



УКРАЇНА

(19) **UA** (11) **140857** (13) **U**
(51) МПК (2020.01)
G10L 15/00
G10L 19/00
G09B 19/04 (2006.01)

МІНІСТЕРСТВО РОЗВИТКУ
ЕКОНОМІКИ, ТОРГІВЛІ ТА
СІЛЬСЬКОГО ГОСПОДАРСТВА
УКРАЇНИ

(12) ОПИС ДО ПАТЕНТУ НА КОРИСНУ МОДЕЛЬ

(21) Номер заявки: u 2019 09225	(72) Винахідник(и): Яшник Оксана Ігорівна (UA), Луньова Світлана Андріївна (UA), Руденко Іван Леонідович (UA)
(22) Дата подання заявки: 09.08.2019	
(24) Дата, з якої є чинними права на корисну модель: 10.03.2020	(73) Власник(и): НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ "КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО", просп. Перемоги, 37, м. Київ-56, 03056 (UA)
(46) Публікація відомостей про видачу патенту: 10.03.2020, Бюл.№ 5	

(54) СПОСІБ ЗНАХОДЖЕННЯ УСЕРЕДНЕНОГО СПЕКТРА МОВИ

(57) Реферат:

1. Спосіб знаходження усередненого спектра мови, що включає звукозапис мовного матеріалу дикторами у звукоізоляційній акустичній камері звукозаписувальним обладнанням з наступним мікшуванням звукозапису; встановлення частотного діапазону для подальшого аналізу частотної залежності спектрального амплітудного складу записаного мовного матеріалу, який відрізняється тим, що для знаходження усередненого спектра мови використовують прецизійну техніку: вимірювальний ненаправлений конденсаторний мікрофон, звукову карту, персональний комп'ютер (ПК) та здійснюють обробку записаного звукового сигналу на ПК в середовищах MatLab і ProTools.
2. Спосіб за п. 1, який відрізняється тим, що для звукозапису мовного матеріалу використовують мовний матеріал щонайменше трьох функціональних стилів мовлення, таких як художній, науковий, офісно-діловий тощо.
3. Спосіб за п. 1, який відрізняється тим, що для звукозапису мовного матеріалу використовують текст, що містить щонайменше 2500 символів без урахування пробілів.
4. Спосіб за п. 1, який відрізняється тим, що частотний діапазон для подальшого аналізу амплітудного складу мовного матеріалу встановлюють в межах від 100 Гц до 14-15 кГц.
5. Спосіб за пп. 1, 2, який відрізняється тим, що додаткову обробку мовного сигналу здійснюють у вигляді компресії та еквалізації.
6. Спосіб за пп. 2, 3, який відрізняється тим, що ідентифікацію мови визначають характерними областями спектра, які знаходять обробкою звукового сигналу масиву аудіокниг.

UA 140857 U

UA 140857 U

Спосіб знаходження спектра мови належить до галузі акустики, а саме до психоакустики, та може бути використаний у медичній галузі, зокрема, у оториноларингології, для налаштування слухових апаратів та кохлеарних імплантів або у таких сферах діяльності, як лінгвістичні дослідження текстів, розробки дикторських приміщень для телебачення та радіо, тестування електроакустичних трактів для передачі мови тощо.

Аналіз характеристик української мови здійснювався шляхом використанням результатів, отриманих в області психолінгвістики для інших мов. Такий підхід не враховує відмінні риси української мови.

Найбільш інформативною інтегральною характеристикою мови є її спектр. Усереднений спектр мови $B_p(f)$ визначають як середній спектральний рівень мовлення в залежності від частоти впродовж тривалого часу. Така характеристика визначає розподіл потужності мовного сигналу в зазначеному частотному діапазоні.

Спектральний рівень мовлення визначають через інтенсивність мовного сигналу у встановленому діапазоні частот за формулою:

$$B_p = 10 \lg \frac{I_1}{I_0},$$

де $I_0 = 10^{-12} \left[\frac{B_T}{\text{см}^2} \right]$ інтенсивність нульового рівня, відносно якого обчислюють рівні інших звуків;

$$I_1 = \frac{I_{\text{ср}}}{\Delta f} \left[\frac{B_T}{\text{см}^2} \right] \text{ інтенсивність звуку в смузі частот шириною } \Delta f.$$

Недоліком є те, що відомий спосіб не визначає особливостей української мови [1].

Відомий спосіб вимірювання миттєвого спектра мови, який використовують для розпізнавання мови і, зокрема, виявлення пауз у мовному сигналі. Для цього мовний сигнал, який приймають мікрофоном, розкладають на гребінці смугових фільтрів на ряд частотних смуг із подальшим виділенням обвідної сигналу.

Недоліком відомого способу є необхідність побудови складного електронного блоку, в якому використовують в якості еталону спектр російської мови. Крім того, оцінюється миттєвий, а не усереднений спектр мови [2].

Найбільш близьким є спосіб визначення усередненого спектра мови, в якому використовують автоматичну установку, яка містить вимірювальний конденсаторний мікрофон, сполучений із підсилювачем, фільтри, пристрій для квадровання, інтегратор та лічильник. Діапазон частот встановлюють від 0 Гц до 8000 Гц. Індивідуальний запис кожного диктора устновлений у звукоізоляційній акустичній камері за допомогою вимірювального мікрофона, встановленого на відстані 8 см від диктора. Спосіб забезпечує отриманням схожих результатів при обробці мовних сигналів, записаних різними групами дикторів, але в якості мовного матеріалу використовували обмежені функціональні стилі мовлення: газетні тексти та художні твори російською мовою.

Недоліком цього відомого способу є використання спеціальної установки для аналізу мови, яка є достатньо складною у виготовленні та вносить певні похибки у статистичні дослідження, фільтрацію отриманого звукозапису мовного матеріалу. Дотримання малої відстані від мікрофону створює залежність від параметрів звукозаписувального обладнання.

Обмежений частотний діапазон та обрані стилі мовного матеріалу, а також тривалість звукозапису зменшують точність знаходження усередненого спектра мови.

В основу запропонованого способу покладено задачу удосконалення способу знаходження усередненого спектра мови шляхом обробки мовного сигналу, що зменшить похибки при знаходженні усередненого спектра мови.

Поставлена задача досягається тим, що у способі знаходження усередненого спектра мови, який включає звукозапис мовного матеріалу дикторами у звукоізоляційній акустичній камері, встановлення частотного діапазону для подальшого аналізу звукового сигналу та отримання розподілу спектральних енергетичних рівнів мовленнєвого сигналу, використовують комп'ютерну обробку звукового сигналу, а також враховують вплив компресії та еквалізації мовленнєвого сигналу звукозаписувальним обладнанням, що підвищить точність результатів.

Новим є те, що для знаходження усередненого спектра мови використовують прецизійну техніку: вимірювальний ненаправлений конденсаторний мікрофон, звукову карту, персональний комп'ютер (ПК) та здійснюють обробку записаного звукового сигналу на ПК в середовищах MatLab і ProTools.

Новим є те, що для звукозапису мовного матеріалу використовують мовний матеріал щонайменше трьох функціональних стилів мовлення, таких як художній, науковий, офісно-діловий тощо.

5 Новим є те, що для звукозапису мовного матеріалу використовують текст, що містить щонайменше 2500 символів без урахування пробілів.

Новим є те, що частотний діапазон для подальшого аналізу амплітудного складу записаного мовного матеріалу встановлюють в межах від 100 Гц до 14-15 кГц.

Новим є те, що здійснюють додаткову обробку записаного мовного сигналу у вигляді компресії та еквалізації, при чому зберігають характерні області спектру.

10 Новим є те, що приналежність мовного тексту до певної мови визначають за характерними областями спектра мови та порівнюють зі спектром мови масиву аудіокниг, а саме мовного матеріалу з різним діалектом відповідно до території використання та часу створення текстів.

Сутність заявленого технічного рішення пояснюється схемою та графіками, де на:

фiг.1 - обробка мовного сигналу:

- 15 1 - звукоізоляційна акустична камера,
- 2 - вимірювальний ненаправлений конденсаторний мікрофон,
- 3 - звукова карта,
- 4 - комп'ютерний обчислювальний пристрій,
- 5 - пристрій збереження даних у цифровій формі,
- 20 6 - програмно-апаратний комплекс Protools,
- 7 - програмно-апаратний комплекс Matlab,
- 8 - збереження даних у 8-блок цифровій формі;

фiг. 2 - спектр мови за записаним сигналом і сигналом після первинної обробки;

фiг. 3 - спектр мови в залежності від компресії звукового сигналу;

25 фiг. 4 - спектр мови в залежності від частоти еквалізації;

фiг. 5 - спектр мови за записаним сигналом і звуковим сигналом масиву аудіокниг.

Спосіб виконують наступним чином:

диктори - молоді люди віком 20-22 роки (4 дівчини та 3 чоловіка), які читають мовний матеріал, наприклад: художній твір Олександра Довженка "Зачарована Десна", нормативно-правові акти Конституції України, текстова інформація конспекту лекцій з конструювання електроакустичних перетворювачів НТУУ "КПІ" ім. І.Сікорського.

Звукозапис мовного матеріалу здійснюють (фiг. 1) у звукоізоляційній акустичній камері 1 за допомогою звукозаписувального обладнання. Акустичні характеристики камери: час реверберації 0,15-2 с в діапазоні частот 63-8000 Гц. Відстань від диктора до мікрофона становить ~30 см.

Як звукозаписувальне обладнання використовують прецизійну техніку звукозапису: вимірювальний ненаправлений конденсаторний мікрофон 2, сполучений зі звуковою картою 3, під'єднанню до комп'ютерного обчислювального пристрою 4, обладнаного пристроєм збереження даних у цифровій формі 5.

40 Частотний діапазон для подальшого аналізу записаного мовного матеріалу встановлюють в межах: 100 Гц - (14-15) кГц.

Записані звукові доріжки оброблюють програмно-апаратними комплексами 6 ProTools® та 7 Matlab®, причому попередньо видаляють зайві елементи запису (вдохи - видихи, паузи між фразами, тривалість яких більше 0,5 с, невдалі дублі тощо).

45 Елементи обробленого мовного запису зводять у єдиний блок 8 акустичного сигналу у цифровому виді).

Аналіз впливу первинної обробки сигналу (фiг. 2), свідчить, що видалення неінформативної частини звукових доріжок не має істотного впливу на результат знаходження усередненого спектра мови (максимальна розбіжність записаного і обробленого сигналу не перевищує 2 дБ).

50 Аналіз впливу компресії на спектральну характеристику сигналу (фiг. 3) здійснюють програмно-апаратним комплексом ProTools при співвідношенні 2:1, 4:1 та 8:1. Час атаки складає ~2 мс, час відновлення сигналу ~50 мс.

При аналізі впливу еквалізації застосовують фільтр низьких частот (ФНЧ) 2-го порядку з крутизою зрізу 12 дБ/окт. Результати впливу еквалізації зображено (фiг. 4) для частоти зрізу: 55 500; 1000; 2000 і 4000 Гц.

Для уточнення характерних областей спектра української мови додатково оброблений масив аудіофайлів (2358 звукових доріжок відповідає приблизно 90 хвилинам звучання). Аудіокниги відбиралися з максимальним представленням різних регіонів України (відмінні риси діалекту) та різного часу написання твору (часові зміни мови більш ніж за 50 років).

Порівняння графіків спектральної щільності потужності, одержаного в результаті обробки аудіофайлів та сигналу, записаного за допомогою дикторів (фіг. 5) свідчить, що характерними ознаками спектру для української мови є: максимум у діапазоні 100-500 Гц (з піком в районі 250 Гц), підйом на частотах 1000-1250 Гц, а також посилення енергетичного рівня мови на високих частотах (5000-10000 Гц).

Результати підтверджують, що українська мова належить до балто-слов'янської групи мов. Для української мови характерний більший енергетичний рівень низькочастотних спектральних складових, що свідчить про грудний характер звучання. Відмінністю від інших мов є відносно зростання енергії на середніх частотах (в районі 1000-1600 Гц), а також флуктуації рівня спектральної густини потужності на високих частотах (4000-9000 Гц), які пояснюються особливостями вживання шиплячих і свистячих звуків.

Заявлене технічне рішення забезпечує високу точність знаходження спектральних характеристик мови та усуває проблему відсутності об'єктивних акустичних характеристик української мови.

Джерела інформації:

1. С.А. Луньова, В.С. Дідковський, О.І. Педченко. Акустика мовотворення/ Навчальний посібник. - LAP: Lambert Academic Publishing, 2018.

3. Е.Г. Жиликов, С.П. Белов, А.С. Белов, В.В. Муромцев. Способ обнаружения пауз в речевых сигналах и устройство для его реализации. - Патент России RUN92399103 МПК G10L 11/00, от 10.09.2010.

4. Н.Б. Покровский. Расчет и измерение разборчивости речи. - М.: Связь издат, 1962.

5. О. Pedchenko and S. Lunova, "ANALYSIS OF UKRAINIAN DIAGNOSTIC ARTICULATION TABLES", EUREKA Phys. Eng., vol. 1, pp. 63-72, Jan. 2018.

ФОРМУЛА КОРИСНОЇ МОДЕЛІ

1. Спосіб знаходження усередненого спектра мови, що включає звукозапис мовного матеріалу дикторами у звукоізоляційній акустичній камері звукозаписувальним обладнанням з наступним мікшуванням звукозапису; встановлення частотного діапазону для подальшого аналізу частотної залежності спектрального амплітудного складу записаного мовного матеріалу, який **відрізняється** тим, що для знаходження усередненого спектра мови використовують прецизійну техніку: вимірювальний ненаправлений конденсаторний мікрофон, звукову карту, персональний комп'ютер (ПК) та здійснюють обробку записаного звукового сигналу на ПК в середовищах MatLab і ProTools.

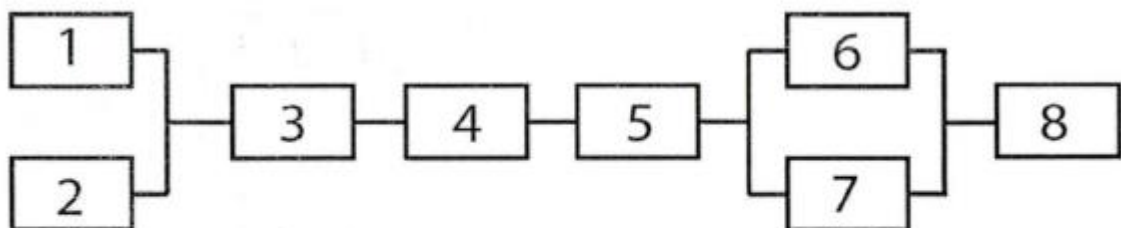
2. Спосіб за п. 1, який **відрізняється** тим, що для звукозапису мовного матеріалу використовують мовний матеріал щонайменше трьох функціональних стилів мовлення, таких як художній, науковий, офісно-діловий тощо.

3. Спосіб за п. 1, який **відрізняється** тим, що для звукозапису мовного матеріалу використовують текст, що містить щонайменше 2500 символів без урахування пробілів.

4. Спосіб за п. 1, який **відрізняється** тим, що частотний діапазон для подальшого аналізу амплітудного складу мовного матеріалу встановлюють в межах від 100 Гц до 14-15 кГц.

5. Спосіб за пп. 1, 2, який **відрізняється** тим, що додаткову обробку мовного сигналу здійснюють у вигляді компресії та еквалізації.

6. Спосіб за пп. 2, 3, який **відрізняється** тим, що ідентифікацію мови визначають характерними областями спектра, які знаходять обробкою звукового сигналу масиву аудіокниг.



Фіг. 1

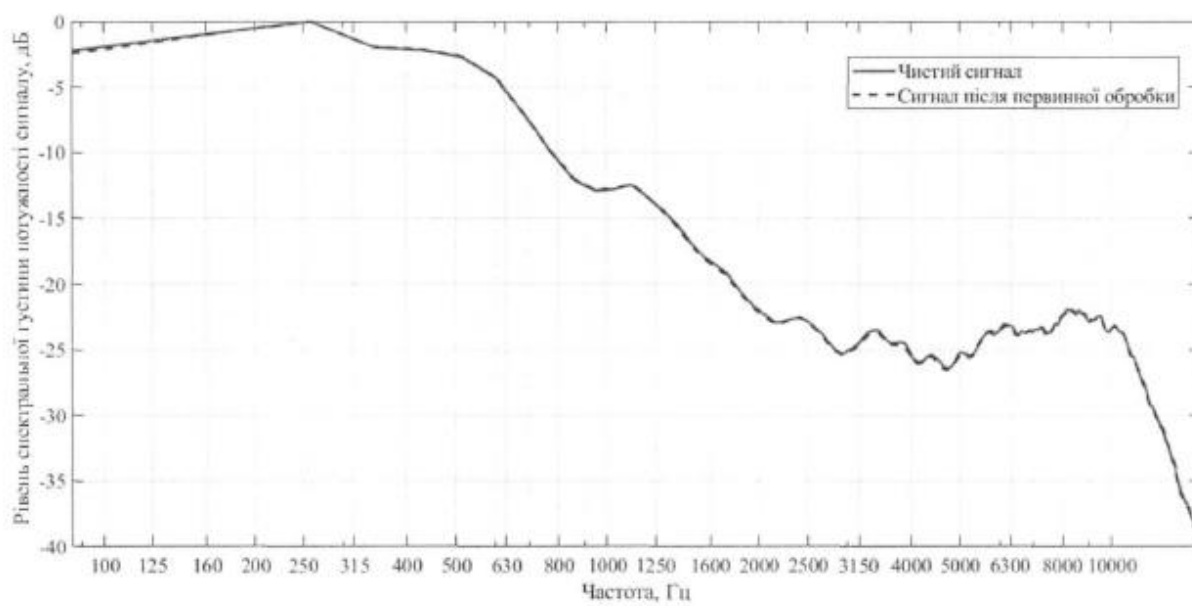


Fig. 2

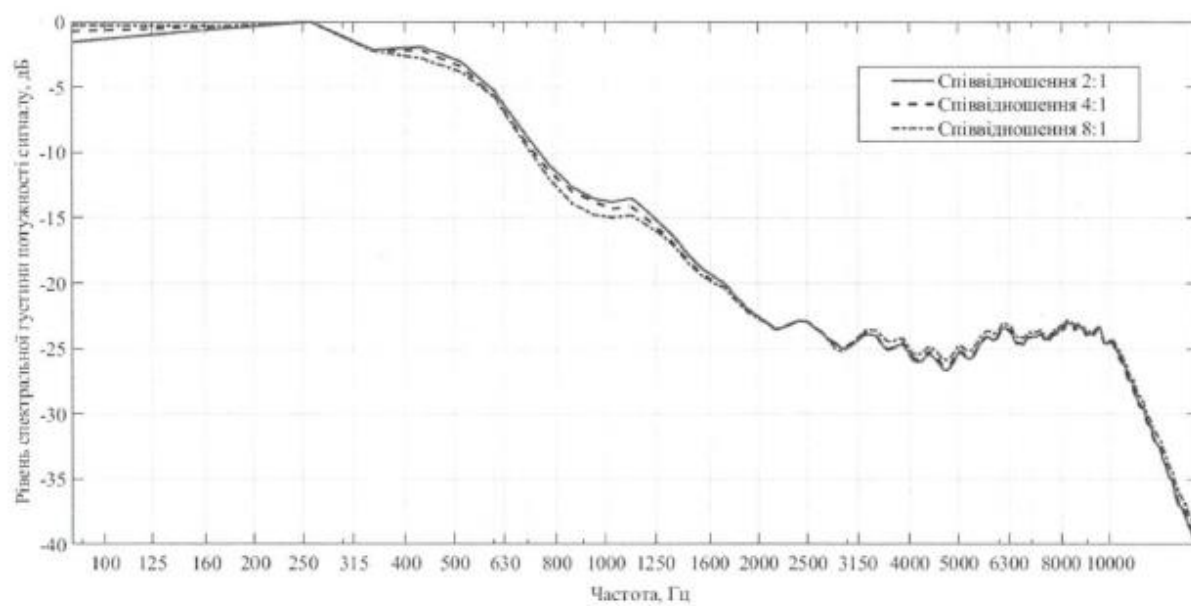
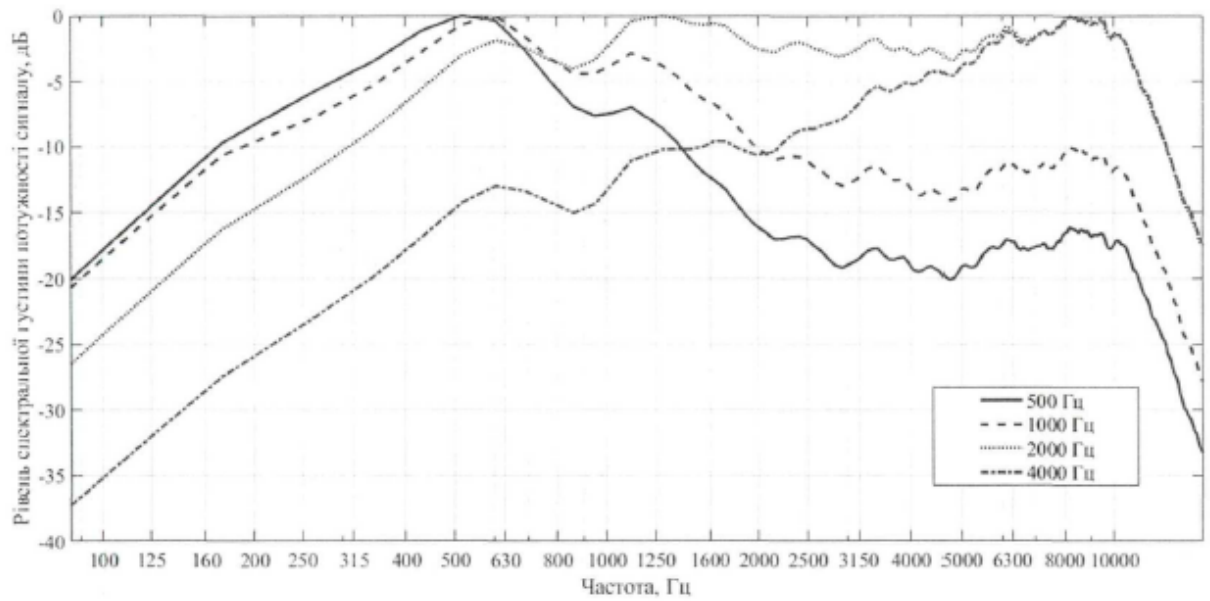
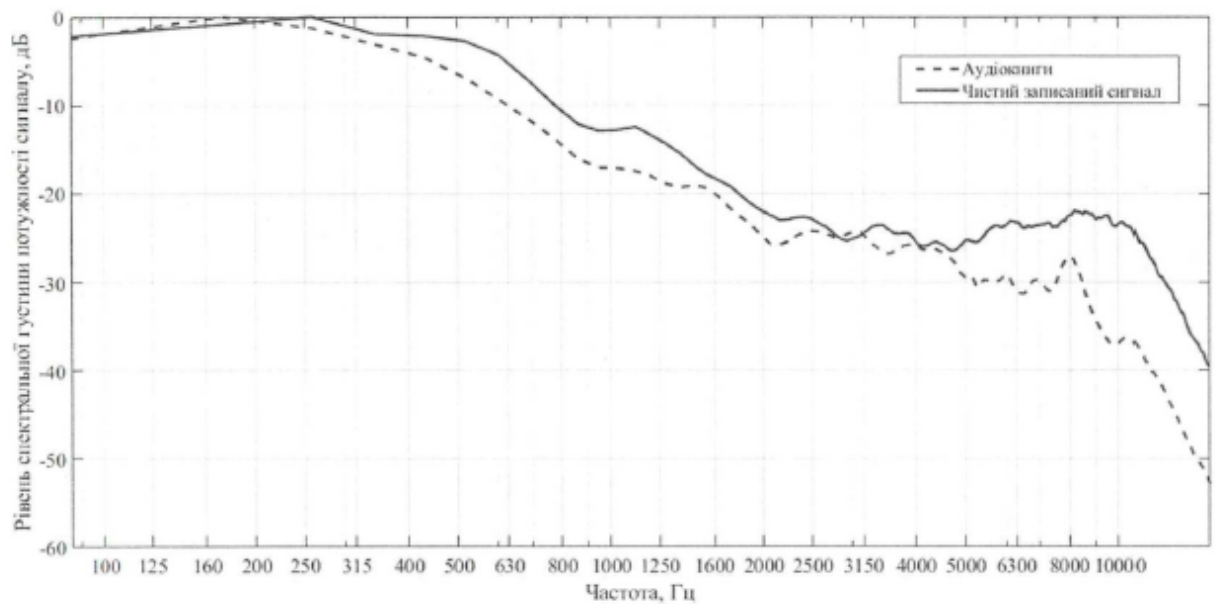


Fig. 3



Фіг. 4



Фіг. 5

Комп'ютерна верстка А. Крулевський

Міністерство розвитку економіки, торгівлі та сільського господарства України,
вул. М. Грушевського, 12/2, м. Київ, 01008, Україна

ДП "Український інститут інтелектуальної власності", вул. Глазунова, 1, м. Київ – 42, 01601