворювач з'єднаний з динаміком, мікроконтролер приєднаний до процесора цифрової обробки сигналу, індикатора, кнопок керування, а всі блоки фразника-перекладача з'єднані з акумулятором, який відрізняється тим, що містить енергонезалежну пам'ять, яка приєднана до мікроконтролера, а також порт приєднання до персонального комп'ютера, який приєднаний до мікроконтролера.

Модель відноситься до галузі розпізнавання мовленнєвих сигналів й призначена для усномовного перекладу фраз з однієї мови на іншу.

До аналогів моделі належать пристрої для перекладу текстів, введенних до комп'ютерної системи за допомогою клавіатури та екрану. Недоліком клавіатурно-екранних пристроїв є низька продуктивність усномовного перекладу, який вимагає введення тексту з клавіатури для його перекладу та збільшені розміри портативних пристроїв за рахунок клавіатури та екрану. До аналогів моделі також відноситься клас пристроїв усномовного перекладу, серед яких кишенькові комп'ютери зі спеціалізованим програмним забезпеченням та портативні пристрої серії Partner та Universal Translator та від ECTACO.

За прототип моделі обрано ECTACO Partner UT-203 (www.ectaco.com). Partner UT-203 є портативним пристроєм, що містить процесор цифрової обробки сигналів й має апаратну систему розпізнавання усномовних фраз. Пристрій призначений для перекладу усномовних фраз з російської та англійської мови на одну з трьох мов: англійську, російську або німецьку. Результат перекладу фрази виводиться на екран й може бути озвучений російською, німецькою, або англійською мовами. Partner UT-203 заявлений як багатодикторна система, тобто така, що не потребує попереднього налаштування на голос диктора. Розпізнавання на пристрої обмежується 3000 фразами, які розбито на 16 підкатегорій. Фрази для перекладу жорстко детерміновані, тобто зміна порядку слів у фразах

заборонена. Це обмеження є основним недоліком Partner UT-203.

Запропонована модель усуває цей недолік.

В основу моделі покладено задачу розширення можливостей усномовного перекладу фраз з однієї мови на іншу, шляхом реалізації автоматичного розпізнавання злитого мовлення, що складається зі слів заданих у вигляді їх фонетичних транскрипцій та ймовірностей слідування слів у фразах.

Поставлена задача вирішується голосовим фразником-перекладачем, що містить мікрофон, цифро-аналоговий та аналого-цифровий перетворювачі, динамік, процесор цифрової обробки сигналів, засоби керування та індикації, акумулятор, при цьому він містить енергонезалежну пам'ять, яка приєднана до мікроконтролера, що з'єднаний з процесором цифрової обробки сигналу та енергонезалежною пам'яттю, а також порт приєднання до персонального комп'ютера, який приєднаний до мікроконтролера.

В результаті розпізнавання обирається оптимальний порядок слів, що відповідає введеній усномовній фразі, або формується відмова від розпізнавання. Результат розпізнавання подається для перекладу до лінгвістичного аналізатора. Таким чином, модель пропонує новий, відсутній в прототипі, спосіб розпізнавання усномовних фраз, який не обмежується певним порядком слідування слів у фразах.

На Фіг.1 зображено спосіб усномовного перекладу, що пропонується даною моделлю. Оцифро-

(13) U
(11) 52350
(19) UA

ваний звуковий сигнал надходить до аналізатора, в якому перетворюється у сукупність векторів, що описують звуковий сигнал з прийнятною для розпізнавання точністю. Блок розпізнавання злитого мовлення в якості вхідних параметрів отримує словник фонетичних транскрипцій слів та ймовірності їх слідування у фразах. В цьому блоці виконується процедура вибору найкращого опису даного звукового сигналу послідовністю слів на основі заданої міри схожості. Обрана послідовність слів подається у лінгвістичний аналізатор, на виході з якого формується результат перекладу.

Словник транскрипцій слів та словосполучень формується у вигляді текстових записів, одиницею яких є текстова позначка фонем. Набір фонем є формальним описом слова або словосполучення. Кожній фонемі відповідає один або декілька векторів, що використовуються при обчисленні міри схожості. Визначення відповідності фонемі її векторному опису здійснюється на етапі навчання розпізнаванню, на основі знання про особливості мовлення диктора, так званого мовного паспорту. Для створення мовного паспорту диктор має вимовити певний текст, на основі якого система налаштується на його голос. Процедура налаштування на голос диктора, тобто створення його мовного паспорту, значно покращує результати розпізнавання у порівнянні з усередненими моделями, що є основою багатодикторних методів розпізнавання.

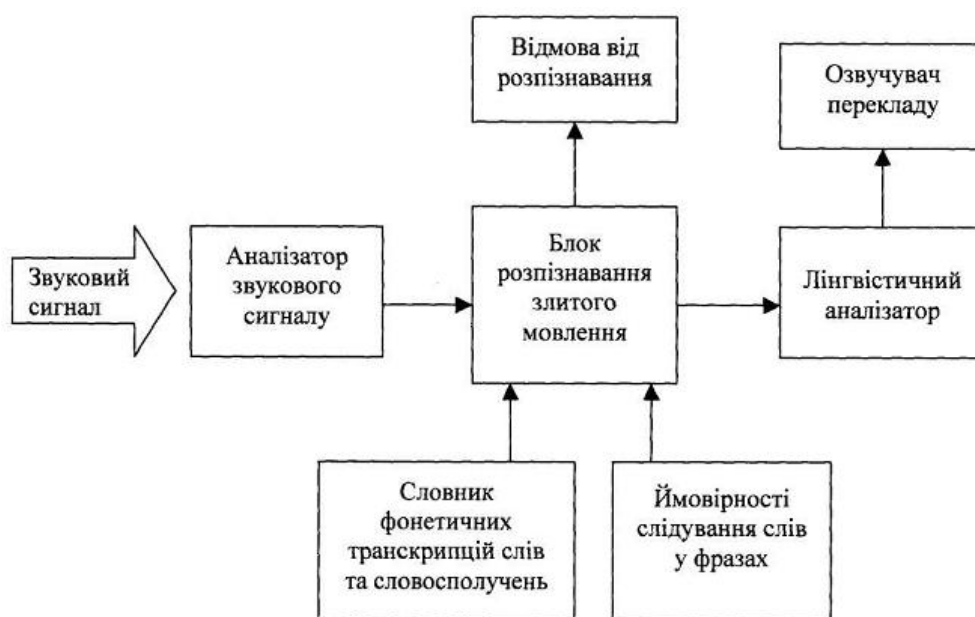
На Фіг.2 зображено структурну схему голосового фразника-перекладача, в якому реалізований поданий спосіб усномовного перекладу. Голосовий фразник-перекладач складається з наступних частин:

- 1 - мікрофона;
- 2 - аналого-цифрового перетворювача;
- 3 - мікроконтролера;
- 4 - процесора цифрової обробки сигналів;
- 5 - енергонезалежної пам'яті;

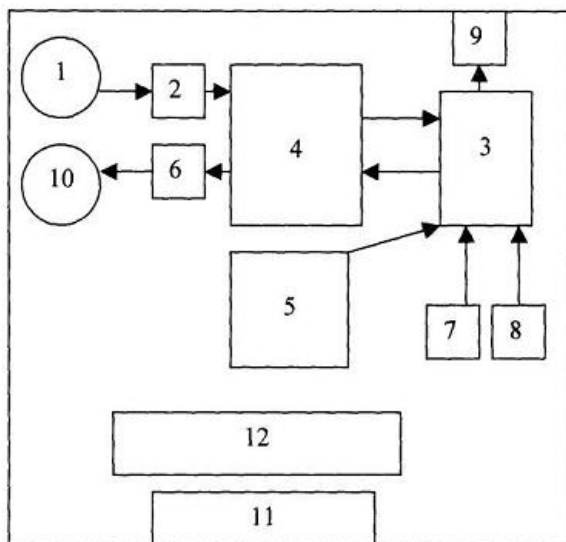
- 6 - цифро-аналогового перетворювача;
- 7 - кнопки включення живлення;
- 8 - кнопки введення усномовного сигналу для перекладу;
- 9 - двох світлових індикаторів;
- 10 - динаміка;
- 11 - виходу для з'єднання диктофона з комп'ютером;
- 12 - акумулятор.

Мікрофон 1 приєднаний до аналого-цифрового перетворювача 2, який з'єднаний з процесором 4 цифрової обробки сигналу. До процесора 4 приєднаний мікроконтролер 3, аналого-цифровий 2 та цифро-аналоговий 6 перетворювачі. Мікроконтролер 3 приєднаний до енергонезалежної пам'яті 5, клавіатури 7, 8, та засобів індикації 9. Живлення пристрою виконують за допомогою акумулятора 12. Для зв'язку з комп'ютером призначений USB-порт 11.

Введення усномовних фраз для перекладу активується натисненням кнопки введення усномовного сигналу. Звуковий сигнал з мікрофона 1 поступає в аналого-цифровий перетворювач 2. Цифрований сигнал подається в процесор 4 цифрової обробки сигналу, в якому сигнал обробляється та розпізнається. Результатом розпізнавання можуть бути переклад або відмова від розпізнавання (наприклад, якщо фраза відсутня в словнику транскрипцій). В залежності від результату розпізнавання, процесор 4 генерує сигнал мікроконтролеру 3, який обирає за заданим алгоритмом фразу для озвучення та викликає її із енергонезалежної пам'яті 5. Фраза для озвучення подається в процесор 4, декодується в прийнятну для озвучення форму, поступає в цифро-аналоговий перетворювач 6 та озвучується. Світлові індикатори 9 індексують режими, в яких працює пристрій: очікування, перекладу, озвучення, з'єднання з персональним комп'ютером тощо.



Фіг. 1



Фиг. 2