



УКРАЇНА

(19) UA (11) 28113 (13) U
(51) МПК (2006)
G10L 15/00

МІНІСТЕРСТВО ОСВІТИ
І НАУКИ УКРАЇНИ

ДЕРЖАВНИЙ ДЕПАРТАМЕНТ
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ

ОПИС ДО ПАТЕНТУ НА КОРИСНУ МОДЕЛЬ

видається під
відповідальність
власника
патенту

(54) ПРИСТРІЙ ДЛЯ ВИДІЛЕННЯ СКЛАДОВИХ СЕГМЕНТІВ У МОВНОМУ СИГНАЛІ

1

2

(21) u200707998

(22) 16.07.2007

(24) 26.11.2007

(72) БИКОВ МИКОЛА МАКСИМОВИЧ, UA, КОВТУН
ВЯЧЕСЛАВ ВАСИЛЬОВИЧ, UA, САВІНОВА
НАТАЛІЯ ГЕННАДІЇВНА, UA

(73) ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ
УНІВЕРСИТЕТ, UA

(56)

(57) Пристрій для виділення складових сегментів у мовному сигналі, що містить акустичний датчик, підключений до входу підсилювача сигналу, вихід якого підключений до блока спектрального аналізу, який містить два фільтри, а також формувач сегментних границь складів, який **відрізняється** тим, що, з метою підвищення надійності виділення складових сегментів у мовному сигналі і спрощення пристрою, в нього введено помножувач сигналів, диференціатор, першу і другу порогові схеми, першу і другу схему АБО, перший лічильний тригер з прямим динамічним керуванням, другий лічильний тригер зі зворотним динамічним керуванням, схему I, причому блок спектрального аналізу містить два смугових фільтри, що настроєні на діапазони частот, в яких зосереджена основна енергія голосних звуків, амплітудні детектори, а також помножувач сигналів, а формувач сегментних границь складів містить диференціатор, першу схему АБО, перший лічильний тригер з прямим

динамічним керуванням, другий лічильний тригер зі зворотним динамічним керуванням, схему I, другу схему АБО, причому вхід першого детектора підключений до виходу першого смугового фільтра, а його вихід з'єднаний з першим входом помножувача сигналів, вхід другого детектора підключений до виходу другого смугового фільтра, а його вихід з'єднаний з другим входом помножувача сигналів, вхід другого детектора підключений до виходу другого смугового фільтра, а його вихід з'єднаний з другим входом помножувача сигналу, вихід помножувача сигналу приєднаний до входу диференціатора формувача сегментних границь складів, вихід якого підключений до перших входів першої і другої порогових схем, напруга на других входах яких визначає відповідні порогові пропускання, вихід першої порогової схеми з'єднаний з першим входом першої схеми АБО, вихід другої порогової схеми з'єднаний з другим входом першої схеми АБО, вихід першої схеми АБО підключений до виходу першого лічильного тригера з прямим динамічним керуванням, до входу другого лічильного тригера зі зворотним динамічним керуванням і до першого входу другої схеми АБО, прямі виходи лічильних тригерів з'єднані зі входами схеми I, вихід схеми I з'єднаний з другим входом другої схеми АБО і є другим виходом пристрою, а вихід другої схеми АБО є першим виходом пристрою.

Корисна модель відноситься до галузі автоматики і обчислювальної техніки, а більш конкретно - до автоматичного розпізнавання мовних образів.

Відомий пристрій для виділення сегментів мовного сигналу [№А.С. СРСР №226310, М. кл. G06k, б. №28, 1968 рік], що містить відтворюючу магнітну головку, установлену з можливістю обертання щодо носія магнітного запису й приєднану до підсилювача, вихід якого з'єднаний з блоком відображення. Відомий пристрій забезпечує виділення сегментів мовного сигналу за рахунок багаторазового переміщення

відтворюючої магнітної головки щодо однієї і тієї самої ділянки носія магнітного запису.

Недоліком подібного пристрою є значна складність процесу виділення сегментів мовного сигналу й невисока точність.

Відомий пристрій для виділення сегментів мовного сигналу [№А.С. СРСР №684599, М. кл. G10L1/02, б. №33, 1979 рік], що містить відтворюючу магнітну головку, установлену з можливістю обертання щодо носія магнітного запису й приєднану до підсилювача, вихід якого підключений до входу кліпувального елемента і з'єднаний з блоком візуального відображення

UA (19) 28113 (13) U

через комутатор, приєднаний до виходу формувача строфа, підключеного до виходу кліпувального елемента й до виходу датчика початку циклу й приєднаного через блок установки початку й кінця строфа до одного з виходів генератора сигналу еталонних частот.

Недоліками даного пристрою є його значна технічна складність, а також необхідність попереднього запису мовного сигналу на магнітофон, що робить пристрій непридатним у системах розпізнавання мови в реальному масштабі часу.

З відомих пристроїв найбільш близьким по технічній сутності до запропонованого рішення є пристрій для виділення складових сегментів у мовному сигналі [Ruske Ct, Schotola F. An approach to speech recognition using syllabic decision units. - Proc. 1978 IEEE ICASSP, Tulsa, 1978. - N.Y., 1978, pp.772-725]. Він містить акустичний датчик, підключений до входу підсилювача сигналу, вихід якого підключений до блоку спектрального аналізу, що містить двадцять два смугових фільтри, виходи яких з'єднані з входами детекторів огинаючих сигналів, виходи яких підключені до входу суматора, вихід якого з'єднаний з входом формувача сегментних границь складів, що містить схеми виділення максимумів і мінімумів сигналу.

До недоліків даного пристрою виділення складових сегментів варто віднести апаратурні витрати, обумовлені необхідністю виділення амплітуд сигналів в 22-х частотних діапазонах, а також наявністю помилкових сегментів складів на ділянках високо енергетичних гучних звуків.

В основу корисної моделі поставлено задачу створення пристрою для виділення складових сегментів у мовному сигналі, в якому за рахунок введення нових елементів та зв'язків досягається можливість підвищити надійність сегментації мовного сигналу на складові сегменти і визначити такі їх ознаки, як тривалість, місцеположення і кількість. Розроблений пристрій може бути застосований при розробці автономних систем розпізнавання мови.

Поставлена задача досягається тим, що в пристрій для виділення складових сегментів у мовному сигналі, що містить акустичний датчик, підключений до входу підсилювача сигналу, вихід якого підключений до блока спектрального аналізу, а також формувач сегментних границь складів, введено помножувач сигналів, а блок спектрального аналізу містить два смугових фільтри, що настроєні на діапазони частот, в яких зосереджена основна енергія голосних звуків, перший і другий амплітудні детектори, а також помножувач сигналів, а формувач сегментних границь складів містить диференціатор, першу схему АБО, перший лічильний тригер з прямим динамічним управлінням, другий лічильний тригер зі зворотним динамічним управлінням, схему І, другу схему АБО, причому вхід першого детектора підключений до виходу першого смугового фільтра, а його вихід з'єднаний з першим входом помножувача сигналів, вхід другого детектора підключений до виходу другого смугового фільтра, а його вихід з'єднаний з другим входом

помножувача сигналів, вхід другого детектора підключений до виходу другого смугового фільтра, а його вихід з'єднаний з другим входом помножувача сигналів, вихід помножувача сигналів приєднаний до входу диференціатора формувача сегментних границь складів, вихід якого підключений до перших входів першої і другої порогових схем, напруга на других входах яких визначає відповідні пороги пропускання, вихід першої порогової схеми з'єднаний з першим входом першої схеми АБО, вихід другої порогової схеми з'єднаний з другим входом першої схеми АБО, вихід першої схеми АБО підключений до входу першого лічильного тригера з прямим динамічним управлінням І, до входу другого лічильного тригера зі зворотним динамічним управлінням і до першого входу другої схеми АБО, прямі виходи лічильних тригерів з'єднані зі входами схеми І, вихід схеми І з'єднаний з другим входом другої схеми АБО і є другим входом пристрою, а вихід другої схеми АБО є першим виходом пристрою.

На Фіг.1 зображена структурна схема запропонованого пристрою для виділення складових сегментів у мовному сигналі, на Фіг.2 зображена функціональна схема формувача сегментних границь слів, на Фіг.3 зображені тимчасові діаграми напруги в пристрої для виділення складових сегментів.

Пристрій для виділення складових сегментів у мовному сигналі містить акустичний датчик 1, приєднаний до входу підсилювача 2, вихід якого підключений до входу першого смугового фільтра 3, причому вхід першого детектора 4 підключений до виходу першого смугового фільтра, вхід другого смугового фільтра 5 підключений до виходу підсилювача, причому вхід другого детектора 6 підключений до виходу другого смугового фільтра 5, а виходи першого й другого детекторів 4 і 6 відповідно з'єднані з першим і другим входами помножувача сигналів 7, вихід якого підключений до входу формувача сегментних границь складів 8.

Акустичний датчик 1 представляє собою мікрофон конденсаторного або електродинамічного типу з частотним діапазоном 60-15000Гц, наприклад, МК-5 або МД-63. Мікрофонний підсилювач 2 представляє собою схему операційного підсилювача в масштабному включенні.

Перший смуговий фільтр 3 настроєний на діапазон частот 250-540Гц, у якому зосереджена основна енергія дзвінких звуків, і являє собою схему фільтра Баттерворта. Перший амплітудний детектор 4 являє собою типову схему амплітудного детектування з більшим динамічним діапазоном. Другий смуговий фільтр 5 настроєний на діапазон частот 800-2500Гц, у якому зосереджена основна енергія голосних звуків і за схемою аналогічний смуговому фільтру 3, описаному вище. Другий амплітудний детектор 6 аналогічний першому амплітудному детектору 4, описаному вище. Помножувач сигналів 7 являє собою схему помножувача на інтегральній мікросхемі К525ПС2.

Формувач сегментних границь складів 8 складається з диференціатора 9, першої і другої порогових схем 10 і 11 відповідно, логічної схеми ПЛІ 12, лічильного тригера з прямим динамічним управлінням 13, лічильного тригера зі зворотним динамічним управлінням 14, логічної схеми І 15, логічної схеми АБО 16.

Диференціатор 9 являє собою практичну схему диференціювання на операційному підсилювачі. Перша й друга порогові схеми 10 й 11 являють собою компаратори. Логічна схема АБО 12 являє собою інтегральну схему диз'юнктора. Перший лічильний тригер 13 з прямим динамічним управлінням являє собою інтегральну схему D-тригера в лічильному включенні. Другий лічильний тригер 14 зі зворотним динамічним управлінням являє собою інтегральну схему JK-тригера в лічильному включенні. Логічна схема І 15 являє собою інтегральну схему кон'юнктора. Логічна схема АБО 16 аналогічна описаній вище схемі 12.

На Фіг.3 наведені тимчасові діаграми напруг у пристрої для виділення складових сегментів на прикладі слова «Миша». Фіг.3 а) зображує огинаючу мовного сигналу U_{d2} в діапазоні 800-2500Гц, отриману за допомогою другого амплітудного детектора 6 на виході другого смугового фільтра 5, а Фіг.3 б) - огинаючу мовного сигналу U_{d1} в діапазоні 250-540Гц на виході першого смугового фільтра 3, отриману за допомогою першого амплітудного детектора 5. Фіг.3 в) зображує сигнал U_n на виході помножувача сигналу 7. На Фіг.3 г) зображені сигнал U_{nd} на виході диференціатора 9 (суцільна лінія) і порогові напруги U_{n1} й U_{n2} , що подаються на другі входи першої і другої порогових схем 10 й 11 відповідно (штрихова лінія). На Фіг.3 д) зображена напруга U_n , одержувана на виході схеми АБО 12. Фіг.3 е) і Фіг.3 ж) зображують напруги логічних сигналів U_{T1} і U_{T2} на прямих виходах першого і другого тригерів 13 й 14 відповідно. На Фіг.3 з) зображений логічний сигнал U_r , (P_{rn}) на виході схеми збігу 15, що визначає положення складових ядер у мовному сигналі. На Фіг.3 й) зображений логічний сигнал P_{rn} , сегментуючий мовний сигнал на відрізки, що відповідають складам мови.

Пристрій для виділення складових сегментів у мовному сигналі працює наступним чином.

Мовний сигнал сприймається акустичним датчиком 1, перетворюється в електричний і надходить на вхід підсилювача 2. Електричний сигнал, посилений до величини, достатньої для роботи наступних каскадів, надходить на входи першого і другого смугових фільтрів 3 і 5. Амплітудний детектор 4, вхід якого підключений до виходу смугового фільтра 3, виділяє огинаючу U_{d2} (Фіг.3б) мовні сигнали в діапазоні 250-540Гц. Амплітудний детектор 6, вхід якого підключений до виходу смугового фільтра 4, виділяє огинаючу U_{d1} (Фіг.3а) мовні сигнали в діапазоні 800-2500Гц. Напруги U_{d1} й U_{d2} , що надходять на перший і другий входи помножувача сигналів 7, перемножуються, в результаті чого на його виході з'являється напруга U_n (Фіг.3в), яка дорівнює:

$$U_n = U_{d1} \cdot U_{d2}$$

Діапазон частот першого і другого смугових фільтрів 3 і 5 відповідно підібраний таким чином,

що в результаті виконання операції множення корелює тільки енергія голосних звуків, що приводить до усунення в результуючій огинаючій максимумів, що відповідають ділянкам високоенергетичних фрикативних звуків (напруга U_n на Фіг.3в). Ця напруга надходить на вхід диференціатора 9 формувача ознак 8 (Фіг.2), на виході якого утвориться напруга U_{nd} (Фіг.3г), пропорційна похідній від напруги U_n . Напруга U_{nd} надходить на перші входи першої і другої порогових схем 10 і 11. На другий вхід порогової схеми 10 надходить позитивна порогова напруга U_{n1} , а на другий вхід порогової схеми 11 надходить негативна порогова напруга U_{n2} , причому U_{n1} й U_{n2} обрані таким чином, що $|U_{n1}| = |U_{n2}| \approx 50\text{mV}$, при цьому усуваються помилкові спрацювання порогових схем через фонові шуми при переході U_{nd} через нульові значення. З виходів першої і другої порогових схем 10 і 11 одержують кліповані сигнали, приведені до стандартних рівнів цифрових сигналів. Ці сигнали надходять на входи першої схеми АБО 12, на виході якої утворюється цифровий сигнал U_n , представлений на Фіг.3 д). Фронти сигналу U_n , що надходить на вхід першого лічильного тригера з прямим динамічним управлінням 13, перемикають щораз його в протилежний стан, у результаті чого на його виході утворюється цифровий сигнал U_{T1} , представлений на Фіг.3 е). Другий лічильний тригер зі зворотним динамічним управлінням 14 перемикається спадами вхідного сигналу U_n і видає на виході сигнал U_{T2} , представлений на Фіг.3 ж). Сигнали U_{T1} і U_{T2} надходять на входи схеми збігу 15, на виході якої утворюються короткі імпульси U_r (Фіг.3 з), що локалізують центри складових ядер у слові. Ці імпульси надходять на перший вхід схеми АБО 16, а так само на вихід пристрою для виділення складових сегментів у слові, будучи ознакою P_{rn} визначення місця розташування центра голосних звуків у складі. На другий вхід другої схеми АБО 16 надходить цифровий сигнал U_n , що схемою АБО поєднується із сигналом U_r , у результаті чого на виході схеми 16 виробляється сигнал $P_{cл}$, що виділяє в слові сегменти, що відповідають положенню складів у слові.

Діапазон частот першого смугового фільтра 3, дорівнює 250-540Гц, вибраний з тих міркувань, що в ньому відсутня енергія високоенергетичних фрикативних звуків типу /ш/ і /ч/, що утворюють помилкові складові ядра, а також зосереджена значна частина енергії всіх дзвінких звуків, у тому числі й голосних. Однак у даному діапазоні енергія сонорних звуків типу /л/, /м/, /н/ сумірна з енергією голосних, тому визначення складових сегментів тільки по огинаючій мовного сигналу в даному діапазоні буде супроводжуватися багатьма помилками. Тому діапазон частот другого смугового фільтра 4 обраний у межах 800-2500Гц, у якому енергія голосних звуків мінімум у два рази перевищує енергію сонорних звуків.

У порівнянні з прототипом технічне рішення, що заявляється, має підвищену надійність і більш просту конструкцію.

Для визначення складових сегментів використовується тимчасова функція U_n , отримана

