

Винахід відноситься до медицини і може бути використаний для обробки мовних сигналів у слухових апаратах з метою поліпшення сприйняття мови пацієнтами, які страждають на нейросенсорну тугоухість.

Відомий спосіб обробки мови, (книга авторів Ліклайдер Д.К.Р., Міллер Д.А. Сприйняття мови. Експериментальна психологія. / Під ред. С.С. Стівенса. М: Іностр література 1963, стор. 93-96), який полягає у видаленні значної частини мовного спектру в області високих або низьких частот. При оцінці сприйняття мовного сигналу в результаті такої фільтрації, у пацієнтів з нормальним слухом зберігається 100% розбірливості мови. Мовний сигнал має властивість надмірності, таким чином, часткове видалення спектрального складу мовного сигналу не впливає на розбірливість мови.

Недолік даного способу полягає в тому, що у ньому не враховуються індивідуальні особливості сприйняття мови. Наприклад, для людей з порушенням слуху даний спосіб фільтрації не є ефективним, через відсутність механізмів поліпшення чіткості сприйняття мови.

Найбільш близьким до пропонованого винаходу є спосіб обробки мови (Смугова фільтрація мовного сигналу - сприйняття мови у нормі та при нейросенсорній тугоухості, С.І. Петров, Вісник оториноларингології", N 3-2000, стор. 55-56), суть якого полягає у фільтрації мовного сигналу шляхом ділення спектру на частотні смуги постійної ширини. У вказаному способі видаляються смуги частот шириною 100Гц з парними номерами, таким чином, отриманий сигнал є сумою парних смуг спектру. Достовірність результатів смугової фільтрації, підтверджується за допомогою програмного пакету "Cool", американської компанії Syntrellium Software Corporation. Дана програма призначена для багатобічного, у тому числі і спектрального аналізу мовного сигналу. Для оцінки впливу даного способу обробки на сприйняття мовного сигналу проводилося порівняння розбірливості обробленого і первинного, такого, що пройшов тільки через аналого-цифровий і цифро-аналоговий перетворювачі, мовного сигналу. Розбірливість слів оцінювалася за допомогою таблиць Грінберга-Зіндера, які враховують фонетичні особливості російської мови, при комфортній для кожного випробовуваного гучності мови. Ніяких додаткових перетворень вихідного сигналу, таких як автоматичне регулювання підсилення або компресія, не проводилося. Оскільки амплітудно-частотні характеристики використовуваних пацієнтами слухових апаратів і підсилювача розрізняються, до початку реєстрації результатів обстеження проводилося тренування. При вимірюванні розбірливості мови, що складається з непарних смуг шириною 100Гц, у хворих на нейросенсорну тугоухість виявлено поліпшення розбірливості мови у межах від 4 до 16%.

Недоліком описаного способу є те, що автоматичне формування фільтру частотних смуг не враховує особливості конкретного мовного сигналу, оскільки частотні смуги, що видаляються, мають детерміновану ширину. Отже, якщо важливі для сприйняття мови складові спектру потрапляють в зони видалення, вони не сприймаються хворим.

Технічне завдання пропонованого способу полягає у підвищенні розбірливості мови шляхом виділення найбільш інформативних областей спектру мовного сигналу, які сприяють поліпшенню сприйняття мови у хворих на нейросенсорну тугоухість. Ця задача вирішена таким чином. У способі підвищення розбірливості мови у хворих на нейросенсорну тугоухість, що здійснює смугову фільтрацію мовного сигналу і ділення спектру на частотні смуги, згідно винаходу, в спектрі мовного сигналу виділяють форманти, потім проводять фільтрацію двох перших формант звуків, що входять в мовну послідовність, з подальшим видаленням частотних смуг, що знаходяться на межі з областю формантного піку, а також накладенням на спектр мовного сигналу фільтру, що відсікає частоти вище 1кГц, причому смугову фільтрацію здійснюють згідно формулам:

$$F_a = F_{\max} - \frac{F_{\max} - F_1}{2} \quad (1.1)$$

$F_a, F_1$  – межі формантних областей, які належить видалити;  
 $F_{\max}$  - пікова частота форманти.

$$F_b = F_{\max} + \frac{F_2 - F_{\max}}{2} \quad (1.2)$$

$F_b, F_2$  - межі формантних областей, які належить видалити;  
 $F_{\max}$  - пікова частота форманти.

На фігурі 1 - зображена схема пристрою, за допомогою якого здійснюється пропонований спосіб.

На фігурі 2 - зображена формантна лінія до фільтрації.

На фігурі 3 - зображена формантна лінія після фільтрації.

На фігурі 4 - зображено спектр звуку «Є» до фільтрації.

На фігурі 5 - зображено спектр звуку «Є» після фільтрації.

На фігурі 6 - зображено спектр звуку «Ж» до фільтрації.

На фігурі 7 - зображено спектр звуку «Ж» після фільтрації.

На фігурі 8 - зображено спектр звуку «Б» до фільтрації.

На фігурі 9 - зображено спектр звуку «Б» після фільтрації.

У таблиці 1 - представлені результати дослідження.

Розглянемо більш детально пропонований спосіб. Спосіб можна здійснити за допомогою пристрою, схема якого дана на фіг. 1. Пристрій складається з блоку керування 1, який здійснює контроль над всіма компонентами системи через шину даних. Вихід блоку керування 1 з'єднаний з буфером тимчасових даних 2, в якому здійснюється зберігання даних. Буфер тимчасових даних 2 з'єднаний зі всіма блоками системи, і здійснює функцію тимчасового зберігання даних. Вихід блоку керування 1 з'єднаний, також з буфером вхідних даних 3, де зберігається мовний сигнал у вигляді цифрового коду. Сигнал з буферу вхідних даних 3 надходить на блок ШПФ (Швидке Перетворення Фур'є) 4, який використовується для отримання спектральної характеристики сигналу. Вихід блоку ШПФ 4 з'єднаний з блоком фільтрації частот 5, де відбувається видалення частот, які знаходяться в області вище 1кГц. Сигнал з блоку фільтрації частот 5 надходить до блоку пошуку максимумів 6, де відбувається виділення всіх максимумів спектральної характеристики. Дані з блоку пошуку максимумів 6 поступають на блок ідентифікації формант 7, в якому виділяються формантні піки і фільтруються помилкові. З блоку ідентифікації

формант 7 сигнал надходить на блок фільтрації областей, що знаходяться на межі з формантним піком 8, згідно формулам 1.1, 1.2. Сигнал з блоку фільтрації областей, що знаходяться на межі з формантним піком, 8 надходить на блок результату обробки мовного сигналу 9.

Будь-який мовний сигнал складається з простих звуків мови, званих фонемами. Відомо, що кожній фонемі відповідає своя форма голосового тракту людини, яка варіюється відповідно до зміни положення язика, губ, зубів, залежно від цього міняються і частоти голосового тракту, у тому числі і резонансні. Області спектральних максимумів мовного сигналу, відповідні резонансним частотам голосового тракту, називаються формантами. Форманта характеризується частотою, шириною та амплітудою. Розпізнавання фонем базується на сприйнятті формант як найбільш інформативних ознак мовного сигналу. Також слід зазначити, що розпізнавання кожної фонемі відбувається в основному по положенню перших двох формант. Ґрунтуючись на особливостях сприйняття мовного сигналу, а саме на сприйнятті формант як найбільш інформативних ознак мови, а також враховуючи особливості сприйняття мови хворими на нейросенсорну тугоухість, такі як втрата сприйняття високочастотних компонент мови, зниження частотної селективності, пропонується видалити із спектру мовного сигналу області, що знаходяться на межі з формантними смугами. В ході експерименту оброблялося декілька різних сигналів різних дикторів. Мовний сигнал перетворювався в цифровий код і оброблявся на ЕОМ. Для ідентифікації і подальшої фільтрації формантних піків використовувався програмний пакет «Cool Edit Pro», який дозволяє проводити спектральний аналіз мовного сигналу. Оскільки найбільший вплив на розпізнавання мовного сигналу роблять перші дві форманти, здійснювалася фільтрація формантних областей перших двох піків. На фіг. 2 зображений моментальний спектр сигналу (одна виділена форманта) до фільтрації. На фіг. 3 зображений моментальний спектр після фільтрації.

Для створення необхідних фільтрів був проведений частотний аналіз голосних і приголосних звуків. Спектральна характеристика містить велике число піків окремих частот, проте велика їх частіша неінформативна і знаходиться в області вище 1500Гц. Основну мовну інформацію несуть у собі піки з відносно великою амплітудою у діапазоні 70Гц - 900Гц. На фіг. 4 представлений спектр звуку «С», де виразно помітні формантні піки у даному діапазоні.

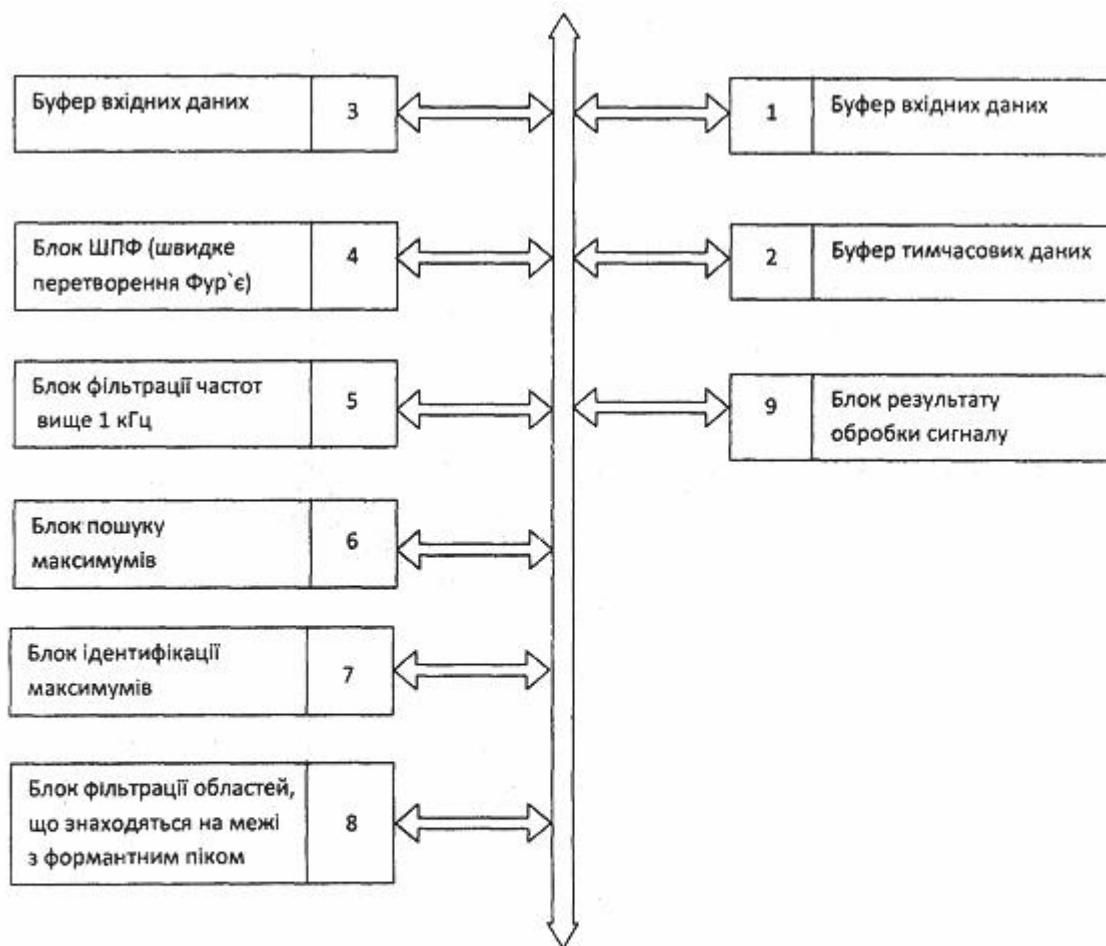
Видалення частотних смуг здійснювалося за допомогою фільтрів, які створювалися в опції меню FFT-фільтр (ШПФ - фільтр), згідно з розрахунковими даними. В результаті отримана спектральна характеристика сигналу, представлена на фіг. 5, звідки можна зробити висновок про звуження формантних областей, а також збільшення амплітуди сигналу в області перших двох формантних піків. На фіг. 6-9 представлені спектральні характеристики звуків Ж, Б до і після фільтрації. Таким чином, проводилася обробка кожного звуку всієї мовної послідовності. Оскільки падіння кривої чутності у хворих на нейросенсорну тугоухість починається з 1кГц, на спектр також накладався фільтр, що відсікає частоти вище 1кГц.

Для оцінки впливу пропонуємого способу обробки на сприйняття мовного сигналу, групі хворих на нейросенсорну тугоухість з 20 чоловік було запропоновано прослухати і порівняти два сигнали: первинний і такий, що пройшов попередню обробку. Всі хворі мали досвід користування слуховими апаратами декілька років. Слід зазначити, що хворі володіли невисокою початковою розбірливістю мови (менше 51%). Для оцінки розбірливості були використані мовні таблиці багатоскладових слів Грінберга, що враховують фонетичні особливості, характерні для російської мови. Вихідний сигнал не піддавався ніяким додатковим перетворенням, окрім регулювання гучності для комфортного сприйняття конкретним пацієнтом. Результати дослідження для 6 чоловік приведені в таблиці 1. Аналізуючи отримані дані можна зробити висновок про підвищення розбірливості мови у хворих в межах 5-18%. Результати дослідження для решти пацієнтів аналогічні, значення розбірливості знаходяться в межах вищезгаданого діапазону. Отже, в результаті фільтрації частотних смуг спектру, що знаходяться на межі з областю формантного піку, збільшується чіткість мовного сигналу. Виділення найбільш інформативних областей спектру сприяє поліпшенню сприйняття мовної послідовності, мова стає виразнішою і такою, що розуміється. Використання пропонованого способу обробки мовного сигналу дозволяє достовірно підвищити розбірливість мовного сигналу хворих на нейросенсорну тугоухість. Таким чином, це дає можливість створення адаптивного алгоритму підвищення словесної розбірливості.

Таблиця 1.

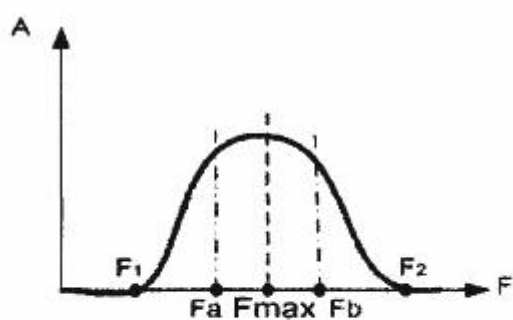
Результати дослідження

Хворі з діагнозом нейросенсорна тугоухість		Розбірливість початкового мовного сигналу (%)	Розбірливість обробленого мовного сигналу (%)
Ступінь тугоухості	Вік хворого		
VI	50	49	54
VI	71	45	55
II	43	50	50
III	45	50	65
IV	54	41	59
III	72	48	5 В



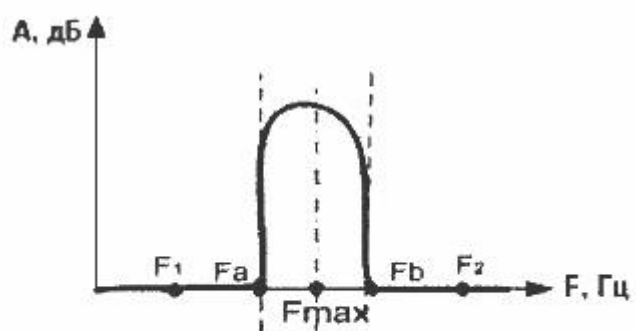
Фіг. 1

Схема пристрою, за допомогою якого здійснюється пропонований спосіб



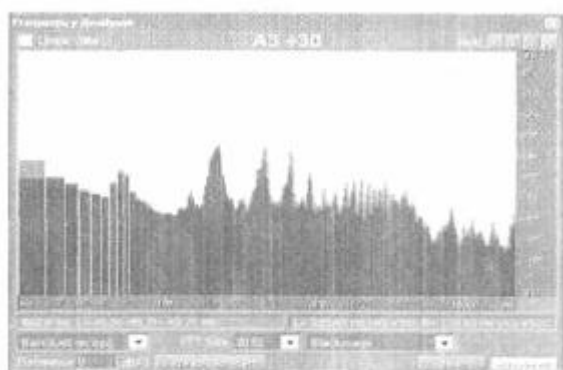
Фіг. 2

Формантна лінія до фільтрації



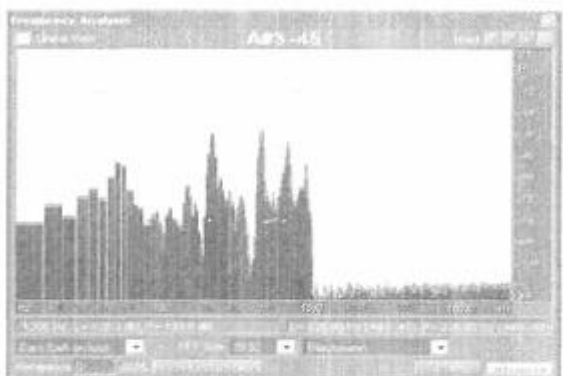
Фіг. 3

Формантна лінія після фільтрації



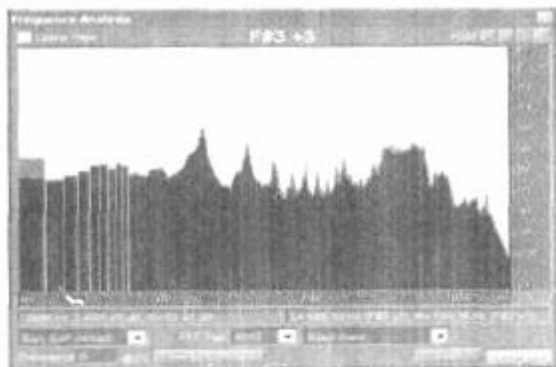
Фіг. 4

Спектр звуку «Є» до фільтрації



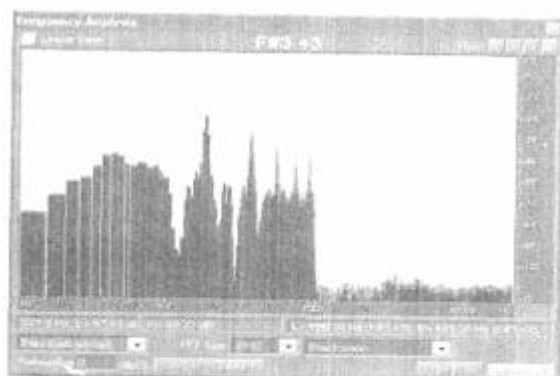
Фіг. 5

Спектр звуку «Є» після фільтрації



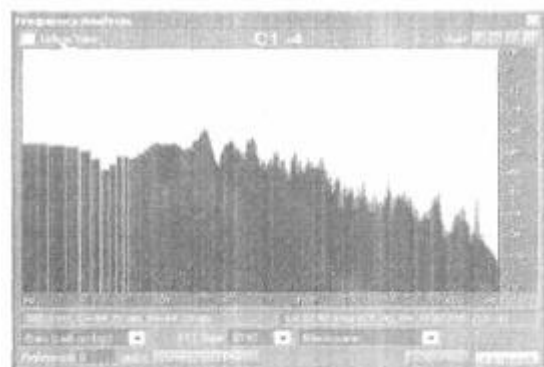
**Фіг. 6**

**Спектр звуку «Ж» до фільтрації**



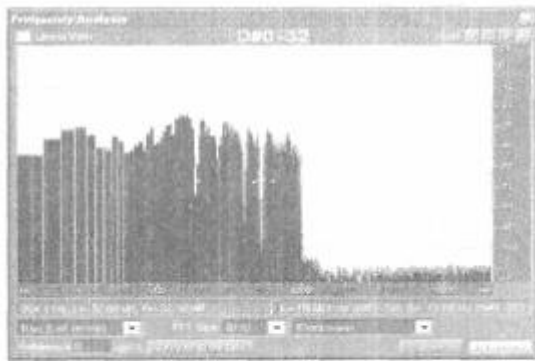
**Фіг. 7**

**Спектр звуку «Ж» після фільтрації**



**Фіг. 8**

**Спектр звуку «Б» до фільтрації**



**Fig. 9**

**Спектр звуку «Б» після фільтрації**