

АНАЛІЗ ВПЛИВУ АНАЛІЗОВАНИХ ДАНИХ ГОЛОСОВОГО СИГНАЛУ НА ЯКІСТЬ ОЦІНКИ ФОРМАНТНИХ ЧАСТОТ

Пастушенко М.О., Пастушенко І.Ю., Пастушенко М.С.

Кафедра математичних методів системного аналізу,
Національний технічний університет "Київський політехнічний інститут
імені Ігоря Сікорського", Україна.

Кафедра інфокомунікаційної інженерії ім. В.В. Поповського,
Харківський національний університет радіоелектроніки,
Україна.

E-mail: pastushenko.maksym@lil.kpi.ua,
ihor.pastushenko@nure.ua,
mykola.pastushenko@nure.ua

Abstract

The paper examines the main directions of improving the quality characteristics of voice authentication systems in various access and speech recognition systems by improving the procedures for extracting formant frequencies from registration materials. One of the main directions for improving the quality characteristics of voice signal processing systems considered in the work is related to the use of information about the registration phase of the materials and the use of the results of the autocorrelation function evaluation based on the data being studied.

Проблеми сучасних систем голосової автентифікації

У сучасному суспільстві голосові технології все частіше використовуються в різних сферах, від криміналістики до введення голосових команд в різні пристрої. Однак сучасні системи голосової автентифікації (СГА) мають в ряді застосувань недостатні якісні характеристики, в першу чергу, це стосується помилок першого і другого типів. На рівень помилок в значній мірі впливає якість оцінок формантних частот голосового сигналу, які входять до складу шаблону користувача системи автентифікації.

Постановка завдання дослідження

В якості методу оцінки формантних частот будемо використовувати спектральний аналіз, а саме, швидке перетворення Фур'є та подальшу порогову обробку отриманого спектру аналізованих даних [1, 2]. В якості експериментального сигналу застосуємо український звук «о» (жіночий голос). Частота дискретизації сигналу становила 64 кГц. Експериментальний голосовий сигнал представлений на рис. 1.

Порівняємо три варіанти оцінки частот формант. Перший варіант пов'язаний з традиційною оцінкою формантної інформації за матеріалами запису цифрового голосового сигналу. У цьому випадку використовуються амплітудно-частотні параметри голосового сигналу користувача.

Другий варіант пов'язаний з аналізом фазового спектру голосового сигналу користувача. Для цього за матеріалами реєстрації формується аналітичний сигнал за допомогою перетворення Гільберта і на основі останнього розраховується фазова інформація. В подальшому використовується швидке перетворення Фур'є та подальша порогова обробка для виділення формантних частот.

Третій варіант пов'язаний з розрахунком автокореляційної функції за матеріалами обробки експериментального сигналу. Тут використовується властивість автокореляційної функції, яка зберігає основні характеристики гармонійного коливання [3]. Тому в ряді застосувань використовують цю властивість функції автокореляції для збільшення відношення сигнал/шум.

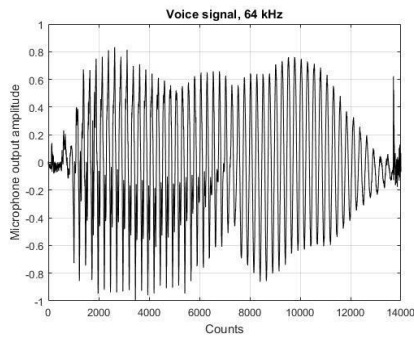


Рис. 1. Експериментальний сигнал

Аналіз результатів досліджень

Результати обробки експериментального сигналу наведені на рис. 2. Амплітудно-частотний спектр (АЧС) досліджуваного експериментального сигналу, розрахований за допомогою швидкого перетворення Фур'є, представлений на рис. 2.а. Аналіз представленого спектру дозволяє визначити вісім формантних частот. У цьому випадку виникають труднощі з визначенням частоти сьомої форманти.

На рис. 2.б. наведено фазово-частотний спектр (ФЧС) досліджуваного сигналу. У цьому спектрі можна виділити десять формантних частот. Як і в попередньому випадку, труднощі виникають при виділенні частоти сьомої форманти.

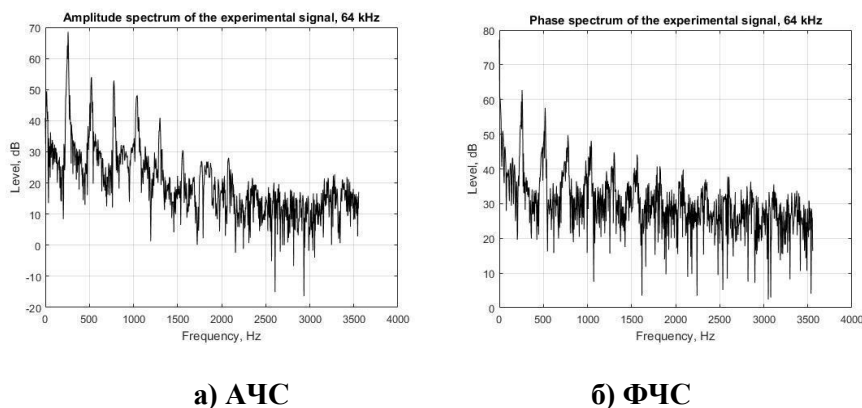


Рис. 2. Амплітудно-частотний та фазочастотний спектри експериментального сигналу

В наведених спектрах є складності з визначенням ширини спектру формантних частот. Особливо це відноситься до чотирьох перших формант, які визначають зміст мови та ознаки користувача.

Подальші дослідження пов'язані з розрахунком автокореляційної функції досліджуваного сигналу. Автокореляційну функцію можна розраховувати ітераційно [3]. На рис. 3.а наведено спектр після першої ітерації обчислення зазначеної функції. У цьому спектрі можна виділити дев'ять формантних частот. Сьома форманта, як і раніше, слабо виражена. Слід відмітити значне збільшення спектральної щільності, що впливає на якість визначення ширини спектру формантних частот.

На рис. 3б наведений спектр після трьох ітерацій розрахунку автокореляційної функції аналізованого сигналу. На цим рисунку бачим шість формантних частот, які мають дуже вузьку ширину спектру формантних частот.

Результати досліджень оцінок формантних частот можуть бути використані для покращення результатів, отриманих за даними АЧС.

Слід зазначити, що ці оцінки будуть корисними для уточнення ширини спектру виділених частот і оцінки огинаючої спектра. Ці оцінки займають особливе місце при уточненні частоти основного тону.

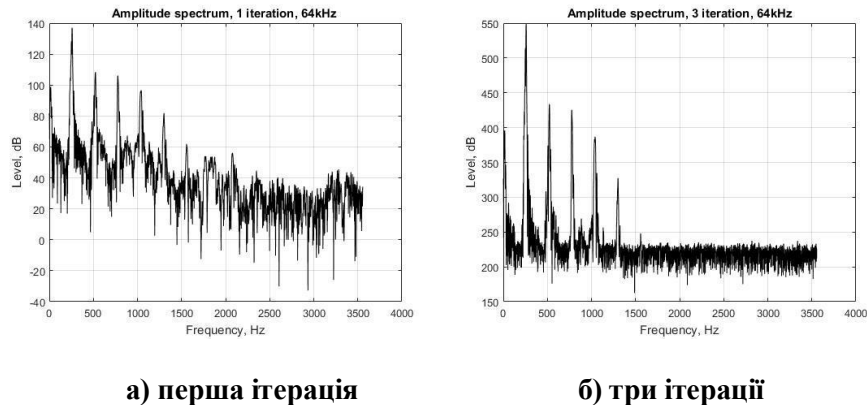


Рис. 3. Спектр експериментального сигналу, розрахований на основі автокореляційної функції

Висновки

Розглянуто актуальну наукову проблему оцінювання формантних частот при використанні в якості вихідних даних різних інформаційних параметрів мовного сигналу. Крім того, один із варіантів оцінювання застосовує відомі властивості функції автокореляції, які можна використовувати для уточнення ширини спектру формантних частот. Отримання близьких значень формантних частот підтверджує достовірність та адекватність результатів досліджень.

Об'єднання результатів розрахунків оцінок формантних частот за різними інформаційними параметрами сигналу приводить до їх уточнення, що надає можливість згенерувати кращий шаблон користувача для системи автентифікації, а це позитивно впливає на зменшення помилок першого та другого типів.

Подальші дослідження будуть зосереджені на об'єднанні отриманих оцінок, а також на уточненні огинаючої спектру.

Література

1. Özbek Arslan I., Demirekler M. Tracking of Visible Vocal Tract Resonances (VVTR) Based on Kalman Filtering. *9th International Conference on Spoken Language Processing/INTERSPEECH 2006*, Pennsylvania, Amerika Birleşik Devletleri, 01 Ocak, 2006. P. 1013-1016 Режим доступу: <https://hdl.handle.net/11511/53260>
2. Beigi H. *Fundamentals of Speaker Recognition*. NY: Springer, 2011. 942 p.
3. Van Trees H. *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. NY: Wiley, 1968. 716 p.
4. Пастушенко М.О., Пастушенко М.С., Романюк В.А. Методологія оцінки формантної інформації мовного сигналу системи автентифікації. *Міжнародна науково-практична конференція, Збірник тез доповідей*. Харків, 2024. С.255-256