



УКРАЇНА

(19) UA (11) 80587 (13) C2  
(51) МПК (2006)  
G10L 15/00  
G10L 17/00

МІНІСТЕРСТВО ОСВІТИ  
І НАУКИ УКРАЇНИ

ДЕРЖАВНИЙ ДЕПАРТАМЕНТ  
ІНТЕЛЕКТУАЛЬНОЇ  
ВЛАСНОСТІ

## ОПИС ДО ПАТЕНТУ НА ВІНАХІД

### (54) СПОСІБ СПЕКТРАЛЬНОЇ ІДЕНТИФІКАЦІЇ МОВНОГО СИГНАЛУ

1

(21) а200505554

(22) 10.06.2005

(24) 10.10.2007

(72) КУЗНЕЦОВ МАКСИМ ВОЛОДИМИРОВИЧ, UA,  
СЕЛЕТКОВ ВІКТОР ЛЕОНІДОВИЧ, UA(73) КУЗНЕЦОВ МАКСИМ ВОЛОДИМИРОВИЧ, UA,  
СЕЛЕТКОВ ВІКТОР ЛЕОНІДОВИЧ, UA

|      |          |          |            |
|------|----------|----------|------------|
| (56) | WO       | 9715915, | 01.05.1997 |
| SU   | 1691852  | A1,      | 15.11.1991 |
| US   | 4827516, |          | 02.05.1989 |
| JP   | 7036475, |          | 07.02.1995 |
| SU   | 518791,  |          | 18.07.1976 |
| US   | 4344031, |          | 10.08.1982 |
| US   | 5694942, |          | 09.12.1997 |
| US   | 4833714, |          | 23.05.1989 |

(57) Спосіб спектральної ідентифікації мовного сигналу в заданій послідовності взаємозв'язаних операцій перетворення мовного сигналу, за яким проводять низькочастотну фільтрацію мовного сигналу в смузі частот 3...10 кГц; розділяють мовний сигнал на сегменти аналізу однакової тривалості 10...30 мс; проводять стандартне 9...12-бітове аналого-цифрове перетворення сигналу з частотою дискретизації 6...20 кГц на кожному

2

сегменті аналізу, який відрізняється тим, що формують на кожному сегменті аналізу дискретні відліки поточної речовинної спектральної щільності Хартлі шляхом застосування стандартного дискретного перетворення Хартлі до речовинних відліків мовного сигналу на сегментах аналізу; виділяють абсолютні значення дискретних речовинних відліків поточних спектрів Хартлі мовного сигналу на сегментах аналізу; для кожної частоти дискретного спектра Хартлі сегмента аналізу мовного сигналу формують варіаційні ряди по сегментах виділених абсолютних значень дискретних речовинних відліків поточних спектрів Хартлі; для кожного варіаційного ряду абсолютних значень відліків виділяють квантілі; формують залежності виділених квантилів від дискретних значень частоти спектра мовного сигналу; порівнюють отримані залежності квантилів від дискретних значень частоти спектра мовного сигналу за відомими критеріями узгодженості, наприклад, Колмогорова, з відповідними залежностями квантилів від цих же значень частоти спектрів Хартлі еталонних або зразкових, або типових мовних реалізацій бази даних.

Передбачуваний винахід відноситься до області техніки автоматичного розпізнавання мови на основі акустичного і фонемного аналізу мовного управління виробничими процесами і автоматичними інформаційно-довідковими системами; мовного введення інформації в ЦВМ; ідентифікації джерел мовних сигналів в системах моніторингу, контролю і охорони.

Спектральна ідентифікація мовного сигналу робиться шляхом перевірки статистичних критеріїв узгодження параметрів поточного спектру мовного сигналу зі однотипними параметрами спектрів еталонних (зразкових, типових) мовних реалізацій і може знайти застосування в широко відомих системах спектральної ідентифікації джерел мовних сигналів.

Широко відомі способи спектральної ідентифікації мовних сигналів шляхом порівняння статистичних характеристик (гістограм функцій

розподілу або щільності розподілу імовірностей, оцінок середніх і дисперсій) значень спектральної щільності потужності або амплітудного спектру у всьому динамічному діапазоні амплітуд і частот спектру мовного сигналу з відповідними характеристиками (гістограмами функцій розподілу або щільності розподілу імовірностей, оцінками середніх і дисперсій) спектральної щільності потужності або амплітудних спектрів еталонних (зразкових, типових) мовних реалізацій бази даних [1, 2].

Найближчим за технічною суттю виконуваних операцій перетворення мовного сигналу і визначення статистичних характеристик значень модулів комплексних амплітуд спектральних компонент в діапазоні частот спектру мовного сигналу є спосіб-прототип, в якому в заданій послідовності виконують взаємозв'язані операції перетворення мовного сигналу: проводять

(13) C2

(11) 80587

(19) UA

низькочастотну фільтрацію мовного сигналу в смузі частот 3...10кГц; розділяють мовний сигнал на сегменти аналізу однакової тривалості 10...30мс; проводять стандартне 9...12-бітове аналого-цифрове перетворення мовного сигналу з частотою дискретизації 6...20кГц на кожному сегменті аналізу; формують на кожному сегменті аналізу дискретні відліки поточної комплексної спектральної щільності мовного сигналу шляхом застосування стандартного дискретного перетворення Фур'є до відліків мовного сигналу на сегментах; виділяють модульні значення дискретних відліків поточних комплексних спектрів сигналів сегментів аналізу; визначають статистичні характеристики (гістограми функцій розподілу або щільності розподілу імовірностей, оцінки середніх і дисперсій) модульних значень спектру по всіх сегментах аналізу для кожної частоти дискретного спектру мовного сигналу; порівнюють отримані статистичні характеристики (гістограми функцій розподілу або щільності розподілу імовірностей, оцінки середніх і дисперсій) мовного сигналу по відомим критеріям узгодженості, наприклад Колмогорова, з відповідними статистичними характеристиками (гістограмами функцій розподілу або щільності розподілів імовірностей, оцінками середніх і дисперсій) еталонних (зразкових, типових) мовних реалізацій бази даних [1, 2].

Основною причиною зниження ефективності відомих систем, що реалізують способи спектральної ідентифікації мовного сигналу, є те, що мовний сигнал є випадковим нестационарним процесом зі змінною дисперсією і складною формою поточної спектральної щільності потужності (амплітудного спектру). В результаті цього багатовимірні функції щільності розподілу імовірностей миттєвих модульних значень дискретних комплексних відліків поточних спектрів Фур'є сигналу на сегментах аналізу є негаусовськими та багатомодальними, що істотно ускладнює застосування відомих статистичних критеріїв узгодження і знижує імовірності тривальної ідентифікації.

В основу винаходу «Спосіб спектральної ідентифікації мовного сигналу» поставлена задача підвищення ефективності відомих систем спектральної ідентифікації мовного сигналу шляхом порівняння залежностей від спектральної частоти мовного сигналу квантилів (квartilів, децилів, процентилів) варіаційних рядів по всіх сегментах аналізу абсолютних значень поточних речовинних спектрів Хартлі мовного сигналу на сегментах з відповідними залежностями квантилів (квartilів, децилів, процентилів) абсолютних значень речовинних спектрів Хартлі еталонних (зразкових, типових) мовних реалізацій бази даних на основі відомих критеріїв узгодження.

Рішення поставленої задачі досягається тим, що в заданій способом-прототипом послідовності взаємозв'язаних операцій перетворення мовного сигналу: проводять низькочастотну фільтрацію мовного сигналу в смузі частот 3...10кГц; розділяють мовний сигнал на сегменти аналізу однакової тривалості 10...30мс; проводять

стандартне 9...12-бітове аналого-цифрове перетворення сигналу з частотою дискретизації 6...20кГц на кожному сегменті аналізу, додатково формують на кожному сегменті аналізу дискретні відліки поточної речовинної спектральної щільності Хартлі шляхом застосування стандартного дискретного перетворення Хартлі до речовинних відліків мовного сигналу на сегментах аналізу; виділяють абсолютні значення дискретних речовинних відліків поточних спектрів Хартлі мовного сигналу на сегментах аналізу; для кожної частоти дискретного спектру Хартлі мовного сигналу формують варіаційні ряди по сегментах виділених абсолютних значень дискретних речовинних відліків поточних спектрів Хартлі; для кожного варіаційного ряду абсолютних значень відліків виділяють квантілі (квartilі, децилі, процентілі); формують залежності виділених квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру мовного сигналу; порівнюють отримані залежності квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру мовного сигналу по відомим критеріям узгодженості, наприклад Колмогорова, з відповідними залежностями квантилів (квartilів, децилів, процентилів) від цих же значень частоти спектрів еталонних (зразкових, типових) мовних реалізацій бази даних.

На Фіг.1 наведена структурна схема пристрою спектральної ідентифікації мовного сигналу шляхом порівняння залежностей від дискретних значень частоти спектру мовного сигналу квантилів (квartilів, децилів, процентилів) варіаційних рядів абсолютних значень відліків поточних речовинних спектрів Хартлі мовного сигналу для кожної частоти дискретного спектру по сегментах аналізу з аналогічними залежностями квантилів (квartilів, децилів, процентилів) від частоти еталонних зразків мови бази даних, який реалізує об'єкт, що заявляється - спосіб спектральної ідентифікації мовного сигналу, де позначено:

1 - блок фільтру нижніх частот фільтрації мовного сигналу в смузі частот 3...10кГц;

2 - блок сегментації мовного сигналу на сегменти аналізу однакової тривалості 10...30мс;

3 - блок аналого-цифрового перетворення мовного сигналу з частотою дискретизації 6...20кГц і стандартним 9...12-бітовим квантуванням на кожному сегменті аналізу;

4 - блок дискретного перетворення Хартлі речовинних відліків мовного сигналу на сегментах аналізу;

5 - блок виділення абсолютних значень дискретних речовинних відліків поточних спектрів Хартлі сигналів на сегментах аналізу;

6 - блок формування варіаційних рядів виділених абсолютних значень по сегментах аналізу для кожної частоти дискретного спектру мовного сигналу;

7 - блок виділення квантилів (квartilів, децилів, процентилів) варіаційних рядів абсолютних значень дискретних відліків поточних спектрів Хартлі по сегментах;

8 - блок формування залежностей квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру мовного сигналу;

9 - блок порівняння залежностей квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру Хартлі мовного сигналу і відповідних залежностей квантилів (квartilів, децилів, процентилів) спектрів Хартлі еталонних (зразкових, типових) мовних реалізацій бази даних;

10 - блок еталонних залежностей квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру еталонних (зразкових, типових) мовних реалізацій бази даних;

11 - блок ідентифікації мовного сигналу за наслідками статистичного порівняння залежностей квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру мовного сигналу і еталонних (зразкових, типових) мовних реалізацій бази даних на основі відомого критерію узгодженості Колмогорова.

Вихід блоку фільтру нижніх частот 1 фільтрації мовного сигналу в смузі частот 3...10кГц, вхід якого є входом пристрою, через послідовно з'єднанні блок сегментації 2 мовних сигналів на сегменти аналізу однакової тривалості 10...30мс, блок аналого-цифрового перетворення 3 мовних сигналів з частотою дискретизації 6...20кГц і стандартним 9...12-бітовим квантуванням на кожному сегменті аналізу, блок дискретного перетворення Хартлі 4 речовинних відліків мовного сигналу на сегментах аналізу, блок виділення абсолютних значень 5 дискретних речовинних відліків поточних спектрів Хартлі мовного сигналу на сегментах аналізу, блок формування варіаційні рядів 6 виділених абсолютних значень по сегментах аналізу для кожної частоти дискретного спектру мовного сигналу, блок виділення квантилів 7 (квartilів, децилів, процентилів) варіаційних рядів абсолютних значень дискретних відліків поточних спектрів Хартлі по сегментах аналізу, блок формування залежностей квантилів 8 (квartilів, децилів, процентилів) від дискретних значень частоти спектру мовного сигналу підключений до першого входу блоку порівняння залежності квантилів 9 (квartilів, децилів, процентилів) від дискретних значень частоти спектру Хартлі мовного сигналу і відповідних залежностей квантилів (квartilів, децилів, процентилів) спектрів Хартлі еталонних (зразкових, типових) мовних реалізацій бази даних, до другого входу якого підключений вихід блок еталонних залежностей квантилів 10 (квartilів, децилів, процентилів) від дискретних значень частот спектру Хартлі еталонних (зразкових, типових) мовних реалізацій бази даних, а вихід підключений до входу блоку ідентифікації мовного сигналу 11 за наслідками статистичного порівняння різниць відповідних залежностей на основі критерію узгодженості Колмогорова, вихід якого є виходом пристрою.

З виходу блоку фільтру нижніх частот 1 фільтрації мовного сигналу в смузі частот 3...10кГц мовний сигнал поступає на вхід блоку сегментації

2 мовного сигналу на сегменти аналізу однакової тривалості 10...30мс, а з його виходу подається на блок аналого-цифрового перетворення 3 з частотою дискретизації 6...20кГц і стандартним 9...12-бітовим квантуванням мовного сигналу на кожному сегменті аналізу. Мовний сигнал в цифровому вигляді з виходу блоку 3 подається на вхід блоку 4 стандартного дискретного перетворення Хартлі речовинних відліків мовного сигналу на сегментах аналізу, при цьому речовинні дискретні значення спектральної щільності Хартлі мовного сигналу кожного сегменту з виходу блоку 4 поступають на вхід блоку 5 виділення абсолютних значень дискретних речовинних відліків поточних спектрів Хартлі сигналів на сегментах аналізу шляхом відкидання знакового біта цифрового дискретного відліку спектру. З виходу блоку виділення абсолютних значень 5 дискретні абсолютні значення спектрів подаються на вхід блоку 6 формування варіаційних рядів виділених абсолютних значень по сегментах аналізу для кожної частоти дискретного спектру мовного сигналу шляхом розміщення виділених абсолютних значень за збільшенням таким чином, що на першому сегменті розміщуються мінімальні величини абсолютних значень, а на останньому відповідно максимальні величини. Впорядковані таким чином за збільшенням величини виділених абсолютних значень з виходу блоку 6 подаються на вхід блоку 7 виділення квантилів (квartilів, децилів, процентилів) сформованих варіаційних рядів абсолютних значень дискретних відліків поточних спектрів Хартлі по сегментах як порядкових статистик вибіркового розподілу імовірностей абсолютних значень дискретних відліків поточних спектрів Хартлі для кожне і частоти дискретного спектру мовного сигналу. Виділені квантилі (квartilів, децилів, процентилів) з виходу блоку 7 поступають на вхід блоку 8 формування залежності квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру мовного сигналу, які подаються на перший вхід блоку 9 порівняння залежності квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру мовного сигналу і відповідних залежностей квантилів (квartilів, децилів, процентилів) еталонних (зразкових, типових) мовних реалізацій бази даних, що поступають на його другий вхід з виходу блоку 10 еталонних залежностей квантилів (квartilів, децилів, процентилів) від дискретних значень частоти спектру еталонних (зразкових, типових) мовних реалізацій бази даних. Вихідний сигнал блоку 9 у вигляді різниць залежностей квантилів (квartilів, децилів, процентилів) спектру Хартлі мовного сигналу і відповідних залежностей квантилів (квartilів, децилів, процентилів) спектрів Хартлі еталонних (зразкових, типових) мовних реалізацій бази даних подається на вхід блоку 11 ідентифікації мовного сигналу за наслідками статистичного аналізу різниць залежностей квантилів (квartilів, децилів, процентилів) спектрів Хартлі від дискретних значень частоти поточного спектру мовного сигналу і еталонних (зразкових, типових) спектрів

Хартлі мовних реалізацій бази даних на основі відомого критерію узгодженості Колмогорова. Вихід блоку 11 є виходом пристрою спектральної ідентифікації мовного сигналу в цілому.

На Фіг.2 наведена залежність квантилів спектру Хартлі еталонного мовного сигналу диктора №1,

а на Фіг.3 наведена залежність квантилів спектру Хартлі еталонного мовного сигналу диктора №2 бази даних.

На Фіг.4 наведена відповідна експериментальна залежність квантилів спектру Хартлі поточного спектру мовного сигналу диктора, що ідентифікується. Зіставлення різниць наведених залежностей квантилів спектрів Хартлі від дискретних значень частоти поточного спектру мовного сигналу і еталонних (зразкових, типових) спектрів Хартлі мовних реалізацій бази вказує на належність мовного сигналу, що ідентифікується, диктору №2. Використання більш інформативних залежностей й децилів та процентилів дозволяє підвищити вірогідність правильної ідентифікації.

У зв'язку з тим, що квантилі (квартилі, децилі, процентілі) варіаційних рядів абсолютних значень дискретних відліків поточних спектрів Хартлі по сегментах як порядкові статистики вибіркового розподілу імовірностей абсолютних значень дискретних відліків поточних спектрів Хартлі для кожної частоти дискретного спектру мовного сигналу розподіляються по нормальному закону незалежно від негаусовського характеру багатовимірних функцій розподілу вибірових значень спектру мовного сигналу, істотно підвищується ефективність критерію узгодженості Колмогорова, орієнтованого на статистику ідентифікацію нормально розподілених випадкових величин, що використовується, що, зрештою, підвищує вірогідність правильної ідентифікації мовного сигналу.

Разом з тим мета способу-прототипу - спектральна ідентифікація мовного сигналу - в об'єкті, що заявляється, реалізується більш ефективно внаслідок того, що обчислювальні (часові) витрати на отримання речовинної спектральної щільності Хартлі мовного сигналу на сегменті менше ніж витрати на отримання комплексної спектральної щільності Фур'є. Крім того, отримання абсолютних значень дискретних відліків поточних спектрів Хартлі реалізується шляхом простого відкидання знакового біта цифрового дискретного відліку спектру, тоді як отримання модулів комплексних амплітуд спектральних компонент Фур'є вимагає реалізації нелінійної операції вилучення квадратного кореня з суми квадратів реальної і уявної частин відліку спектральної густини Фур'є [3].

У свою чергу, спектральна щільність Хартлі мовного сигналу, на відміну від спектральної щільності потужності або амплітудного спектру Фур'є, несиметрична щодо нульової частоти поточного спектру мовного сигналу, оскільки у часовій області реалізації мовних сигналів не є парними функціями. Ця обставина дозволяє використовувати для підвищення якості ідентифікації мовного сигналу додаткову

інформаційну ознаку - ступінь несиметричності реалізації мовного сигналу на сегменті аналізу [3].

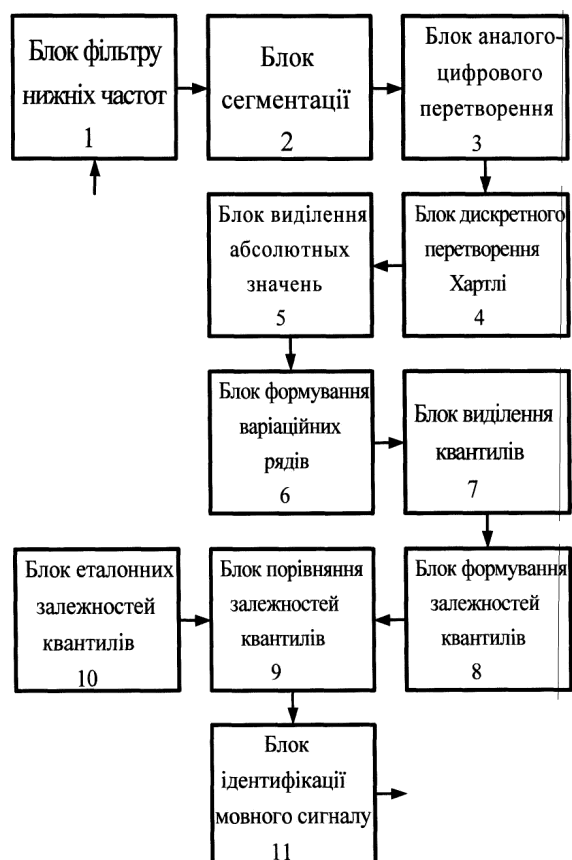
Комплексне використання перерахованих відмітних особливостей і пов'язаних з ними позитивних ефектів дозволяють забезпечити підвищення ефективності спектральної ідентифікації мовного сигналу шляхом порівняння залежності від спектральної частоти мовного сигналу квантилів (квартилів, децилів, процентилів) варіаційних рядів по всіх сегментах аналізу абсолютних значень поточних речовинних спектрів Хартлі мовного сигналу на сегментах з відповідними залежностями квантилів (квартилів, децилів, процентилів) абсолютних значень речовинних спектрів Хартлі еталонних (зразкових, типових) мовних реалізацій бази даних на основі відомих критеріїв узгодженості.

джерела інформації:

1. Сапожков М.А. Речевой сигнал в кибернетике и связи (Преобразование речи применительно к задачам техники связи и кибернетики). - М: Гос. изд-во л-ры по вопр. связи и радио, 1963.-С.368-377.

2. Шелухин О.И. Лукьянцев Н.Ф. Цифровая обработка и передача речи. / Под ред. О.И.Шелухина.- М.: Радио и связь, 2000.- 456 :ил. - С.98-106.

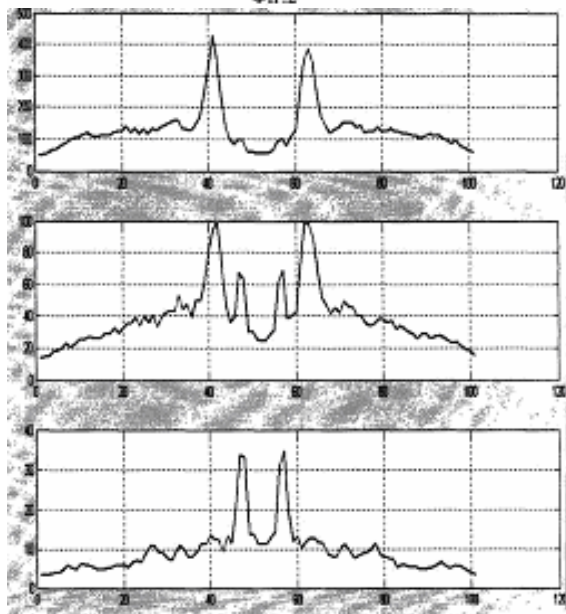
3. Денисенко А.Н. Сигналы. Теоретическая радиотехника. Справочное пособие. - М.: Горячая линия - Телеком, 2005,- 704с: ил. - С.99-105.



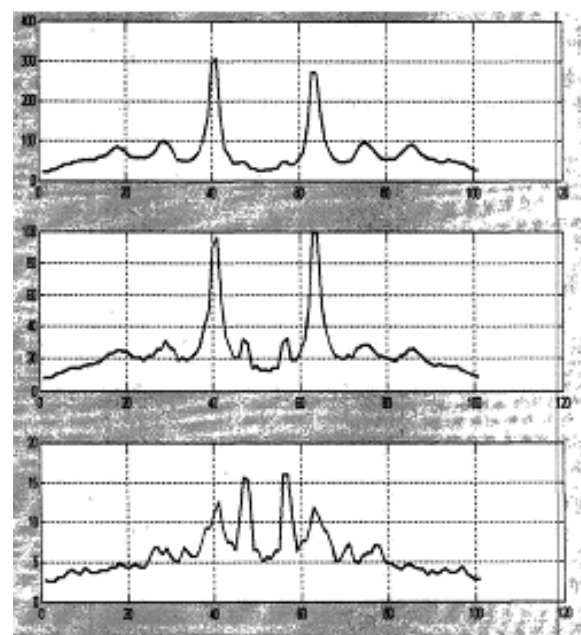
Фіг.1



Φir.2



Φir.3



Φir.4.