Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
(повна назва)
Кафедра Інфокомунікаційної інженерії імені В.В. Поповського
(повна назва)

АТЕСТАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Дослідження кепстральних коефіцієнтів голосового сигналу користувача системи автентифікації
(тема)

Виконав:
студент 2 курсу, групи ІКІм-19-1 .
Заіка М. В. .
(прізвище, ініціали)

Спеціальність: 172 Телекомунікації та радіотехніка .
(код і повна назва спеціальності)
Тип програми: освітньо-професійна .
(освітньо-професійна або освітньо-наукова)
Освітня програма: Інфокомунікаційна інженерія .
(повна назва освітньої програми)

Керівник: професор кафедри ІКІ ім. В.В. Поповського
Пастушенко М.С.
(посада, прізвище, ініціали)

Допускається до захисту
Зав. кафедри _____ Лемешко О.В.
(підпис) (прізвище, ініціали)

2020р.

Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
(повна назва)
Кафедра Інфокомунікаційної інженерії імені В.В. Поповського
(повна назва)
Рівень вищої освіти другий (магістерський)
Спеціальність 172 Телекомунікації та радіотехніка .
(код і повна назва)
Тип програми освітньо-професійна .
(освітньо-професійна або освітньо-наукова)
Освітня програма Інфокомунікаційна інженерія .
(повна назва)

ЗАТВЕРДЖУЮ
Зав. кафедри_____
(підпис)
«____» _____2020р.

ЗАВДАННЯ
НА АТЕСТАЦІЙНУ РОБОТУ

студентові Заіка Максим Володимирович .
(прізвище, ім'я, по батькові)

1. Тема роботи: Дослідження кепстральних коефіцієнтів голосового сигналу користувача системи автентифікації.
затверджена наказом по університету від « 20 » жовтня 2020р. № 1396 Ст.
2. Термін подання студентом роботи до екзаменаційної комісії 10.05.2020р.
3. Вихідні дані до роботи: ISO/IEC TR 24741:2007 Information technology – Biometrics tutorial (ГОСТ Р 54412-2011), ISO/IEC/TR 24722:2007 Information technologies. Biometrics. Multimodal and other multibiometric fusion. (ГОСТ Р 54411-2011)
4. Перелік питань, що потрібно опрацювати в роботі:
1) Аналіз поточного стану біометричних систем автентифікації користувачів.
2) Процедури цифрової обробки, математична модель та схема проведення експериментальних досліджень голосових сигналів.
3) Результати експериментальної оцінки мел-частотних кепстральних коефіцієнтів за амплітудною та фазовою інформацією.
4) Пропозиції щодо використання фазової інформації при вдосконаленні систем голосової автентифікації.

5. Перелік графічного матеріалу із зазначенням креслень, плакатів, комп'ютерних ілюстрацій: Демонстраційний матеріал у вигляді ppt-презентації; структурна схема біометричної системи; методика експериментальних досліджень; результати оцінки мел-частотних кепстральних коефіцієнтів голосового сигналу.

6. Консультанти розділів роботи

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		(підпис)	(дата)
Основна частина	професор Пастушенко Микола Савелійович		

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання	05.10.2020	Виконано
2	Збір матеріалів для дослідження	01.11.2020	Виконано
3	Розробка 1 розділу	05.11.2020	Виконано
4	Розробка 2 розділу	25.11.2020	Виконано
5	Розробка 3 розділу	05.12.2020	Виконано
6	Оформлення атестаційної роботи	10.12.2020	Виконано

Дата видачі завдання 5 жовтня 2020 року

Студент _____ Заіка М.В.
(підпис) (прізвище, ініціали)

Керівник роботи _____ професор Пастушенко М.С.
(підпис) (посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: 81 с., 22 рис., 1 табл., 73 джерел

АВТЕНТИФІКАЦІЇ, АНАЛІЗ, БІОМЕТРІЯ, НАДІЙНІСТЬ, ПАРОЛЬ, СИСТЕМА ДОСТУПУ, ФАЗА СИГНАЛА.

Об'єктом дослідження є процес цифрової обробки сигналів в голосових системах автентифікації.

Метою даної роботи є дослідження напрямків підвищення якості голосових систем автентифікації на основі використання фазових даних.

Методи досліджень – аналіз, спостереження, вимірювання, моделювання та експеримент, узагальнення результатів і формування висновків.

В роботі виконано аналіз поточного стану систем голосової автентифікації користувачів. Розглянуто їх переваги та недоліки. Особливу увагу приділено використанню голосових систем автентифікації користувачів при побудові сучасних інформаційно-комунікаційних систем і локальних мереж, які мають суттєві переваги в порівнянні з випадком, коли в якості системи доступу використовуються інші методи автентифікації користувача.

Обґрунтовано, що основним напрямком підвищення якості систем голосової автентифікації є використання фазових даних оброблюваних матеріалів реєстрації. На прикладі оцінки мел-частотних кепстральних коефіцієнтів проведено дослідження фазового спектру голосового сигналу користувача, отриманого в процесі модельного експерименту. Виконані дослідження і розроблені процедури більш ефективної оцінки мел-частотних кепстральних коефіцієнтів за рахунок використання фазових даних голосового сигналу користувача системи автентифікації.

ABSTRACT

The report contains: 81 p., 22 pictures, 1 tables, 73 sources

AUTHENTICATION, ANALYSIS, BIOMETRICS, PASSWORD, ACCESS SYSTEM, RELIABILITY, SIGNAL PHASE.

The object of the study is the process of digital signal processing in voice authentication systems.

The purpose of this work is to investigate the directions of improving the quality of voice authentication systems based on the use of phase data.

Research methods – analysis, observation, measurement, simulation and experiment, generalization of results and conclusions.

The current state of users' voice authentication systems is analyzed. Their advantages and disadvantages are considered. Particular attention is paid to the use of voice authentication systems for users in the construction of modern information and communication systems and local networks, which have significant advantages over the case when other methods of user authentication are used as the access system.

It is substantiated that the main direction of improving the quality of voice authentication systems is the use of phase data of processed registration materials. The phase spectrum of the user's voice signal obtained in the process of the model experiment was studied on the example of estimation of chalk-frequency keptral coefficients. Researches are carried out and procedures of more effective estimation of chalk-frequency keptral coefficients due to use of phase data of a voice signal of the user of authentication system are developed.

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАК, СИМВОЛІВ, ОДИНИЦЬ І ТЕРМІНІВ

БД – база даних
БНЯ – безпека, надійність і якість
БШП – блок швидкого перетворення
ВЕВ – віддалена електронна взаємодія
ДІ – достовірність ідентифікації
ЕП – електронний підпис
ІА – ідентифікації та автентифікації
ІС – інформаційних систем
МО – математичне очікування
СКВ – середньоквадратичне відхилення
СКПЕП – сертифікат ключа перевірки електронного підпису
СГА – система голосового автентифікації
СКМ – система комп'ютерної математики
ТМЗІ – технологій, механізмів і засобів ідентифікації
ЧОТ – частота основного тону
ШПФ – швидке перетворення Фур'є
EER – Equal Error Rate
CELP – Code Excited Linear Prediction
IBIA – International Biometric Industry Association
FAR – False Acceptance Rate
FRR – False Rejection Rate
GMM – Gaussian Mixture Model
LPC – Linear Prediction Coding Coefficients
MFCC – mel-frequency cepstral coefficients
HMM – Hidden Markov Models
SVM – Support Vector Machine

ВСТУП

В даний час загострюється проблема забезпечення безпеки фінансів, інформації, послуг і ресурсів, доступ до яких здійснюється за допомогою сучасних телекомунікаційних і комп'ютерних систем різного призначення. Про це свідчать численні періодичні повідомлення в пресі. Тут же слід зауважити, що західні фінансові установи намагаються не афішувати випадки розкрадання коштів до 100 тисяч USD.

Після драматичних подій 11.09.2001 року країнами G8 прийнято рішення, яке орієнтоване на зниження ризиків і підвищення ефективності систем доступу різного призначення. Для цього запропоновано використовувати в таких системах біометричні характеристики користивуча. В якості основних характеристик рекомендовано використовувати фізіологічні (статичні) ознаки користувача, а саме, папілярний візерунок пальців, зображення особи і райдужну оболонку ока.

Обумовлено це тим, що дактилоскопія, як і фото, дуже широко і ефективно застосовуються в криміналістиці для ідентифікації злочинців. Більш того, накопичені і інтенсивна поповнюються великі бази відбитків, особливо в Західних країнах. Тут же зауважимо, що зазначені статичні біометричні ознаки мають обмежену інформативність.

За останнє десятиліття біометричні технології стали активно застосовуватися в багатьох областях, пов'язаних із забезпеченням безпеки доступу до інформації та матеріальним об'єктам, а також в задачах унікальної ідентифікації особистості. Багато в чому цьому сприяло поширення мікропроцесорних технологій.

Разом з тим, запропоновані біометричні ознаки не дозволили істотно підвищити надійність систем доступу. Обумовлено це тим, що як в криміналістиці, так і будь-який біометричної системи основними характеристиками є два числа – False Acceptance Rate (FAR, помилковий доступ в систему) і False Rejection Rate (FRR, помилковий відмова в доступі) [1]. Стосовно до системи автентифікації перше число характеризує ймовірність помилкового збігу біометричних характеристик двох людей. Друге - ймовірність відмови доступу людині, що має допуск. Система тим краще, чим менше значення FRR при однакових значеннях FAR. На відміну від криміналістики, система автентифікації повинна бути стійка до підробки. Останнє не притаманне криміналістиці. Стійкість до підробки - це емпірична

характеристика, яка узагальнює те, наскільки легко обдурити біометричну систему.

Стосовно до статичних біометричним ознаками можна констатувати, що вони не задовольняють вимогам по стійкості до підробки. Наприклад, давно відомі досліди японського криптографа Цутому Мацумото (Tsutomu Matsumoto), які дозволили розкрити від 80 до 100% тестованих дактилоскопічних систем доступу. Через низьку стійкість переходять від просторових до тривимірним зображенням обличчя користувача. З'явилися повідомлення про підробку райдужної оболонки ока, які формують на основі декількох фото за допомогою сучасної фотоапаратури.

У зв'язку з цим все більше уваги приділяється поведінковим (динамічним) ознакам користувача, а саме, підпис (форма букв, манера письма, натиск), голос, клавіатурний почерк і ін [2]. Основна перевага зазначених ознак – оперативне нарощування аналізованої послідовності на вимогу системи. У загальному випадку можна стверджувати, що зазначені ознаки мають необмежену інформативність. Це істотно впливає на зниження величин FRR і FAR, а також підвищує стійкість до підробки.

Зазначене перевага особливо яскраво проявляються для систем голосової автентифікації (СГА). Поряд із зазначеним вище, СГА мають ряд додаткових переваг, таких як: простота, компактність, дешевизна, можливість віддаленої автентифікації з використанням телефонних каналів зв'язку та ін. Питанням розвитку і впровадження цих систем сьогодні присвячені численні дослідження і розробки, окремі питання яких розглянуті в роботах Г. Фанта, Р.М. Болла, Г.С. Рамішвіллі, Ф. Россе, В.Н. Сорокіна, Г. Холлиена та ін.

Тут же слід зазначити, що в сучасних СГА для ідентифікації використовуються переважно спектральні характеристики амплітуди і частоти мовного сигналу користувача, а фаза матеріалів реєстрації ігнорується. При цьому, давно відомо, що фаза є більш інформативним параметром реєстрованого сигналу.

Тому в роботі досліджується актуальна науково-технічна задача підвищення якісних характеристик СГА за рахунок використання фазових даних голосового сигналу.

Виконаний аналіз робіт в області голосової автентифікації показав, що відкритим залишається питання оцінки мел-частотних кепстральних коефіцієнтів (MFCC - Mel-frequency cepstrum coefficients) за фазовою інформацією голосового

сигналу користувача системи автентифікації. Тому мета даної роботи – аналіз інформативності фазової інформації голосового сигналу при оцінці мел-частотних кепстральних коефіцієнтів. Результати роботи опубліковані в [3-5].

1 ЗАГАЛЬНА ХАРАКТЕРИСТИКА ЗАВДАНЬ ІДЕНТИФІКАЦІЇ І АВТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ ІНФОКОМУНІКАЦІЙНИХ СИСТЕМ

Вплив інформаційної сфери на розвиток сучасного суспільства безперервно зростає. У зв'язку з цим забезпечення інформаційної безпеки стає одним з пріоритетів національної політики держави. Зміст проблем, що об'єднуються поняттям «інформаційна безпека», в останній період часу визначається перш за все швидким поширенням нових інформаційних технологій. Ці технології призвели до проникнення комп'ютеризації в усі сфери людської діяльності, вони збільшують залежність від інформаційних систем і послуг, а також створюють нові види загроз для інтересів окремої особистості, підприємств і організацій, суспільства в цілому.

1.1 Сучасні підходи до вирішення завдань ідентифікації та автентифікації користувачів інфокомунікаційних систем

Вразливість приватних осіб, організацій і держави по відношенню до загрозам інформаційної безпеки особливо зростає при використанні інформаційних мереж – як загального користування, так і корпоративних. Цьому сприяє також поширена тенденція до розподіленої обробки даних, пов'язана з використанням дистанційного режиму і телекомунікаційних технологій (зокрема, розширюється сфера діяльності співробітників і залучених осіб поза відповідної організації).

Все більших масштабів приймають кримінальні напрямки комп'ютерної діяльності. Згідно уніфікації Комітету міністрів Європейської Ради, до цих напрямків можна віднести комп'ютерне шахрайство, несанкціонований доступ до інформації, підробку комп'ютерної інформації, несанкціоноване перехоплення даних і інші види злочинних дій. У зв'язку з цим найважливішим завданням стає створення і застосування нових ефективних методів і засобів захисту інформації.

Розвиток нових методів і засобів забезпечення інформаційної безпеки покликане, перш за все, запобігти загрозам доступу до інформаційних ресурсів сторонніх осіб, які не мають доступу. Для вирішення цього завдання необхідна наявність ідентифікаторів і створення процедур ідентифікації для всіх користувачів.

Сучасні ідентифікація та автентифікація включають в себе різні системи, в тому числі, і способи біометричної ідентифікації особистості.

Розвиток систем ідентифікації особистості, заснованих на біометричних вимірах, пов'язане з цілим комплексом переваг: такі системи більш надійні, оскільки біометричні показники складніше підробити; сучасна мікропроцесорна техніка робить біометричні методи (невід'ємні біометричні ідентифікатори) більш зручними у порівнянні зі звичайними методами ідентифікації (електронними ідентифікаторами); нарешті, вони значно простіше піддаються автоматизації вимірювань.

Методи біометричної ідентифікації зазвичай поділяють на дві групи: фізіологічні та поведінкові (враховують підсвідомі дії людини). До фізіологічних методів ідентифікації відноситься використання таких характеристик, як відбитки пальців рук або долонь, райдужної оболонки або сітківки ока, 2-мірних і 3-мірних зображень особи тощо.

Та чи інша конкретна біометрична технологія може володіти певними перевагами в залежності від різних конкретних умов і вирішуваних завдань. Однією з найбільш поширених біометричних характеристик людини є його голос, що володіє набором індивідуальних особливостей, відносно легко піддається виміру (наприклад, частотний спектр голосового сигналу). До переваг голосової ідентифікації відносяться також зручність застосування і використання, досить невисока вартість пристроїв, застосовуваних для ідентифікації (наприклад, мікрофонів) та ін.

Необхідність розмежування доступу до постійно зростаючих обсягів інформації в сучасному світі гостро ставить проблему перевірки справжності користувача ресурсів. Зростання інфокомунікаційних мереж та інтенсивності їх використання в повсякденному житті суттєво спрощує завдання зловмисника по отриманню несанкціонованого доступу до даних або певним сервісів, що надаються телекомунікаційними системами.

Окрім забезпечення розмежування прав доступу до конфіденційної інформації, сучасні автоматизовані системи вирішують ряд суміжних завдань. Двома основними процедурами, виконуваними подібними системами, є ідентифікація і автентифікація суб'єкта доступу. В загальному випадку таким суб'єктом для аналізованих систем може бути не тільки конкретна людина, а й будь-який обчислювальний процес, виконуваний видалено або локально.

Однак, за даними аналізу статистики експертами в області комп'ютерної безпеки більшість випадків, пов'язаних з витоком конфіденційної інформації пов'язані з факторами, до яких безпосередньо причетний конкретна людина. Причинами доступу третіх осіб до конфіденційних даних на порядок частіше ставала робота інсайдерів і хакерів, ніж дії шкідливого програмного забезпечення. Таким чином, можна зробити висновок про першорядної важливості проблеми ідентифікації і автентифікації в області комп'ютерної безпеки.

У більшості сучасних інформаційних систем перевірка особистості користувача здійснюється за допомогою введення логіна і пароля. А нині існують і інші методи, які хоча і не отримали такого широкого поширення, потенційно є набагато більш надійними. Зокрема, останнім часом, широко використовується цілий клас перспективних біометричних систем.

В основі науки про ідентифікацію особистості лежать ідеї вимірювання тіла людини і його частин. Ці ідеї вперше сформулював французький криміналіст Альфонс Бертільон (Alphonse Bertillon) – співробітник паризької префектури, який займався реєстрацією злочинців. У 1879 р він представив систему ідентифікації злочинців, яка отримала назва антропометрії і включала в себе: вимір їх зріст людіни, довжини і обсягу її голови, довжини рук, пальців, стоп тощо, а також словесний портрет злочинця, фото портрет в анфас і в профіль, а також опис особливих прикмет.

Сучасна криміналістика досі також використовує цю систему, доповнивши її антропоскопії, дактилоскопії, фотороботами, новими методами опису особливих прикмет на обличчі й тілі людини і технологіями їх реалізації. Однак поняття біометрії як окремої науки було сформульовано десятиліттям пізніше. Біля витоків ранньої біометрії стояв англійський дослідник Френсіс Гальтон (Francis Galton). У книзі, присвяченій природної спадковості і виданої в 1889 р, він вперше ввів поняття біометрії (biometry) як науки, що займається кількісними біологічними експериментами з залученням методів математичної статистики.

У той же час, термін «Biometrics» – біометріка, з'являється в англомовній літературі, як нова гілка біометрії, яка охоплює галузь знань, що представляє методи вимірювання персональних фізичних і поведінкових характеристик людини і методи їх використання для цілей ідентифікації або автентифікації. У російськомовній науково-популярній літературі термін «біометріка» (це відповідає прямому перекладу англійського слова «biometrics») так само зустрічається. Однак в

тому ж значенні в російськомовній літературі використовується і термін «біометрія», який, в свою чергу, хоч і з'явився як переклад слова «biometry», що в англійській традиції позначає «біостатистика», в україномовних документах повноцінно використовуються і в значенні «біометріка».

Таким чином, в сучасній науковій літературі часто обидва терміни і «біометрія» і «біометріка» використовуються для цілей позначення біометричної ідентифікації і автентифікації. Але, звичайно, варто звертати увагу на контекст, тому термін «біометрія» теж може бути використаний в значенні «біостатистика».

В зв’язку зараз вважають, що біометрія – це наука, заснована на описі і вимірі характеристик тіла живих істот. У застосуванні до систем автоматичної ідентифікації під біометричними розуміють ті системи і методи, які засновані на використанні для ідентифікації або автентифікації будь-яких унікальних характеристик людського організму. Наше життя наповнене ситуаціями, коли нам потрібно довести, хто ми. Такими ситуаціями наповнена як особиста, так і професійна сфера.

Неважко перерахувати широкий спектр галузей які вимагають швидкою, надійною і зручною автентифікації користувача: доступ до персонального комп'ютера або смартфону, доступ до електронної пошти, банківські транзакції, відкриття дверей і запуск двигуна вашого автомобіля, контроль доступу в приміщення, перетин державних кордонів, і, взагалі, як правило, будь-яка взаємодія з державними органами влади вимагає ідентифікації.

Таким чином, ідентифікація і автентифікація нашої особистості стали наріжним каменем в сучасному суспільстві, забезпечуючи безпечну взаємодію, запобігаючи шахрайство і злочинність.

Біометричну ідентифікацію часто називають чистою або реальною автентифікацією, так як використовується не віртуальний, а реально має ставлення до людини біометричну ознаку (ідентифікатор). Специфічною особливістю біометричної ідентифікації буде великий розмір біометричної бази даних. Оскільки, кожен з біометричних зразків повинен бути зіставлений з усіма записами, які є в базі даних (зіставлення 1:N або «один до багатьох»). Для використання в реальному житті така система вимагає високої швидкості зіставленні біометричних ознак.

Наприклад, чисельність співробітників навіть великого підприємства від декількох сотень до декількох тисяч чоловік. Візьмемо для прикладу чисельність співробітників 10 000 чоловік. Значить розмір бази даних (виходимо що для однієї

людини використовується один відбиток пальця) становитиме 10 000 відбитків пальців. При прикладанні пальця до зчитувача відбитків система буде виробляти зіставлення 1:10 0000. Що дуже небагато для сучасних систем. Саме тому всі системи контролю доступу або обліку робочого часу працюють в режимі біометричної ідентифікації.

На іншій стороні полюса – верифікаційні системи, вони здійснюють, як правило, тільки одне зіставлень в режимі 1:1. Тобто пред'явлена біометрична ознака порівнюється з одною біометричною ознакою з бази даних. Тобто система відповідає на питання, чи той ти за кого себе видаєш.

Термін «біометрична автентифікація» використовується досить часто, незважаючи на його важливість, досить часто виникає плутанина. Тому що в різних типах систем визначення цього терміну відрізняються, наприклад в банківських і юридичних системах. Дамо визначення цих термінів для біометричних систем. Автентифікація (від англійського – authentication) – процедура перевірки приналежності суб'єкту доступу пред'явленого їм ідентифікатора.

Простий приклад автентифікації – підтвердження особи користувача шляхом порівняння введеного ним логіна з паролем в базі даних ідентифікованих раніше користувачів. В даному прикладі автентифікацією є процес порівняння паролів, і подальше або надання доступу або відмову, а ідентифікатором буде як раз логін.

Способи автентифікації були згруповані в 1994 році в три основні категорії, засновані на так званих факторах автентифікації (див. табл. 1.1):

- з використанням знань – користувач вводить в систему якусь секретну фразу, наприклад, логін і пароль;

- з використанням власності – користувач пред'являє системі деякий фізичний предмет, наприклад, пропуск, смарт-карту або usb-токен;

- з використанням своїх характеристик (ознак людини) – користувач пред'являє системі свої фізіологічні або поведінкові параметри.

Кожен фактор автентифікації охоплює ряд елементів, використовуваних для автентифікації або перевірки особи до надання доступу, затвердження запиту транзакції, підписання документа, надання повноважень іншим тощо.

Перше масове застосування біометрії запустила компанія Apple 10 вересня 2013 року, представивши публіці вбудований в iPhone 5s зчитувач відбитків пальців – Touch ID. Обсяг продажів за 2017 склав 1,5 млрд штук.

Друга віха – біометричні паспорти. У ряді країн після 2000 року почали видавати паспорти нового покоління містять електронний носій інформації – безконтактний чіп. Дані на чіпі будь-якого паспорта захищені за допомогою технології контролю доступу BAC (Basic access control) і містять, як правило, фотографію власника паспорта, відбитки пальців, інформацію про дату і місце народження власника, дату видачі паспорта та орган, що видав документ.

Таблиця 1.1 – Основні категорії способів автентифікації

Фактори знання (логін, пароль)	Фактори володіння (сім-карта, телефон)	Фактори ознак (відбитки пальців)
		

Очевидно, що така популярність, може бути продиктована тільки перевагами над будь-якими іншими методами ідентифікації та автентифікації.

Переваги біометрії вже привели до широкого поширення сенсорів відбитків пальців в мобільних пристроях, таких як смартфони і планшети. Але типів біометричних технологій набагато більше, ніж тільки відбиток пальця, в найближчому майбутньому вони отримають найширше поширення.

1.2 Поточний стан систем біометричної ідентифікації

У зв'язку з інтенсифікацією інформатизації сучасного суспільства, переходом до хмарних обчислень і залученням до цих процесів різних державних органів, підприємств, організацій і громадян дуже актуальними стають питання достовірної ідентифікації та автентифікації (ІА) учасників електронної взаємодії. розвиток і модернізація інформаційних систем (ІС), що містять відкриту інформацію та інформацію обмеженого доступу різного рівня, а також необхідність їх більш тісної взаємодії ставить одним з першочергових завдання організації захищеного авторизованого доступу користувачів до інформаційних ресурсів, в тому числі містить конфіденційну інформацію.

Згідно з державними законами інформація, в тому числі і про стан здоров'я громадянина є однією з найбільш чутливих до розголошенню видів конфіденційної інформації. управління доступом користувачів неможливо без коректного вирішення завдань ІА. розвиток системи державних і муніципальних послуг, електронної торгівлі, дистанційного банківського обслуговування і освіти, а також електронної охорони здоров'я (e-Health) вимагає створення і практичного застосування надійних методів визначення сторін віддаленої електронної взаємодії (ВЕВ), що дозволяють с певною часткою впевненості говорити про достовірність ідентифікації особистості, як правило, виступає однією зі сторін ВЕВ.

Одним з методів ІА особистості, що найбільш інтенсивно розвивається, є ідентифікація по біометричним характеристикам. Біометрія привертає розробників тим, що користувачеві не треба запам'ятовувати або записувати ідентифікаційну і автентифікаційну інформацію. За останні два десятиліття розроблено кілька десятків методів ідентифікації. У маркетингових матеріалах виробники наводять дуже привабливі дані по точності ідентифікації, проте на практиці ці дані виявляються завищеними.

Програмні біометричні рішення розвиваються в світі випереджаючими темпами в порівнянні з апаратними. так оцінює перспективи ринку біометрії експерти компанії Mercator Advisory Group.

До недавнього часу системи ідентифікації користувачів і контролю доступу, побудовані переважно на програмних засобах, які не були розвинені. У цій ситуації комерційні замовники, наприклад банки, вкладалися в апаратні рішення з поверненням інвестицій через період близько п'яти років. Тепер це стає для них менш вигідним, оскільки програмні, швидко окупаються рішення не поступаються апаратним по надійності і зручності використання, але дають набагато більш швидку віддачу.

Ідентифікація за відбитками пальців продовжує залишатися найбільш поширеним видом біометрії. Друге місце серед апаратних рішень займають системи на основі райдужної оболонки ока.

Програмно-орієнтовані засоби – це, в першу чергу, швидко проникає в багато сфер розпізнавання осіб, а також просте і зручне, але поки недооцінене розпізнавання голосу. ціле сімейство біометричних факторів з програмною реалізацією складають поведінкові технології ідентифікації.

Загальне збільшення числа смартфонів і їх проникнення практично в усі сфери життя сприяє впровадженню біометрії і впливає на характер її використання. Доступ до самого смартфону також все частіше контролюється біометричними засобами. За рахунок цього відбувається швидше прилучення користувача до біометрії. Далі ідентифікація по біометричним ознаками легше поширюється по різним вертикальним ринкам. Сама автентифікація людини за біометричними ознаками завдяки поширенню смартфонів стає мобільною.

Згідно із прогнозом дослідницької компанії Acuity Market Intelligence, в 2019 році всі випущені смартфони будуть обладнані біометричними технологією. До 2022 року на руках у користувачів залишаться тільки такі смартфони [6].

Експерти стверджують, що в майбутньому автентифікація буде не разовою подією, а постійним процесом, який не потребує будь-яких дій з боку користувача. Вона буде багатофакторної, використовує одночасно географічну локацію, аналіз обстановки, розпізнавання обличчя і голосу і поведінкові характеристики людини.

Біометричний ринок досяг переломною точки – бажання різних організацій краще автентифікувати користувачів тепер цілком відповідає неприйняття користувачами систем, в яких від них потрібно що-небудь пам'ятати, наприклад пароль. В результаті біометричні рішення із збільшеною швидкістю проникають в призначені для користувача, комерційні та державні системи.

Так описують поточний стан ринку біометрії фахівці компанії Tractica. У проведеному ними дослідженні обраховані і обговорені основні тенденції в розвитку даного ринку. Робота спиралася на докладний аналіз 142 сценаріїв використання біометрії.

Дослідники вважають, що в розвитку таких сценаріїв укладені драйвери зростання ринку біометрії. Окремо вивчалися бізнес-функція кожного сценарію, його додаток до певного сегменту економіки і процедура. Далі аналізувалася реалізація сценаріїв використання біометрії в різних регіонах світу, з урахуванням відмінностей в їх економічному розвитку.

Згідно з даними, які зібрала компанія, загальний обсяг продажів апаратних і програмних засобів біометрії склав в 2016 році у всьому світі $ 2,4 млрд. Фахівці компанії прогнозують зростання цього показника на 23% в рік. В такому випадку світовий обсяг продажів біометричних систем до 2025 року перевищить $ 15 млрд, а всього за цей період буде випущено продукції на суму $ 70 млрд [6].

Традиційно в питаннях застосування біометричних методів споживча середовище та організації чітко відділялися один від одного, проте зараз це розмежування зникає. Фінансові організації, наприклад, дають можливість і фізичним особам, і представникам юридичних осіб проходити автентифікацію в онлайнових банківських системах через розпізнавання голосу або райдужної оболонки ока і не вводити паролі з клавіатури.

Сегменти ринку біометрії, на які припадає найбільший обсяг продажів, – це, як визначили фахівці компанії Tractica, розпізнавання відбитків пальців, голосу, райдужної оболонки ока і осіб. Найбільші вертикальні ринки – споживчий, фінансовий, державний і промисловий.

За даними консалтингової компанії Frost & Sullivan, інтеграція штучного інтелекту і глибокого навчання підвищить ефективність біометричних методів, що має привести до зростання європейського ринку біометрії з 4,98 до 11,5 млрд доларів до 2023 року [6].

Нове дослідження «Європейський ринок біометрії, прогноз до 2023 року» аналізує поточні і майбутні ринкові тенденції в урядовому і комерційному сегментах. Згідно зі звітом, розпізнавання відбитків пальців залишається домінуючим фактором на ринку, хоча після 2017 року розпізнавання осіб стало основним напрямком галузі для банківських і фінансових послуг.

«Поведінкова» біометрична характеристика, заснована на штучному інтелекті, стане основним елементом автентифікації особистості в двухфакторної або багатофакторної автентифікації, – говорить аналітик Frost & Sullivan Рам Рави.

Аналітики також прогнозують велике майбутнє за біометричними технологіями, заснованими на блокчейне.

Дослідження прогнозує майбутнє за тими постачальниками біометрії, які будуть пропонувати єдиний комплекс біометричних послуг, який вирішує весь спектр проблем замовника, та розробляють програмне забезпечення і технологічні рішення під конкретні цілі та потреби замовника.

1.3 Якісні характеристики систем ідентифікації і автентифікації

Результати ідентифікації (порівняння пред'явленого значення ідентифікатора з занесеним в базу даних при реєстрації користувача значенням) за своєю природою повинні бути предметом вивчення за допомогою теорії ймовірності. Іноді

пропонується введення рівнів довіри до результатів ідентифікації (апріорі імовірнісна характеристика) і автентифікації (надійність і якість якої теж є ймовірнісної характеристикою). Для спрощення завдання можна ввести 2 рівня ідентифікації: спрощена (по паролю) і стандартна, а також 3 рівня автентифікації: проста, посилена і сувора.

Такий підхід узгоджується з дослідженням процесів і результатів ІА, проведених на основі аналізу ризиків ідентифікації і автентифікації і аналізу надійності. Слід розрізняти достовірність ідентифікації (ДІ) при первинному зверненні суб'єкта (реєстрації нового користувача) в ІС і ДІ при вторинних (найчастіше повторних) зверненнях. При цьому слід враховувати, що ДІ особистості при первинному зверненні заявника залежить від наступних факторів:

- якість ідентифікації – відмінність одного суб'єкта від іншого шляхом порівняння пред'явлених ідентифікаторів з даними, занесеними в базу при реєстрації;

- в процесі ідентифікації є помилки першого (зловмисник ідентифікований як легальний user) і другого роду (легальний користувач не ідентифікований);

- необхідність введення рівнів довіри до результатів порівняння в залежності від числа ідентифікаторів і, головне, – від надійності і безпеки механізмів порівняння;

- необхідність протоколювання результатів процесів підтвердження збігів ідентифікаторів з державних баз даних для розбору конфліктних ситуацій.

При вторинної ідентифікації (повторних зверненнях до ресурсів ІС) процес ідентифікації зводиться до процедури порівняння пред'явлених користувачем ідентифікаторів з занесеними раніше даними в інформаційну базу при реєстрації. У найпростішому випадку це може бути один ідентифікатор (наприклад, логін), в більш складних схемах ідентифікації це може бути послідовна процедура пред'явлення заданого системою кількості ідентифікаторів. Наприклад, при первинному зверненні громадянина до порталу держпослуг, як мінімум, необхідно пред'явити номер паспорта.

Достовірність автентифікації також важлива як для управління доступом користувача до прикладного програмного забезпечення, що викликає процедуру електронного підпису, так і для організації процедури волевиявлення власника засобом електронного підпису (ЕП) в момент підписання документа або повідомлення.

Пропоновані вище прості рівні автентифікації можуть також бути розбиті на підрівні достовірності залежно від використовуваних технологій і механізмів автентифікації. По суті, ці рівні також пов'язані з ризиками авторизації зловмисника під ім'ям легального користувача. Застосовуваний в західних нормативних актах і стандартах (обов'язкових до виконання) термін «ідентифікація» включає в себе як ідентифікацію, так і автентифікацію.

Всі біометричні методи засновані на імовірнісних та статистичних методах [7]. Надійність методів може оцінюватися кількома способами, в найбільш поширеному підході в якості основних характеристик можна прийняти помилки першого і другого роду. Помилка першого роду (FRR – False Rejection Rate) – це ймовірність помилкового відмови в доступі користувачеві, що має право доступу. Помилка другого роду (FAR – False Acceptance Rate) – це ймовірність помилкового доступу, коли система помилково пізнає чужого як свого. Одним з критеріїв роботи системи може бути підхід, що полягає в наступному: система тим краще, чим менше значення FRR при однакових значеннях FAR.

Іноді використовується порівняльна характеристика EER (Equal Error Rate, рівний коефіцієнт помилок). Ця характеристика визначає точку, в якій величини FRR і FAR рівні. Справедливість цього твердження пояснимо нижче.

Для пояснення розглянутих статистичних характеристик біометричних систем, розглянемо дві щільності розподілу ймовірності, які характеризують шаблон користувача і шаблон хакера (див. рис. 1.1, [7]). Припустимо, що ці шаблони можуть бути задані у вигляді нормальних розподілів. Оскільки шаблон користувача отриманий в процесі навчання системи, то його основна характеристика (математичне значення) q1 має велике значення. Для користувача-хакера, який намагається проникнути в обчислювальну систему, шаблон отриманий оперативно в процесі застосування біометричної системи за призначенням має характеристику $q0$.

Величина qh – це певний поріг, який визначає величини, зазначених вище, помилок і бере участь у формуванні прийнятих рішень. В результаті оперативного аналізу отриманих біометричних даних користувача системи автентифікації і шаблонів, що зберігаються в базі зазначеної системи, приймається рішення про допуск поточного користувача, тобто віднести його до користувачів системи або користувачам (хакерам), які не допущені до ресурсів системи. З математичної точки дана задача відноситься до класу перевірки статистичних гіпотез, на базі якої син-

тезується вирішальне правило. У найпростішому випадку рішення приймається при розгляді двох взаємно виключаючих умовах:

- аналізовані біометричні характеристики належать користувачеві системи (умова A_1);

- аналізовані біометричні характеристики належать користувачу-хакеру, який не допущений до ресурсів системи (умова A_0).

У процесі автоматичного прийняття рішення в системі ці умови невідомі, а рішення приймаються на основі матеріалів реєстрації.

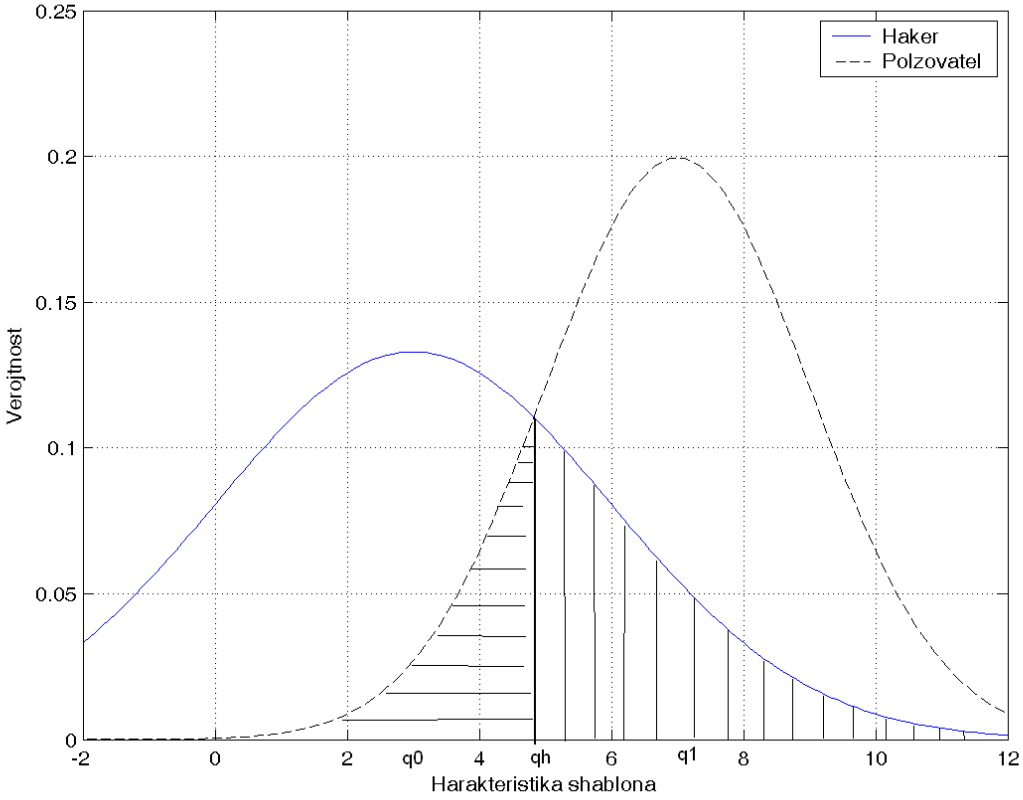

Рисунок 1.1 – До питання аналізу щільності розподілу ймовірностей аналізованих шаблонів

При синтезі вирішального правила за критерієм мінімуму середнього ризику, це правило зводиться до так званого вагового критерію [7]

$$D - l_0 F = \max, \qquad (1.1)$$

де D – умовна ймовірність: правильного допуску до ресурсів;

F – умовна ймовірність помилкового розпізнавання (допуск хакера до ресурсів системи, див. вертикальне штрихування на рис. 1.1).

Множник l_0, який називається ваговим, визначається співвідношенням вартостей помилкових рішень, а також величинами апріорних ймовірностей, розглянутих умов прийняття рішення A_1 і A_0.

Розглянутий критерій свідчить про наступне.

За сукупністю вимог – підвищення умовної ймовірності правильного допуску авторизованого користувача до ресурсів D, а також зниження умовної ймовірності помилкового розпізнавання F, слід прагнути до збільшення «зваженої» різниці .

$$D - l_0 F. \tag{1.2}$$

Звернемо увагу на наступне. У разі рівного розподілу апріорних ймовірностей розглядаються умов і вартості помилкових рішень величина множника l_0.=1. В цьому випадку синтезований критерій (1.1) перетворюється до наступного вигляду

$$D - F = \max, \tag{1.3}$$

або, що еквівалентно F = (1–D). У ряді робіт зазначений режим роботи біометричної системи вважається оптимальним. Виконаємо перевірку достовірності зазначеного твердження, проаналізуємо його більш детально.

Рівність $P(A_0) = P(A_1)$ свідчить про те, що апріорні ймовірності появи авторизованого користувача і хакера на вході біометричної системи автентифікації однакові і рівні 0.5.

На практиці $P(A_1) >> P(A_0)$. Разом з тим, не рівні і вартості помилкових рішень. Природно припустити, що вартість помилкової заборони доступу зареєстрованому користувачу значно менше вартості допуску до ресурсів хакера, яка характеризує умовну ймовірність помилкового розпізнавання.

У зазначених умовах $l_0>1$, а значить і (1–D) > F. Зауважимо, що помилки FRR мають менш важкі наслідки (вимагають повторної реєстрації користувача), на відміну від помилок FAR (помилкове розпізнавання), які призводять до допуску хакера до ресурсів і послуг телекомунікаційної системи.

Таким чином, порівняльна характеристика EER в практичних додатках не є оптимальною і доцільною для використання.

1.4 Огляд літератури за темою досліджень

Можливості ідентифікації особистості за голосовими даними захоплюють вельми широкий спектр завдань, що виділяє їх серед інших біометричних систем. Перш за все, голосова ідентифікація досить давно і широко використовується в різних системах розмежування доступу до фізичних об'єктів і інформаційних ресурсів [8,9]. Голосова ідентифікація є частиною окремого наукового напрямку – теорії мовотворення [10,11]. Перспективним представляється її нове застосування в системах, заснованих на телекомунікаційних каналах зв'язку. Як приклад, в мобільного зв'язку за допомогою голосу можна здійснювати управління послугами, причому впровадження голосової ідентифікації сприяє захисту від шахрайства [12,13].

Велика роль голосової ідентифікації обумовлена також рішенням такого важливого завдання, як захист мовної інформації. Ця ідентифікація застосовується при створенні нових технічних засобів і програмно апаратних пристроїв захисту мовної інформації [14], зокрема, від витоку акустичним, віброакустичним та іншим каналам.

Особливе місце ідентифікація особистості по голосу займає при розслідуванні злочинів, в тому числі в сфері комп'ютерної інформації [15-17], і при формуванні доказової бази такого розслідування. У цих випадках часто виникає необхідність проведення ідентифікації невідомого голосового запису. Проведення голосової ідентифікації – важлива практична задача при пошуку підозрюваного по запису голосу в телекомунікаційних каналах зв'язку. Визначення таких характеристик по голосу диктора, як стать, вік, національність, діалект, емоційне забарвлення мови, також важливі в галузі криміналістики і антитерористичних дій [18-20]. Результати ідентифікації важливі при проведенні фоноскопічних експертиз, при здійсненні експертного криміналістичного дослідження на основі теорії криміналістичної ідентифікації [21,22].

Істотний інтерес представляє розвиток методів голосової ідентифікації для суміжних напрямів, саме, для нових мовних технологій, пов'язаних з розпізнаван-

ням усного мовлення [23], управлінням комп'ютерними системами за допомогою голосових команд [24].

Окрему важливу складову голосової ідентифікації особистості представляє формування баз голосових даних. Роль таких баз даних істотно зросла у зв'язку з розвитком нових технічних засобів обробки і зберігання голосової інформації [25]. Бази голосових даних необхідні, зокрема, при апробації нових методів оцінки захищеності мовної інформації, а також при перевірці надійності технічних пристроїв її захисту.

Потрібно окремо відзначити, що при виконанні особливо важливих робіт необхідно використовувати поєднання різних технологій і методів, що забезпечує найбільш надійну ідентифікацію та автентифікацію. Іншими словами, в цих випадках доцільно поєднувати біометричні голосові методи ідентифікації зі спеціальними фізичними пристроями доступу з пам'яттю (token) і з мікропроцесорними картами (смарт-карти) [26].

В роботах [27-31] розглянути окремі питання підвищення якості систем голосової автентифікації за рахунок використання фазових даних голосового сигналу користувача.

2 АНАЛІЗ БІОМЕТРИЧНИХ МЕТОДІВ ІДЕНТИФІКАЦІЇ І АВТЕНТИФІКАЦІЇ КОРИСТУВАЧІВ ІНФОКОМУНІКАЦІЙНИХ СИСТЕМ

2.1 Загальна характеристика завдання оцінки достовірності методів ідентифікації і автентифікації користувачів

Одним методів ідентифікації і автентифікації особистості, що найбільш інтенсивна розвивається в останні роки, є ідентифікація по біометричних характеристиках. Біометрія привертає розробників тим, що користувачеві не треба запам'ятовувати або записувати ідентифікаційну і автентифікаційну інформацію. За останні два десятиліття розроблено кілька десятків методів ідентифікації. У маркетингових матеріалах виробники наводять дуже привабливі дані щодо точності ідентифікації, проте на практиці ці дані виявляються завищеними.

В даному розділі розглядається достовірність деяких найбільш вживаних методів ідентифікації, наводиться огляд і короткий аналіз найбільш використовуваних способів біометричної ідентифікації. Проблема достовірності ідентифікації особистості в сучасних умовах є дуже актуальною. Під достовірністю ідентифікації (ДІ) будемо розуміти повноту і точність ідентифікаційної інформації про користувача. ДІ обернено пропорційна ймовірності виникнення помилок при виникненні в ІС ідентифікаційної та автентифікаційної інформації, а також під час їх зберігання, передачі і обробки в ІС. Іншими словами, ДІ визначається надійністю і безпомилковістю процесу ідентифікації особистості.

Питанням достовірності ідентифікації сторін взаємодії поки не приділялася належна увага, можливо через те, що тільки зараз до безлічі закритих раніше корпоративних систем почали пред'являти вимоги Web-доступу та обміну з іншими ІС. Число зовнішніх користувачів багатьох ІС інтенсивна росте, наприклад, ІС держави вже містять десятки мільйонів зовнішніх користувачів, і потенціал зростання ще не вичерпано.

Основною проблемою при цьому є визначення технологій, механізмів і засобів ідентифікації (ТМЗІ), що дозволяють з певним ступенем упевненості довіряти результатам ідентифікації як для невеликих ІС, так і для ідентифікаційні системи загального користування. Завдання ускладняється тим, що вітчизняна нор-

мативна база не містить вимог до безпеки, надійності і якості (БНЯ) виконання процесів ідентифікації і автентифікації.

Питання вибору тих чи інших ТМЗІ віддано на відкуп власникам інформаційних систем. Потрібно знайти критерії вибору, які відповідають очікуванням БНЯ і охоплюють всі існуючі і перспективні рішення на ринку. В якості одного з можливих підходів до вирішення завдання розглядається використання для ідентифікації учасників ВЕВ сертифікатів ключа перевірки електронного підпису (СКПЕП), прямим призначенням якого є в числі інших і функція ідентифікації власника. Однак дослідження достовірності ідентифікації *D* особистості власника СКПЕП, наведене в ряді робіт, показало занадто низькі значення достовірності, що розраховуються за формулою:

$$D = 1 - \prod_{i=1}^{n} p_i \qquad (2.1)$$

де p_i - ймовірність відсутності помилки ідентифікації при пред'явленні *i*-го ідентифікатора.

Відомо, що для ІС загального користування, де порядок числа користувачів становить, як правило, 10^5 - 10^7, необхідно мати надійні механізми ідентифікації, що забезпечують точність ідентифікації порядку 10^{-8}. При збільшенні числа користувачів порядок точності оцінюється як 10^{-n+1}, де *n* – кількість користувачів системи.

2.2 Коротка характеристика фізіологічних біометричних систем автентифікації

Найбільш поширена і має великий термін використання – біометрична система на основі аналізу відбитків пальців (дактилоскопія, див. рис. 2.1). Алгоритм її використання наступний. Оптичний сканер, розташований, наприклад, на клавіатурі, флешці, смартфоні, або на будь-якому іншому девайсі. Сканер зчитує відбиток пальця і перетворює його в картинку. Далі відбувається пост-обробка отриманого «зображення» і порівняння з збереженим зображенням.

Переваги методу:

- широке поширення методу;
- низька вартість пристроїв;

- проста процедура ідентифікації;

- в стані стресу або під впливом інших причин відбитки пальців не змінюються, що підвищує надійність методу.

Рисунок 2.1 – Використання автентифікації за відбитками пальців

Статистичні характеристики методу: FAR від 0.0001% до 0.1%; FRR від 0.3% до 0.9%.

Основні виробники: SecBayometricInc., DigitalPersonaInc., BioLink, Сонда, СмартЛок.

Недоліки методу:

- зниження точності ідентифікації за рахунок пошкодження пальця, забрудненості або вологості;

- можливість підробки відбитків.

Таким чином, до переваг ідентифікації за відбитками пальців можна віднести те, що відбитки не змінюються з віком, мають високу надійність, вартість системи ідентифікації відносно низька. До недоліків же можна віднести те, що папілярний візерунок відбитків пальців дуже легко пошкоджується, наприклад, механічних, хімічних впливами, також недоліком є схильність пропуску за підробленими зображень відбитків.

Системи на основі аналізу райдужної оболонки ока (див. рис. 2.2). Технологія включає пристрій захоплення зображення, його первинну обробку і передачу даних обчислювачеві. Далі обчислювач порівнює зображення з зображеннями в базі даних і передає команду про допуск виконавчому пристрою.

Рисунок 2.2 – Автентифікація за райдужною оболонкою ока

Переваги методу:

- відсутність необхідності фізичного контакту з пристроєм;
- відсутність часових змін райдужної оболонки ока;
- низька ймовірність існування двох абсолютно ідентичних рисунків на райдужній оболонці.

Статистичні характеристики методу: FAR від 0.0001% до 0.1%; FRR від 0.065% до 0.5%.

Основні виробники: LG Electronics, Panasonic, OKI, Iris Access 2200.

Недоліки методу:

- висока вартість системи;
- низька доступність готових рішень;
- повільне поширення технології через патентні обмеження;
- можливість підробки.

Таким чином, біометрична технологія, заснована на райдужці оболонки очей, має високу надійність, з часом райдужні оболонки очей практично не змінюються і стійкість до підробки може забезпечуватися різними методами захисту. Істотним недоліком є висока вартість.

Автентифікація по геометрії руки (див. рис. 2.3). Зазначена автентифікація використовує до 90 характеристик: вигини пальців, їх товщину і довжину, відстань між суглобами і структуру кістки і ін. Якщо кожний вимір вкладається в певні допустимі рамки зареєстрованого еталонного набору даних, то результат автентифікації буде для користувача позитивним.

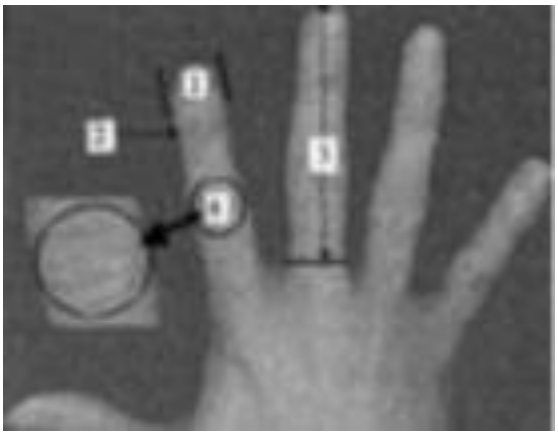

Рисунок 2.3 – Автентифікація по геометрії руки

Переваги методу:

- висока точність визначення: ймовірність існування двох кистей рук з однаковою геометрією надзвичайно мала;

- широке застосування;

- зручність: система не пред'являє підвищених вимог до вологості, температури, кольору.

Статистичні характеристики методу: немає достовірних даних.

Основні виробники: Recognition Systems, BioMet Partners, Escape.

Недоліки методу:

- облік безлічі параметрів;

- вплив пошкоджень або забруднення тканин, а також віку;

- вузьке застосування методу;

- великі габарити пристрою;

- вимоги по розташуванню долоні.

Технологія, що базується на розпізнаванні особи користувача (2D і 3D розпізнавання). За допомогою оцифрування зображення в кадрі вибирається обличчя людини. Фотографія і цифровий опис особи заносяться в базу даних, з якої згодом порівнюється особа, що розпізнається.

Переваги методу:

- широке застосування;

- не дуже дороге обладнання;

- не потрібен фізичний контакт з пристроєм;

- несприйнятливість до змін зовнішності;

- високий рівень захисту;

Рисунок 2.4 – Технологія на розпізнаванні особи користувача

- можливість проведення прихованої процедури.

Статистичні характеристики методу: FAR від 0.0001% до 0.1%; FRR від 9% до 2.5%.

Основні виробники: Geometrix, Inc. Artec Group, Cognitec Systems GmbH, Bioscrupt.

Недоліки методу:

- вплив віку;

- наявність вимог до освітлення, до виразу обличчя та його положення;

- вимоги до фронтального зображення обличчя, з досить невеликими відхиленнями.

Ідентифікацію по обличчю поділяють на два види: 2-мірну і 3-мірну. Перевагами 2-мірної є те, що можлива ідентифікація на деякій відстані і вартість системи низька, а до недоліків можна віднести те, що особа змінюється з віком, низька надійність, чутливість до зовнішніх факторів. Для 3-мірної ідентифікації істотно нижче чутливість до зовнішніх факторів і високий рівень надійності. Недоліком 3-мірної ідентифікації є висока вартість обладнання.

На жаль, розглянуті біометричні системи не забезпечують сучасним вимогам до систем автентифікації. Тому все частіше дослідники звертають на поведінкові біометричні системи автентифікації.

2.3 Аналіз голосових систем автентифікації користувачів

З швидкими темпами зростання інформатизації суспільства зростає потреба в захисті конфіденційної інформації. Однією з поширених заходів захисту інформації є розмежування доступу. Розмежування доступу включає в себе автентифікацію – процедуру перевірки автентичності, і ідентифікацію – привласнення іден-

тифікатора і (або) порівняння пред'явленого ідентифікатора з переліком ідентифікаторів [31]. Існує більш просте завдання – верифікація, що має на увазі порівняння запропонованих ознак, заявлених як відповідні відомому об'єкту, з відомими ознаками того ж об'єкта.

Серед безлічі різних методів розмежування доступу особливе місце займають біометричні технології. Зростаючий інтерес до біометричних технологій пов'язаний в основному із зручністю їх застосування. Часто під біометричними технологіями, в контексті розмежування доступу, розуміють методи біометричної ідентифікації. Завдання ідентифікації має на увазі порівняння ознак одного об'єкта з ознаками інших відомих об'єктів.

Більшість методів біометричної ідентифікації можна розділити на дві групи: фізіологічні (або статичні) і поведінкові (або динамічні, що враховують підсвідомі дії). Основні фізіологічні методи ідентифікації були розглянуті вище.

Поведінкові методи ідентифікації, або поведінкова біометрія (behavioral biometrics): по голосу, по підпису (почерку), по динаміці роботи з клавіатурою або мишею, по ході, крім того, досліджується також можливість ідентифікації по електромагнітних хвилях мозку.

Для визначення ефективності технологій оцінюють такі критерії, як надійність, стійкість до підробки, стійкість до навколишнього середовища (завадостійкість), стабільність ознаки від часу, швидкість, вартість і зручність застосування. Переваги і недоліки деяких з фізіологічних біометричних технологій розглянуті вище.

Всі оцінки критеріїв ефективності і надійності тієї чи іншої технології великою мірою залежать від використовуваних баз даних. У лабораторних умовах створені бази даних біометричних ознак при ідентифікації можуть забезпечувати високу надійність, але в реальних умовах, де впливають різні зовнішні перешкоди, надійність може виявитися значно нижчою від заявленої. Незважаючи на різну ефективність цих технологій, кожна біометрична технологія може бути безумовно краще за інших для певного специфічного завдання.

Голосова ідентифікація володіє такими перевагами, як зручність застосування і невисока вартість. Недоліком такої ідентифікації є низька надійність. Одним з перспективних шляхів підвищення надійності голосової ідентифікація є залучення інших характеристик (наприклад, фазу) сигналу, який обробляється, що активно використовується при ідентифікації за підписом [32]. З іншого боку, є об-

ласті застосування, в яких голосова ідентифікація є найбільш зручною, наприклад, віддалений доступ за допомогою телекомунікаційних каналів зв'язку.

У сучасних системах голосової ідентифікації для підвищення надійності застосовують текстозалежну ідентифікацію, наприклад, проголошення парольної фрази, яка кожного разу генерується випадковим чином. Використання індивідуальних ознак і збіг згенерованої і розпізнаної парольної фрази підвищує надійність. Текстонезалежна ідентифікація має на увазі використання тільки індивідуальних ознак.

Важливою характеристикою системи голосової ідентифікації є швидкість (швидкодія) визначення особистості. Швидкодія особливо важлива для додатків, що обробляють великі бази голосових даних і працюють в реальному режимі часу. Підвищення швидкодії може бути досягнуто за рахунок використання нових швидких алгоритмів обробки даних. Таким чином, голосова ідентифікація особистості, незважаючи на зазначені у цій роботі недоліки, в певних умовах має істотні переваги, які необхідно розвивати.

Підвищення надійності голосової ідентифікації важливо не тільки для такого напрямку, як розмежування доступу до фізичних і інформаційних об'єктів, наприклад, доступу до операційної системи персонального комп'ютера або віддаленого доступу до телекомунікаційних каналів зв'язку по аналізу голосових даних. Певний інтерес є і для суміжних напрямів мовних технологій: розпізнавання усного мовлення, управління голосовими командами і інші. На сьогоднішній день широкого поширення набув електронно-цифровий підпис для захисту конфіденційних документів у вигляді захищеного електронного пристрою (token), в зв'язку з цим перспективним напрямком є розробка захисту конфіденційних документів на основі мовного підпису [33]. Крім того, практичні застосування таких досліджень корисні для правоохоронних органів, наприклад, пошук підозрюваного по голосу в телекомунікаційних голосових каналах зв'язку. Пошук диктора за голосом може успішно застосовуватися для виявлення диктора (виявлення ознак диктора в аудіо-потоці) [34] і протоколювання дикторів (визначення ділянок усного мовлення диктору окремо в аудіо-потоці) [35].

Подібно відбитками пальців у криміналістичному вченні – дактилоскопії, в фоноскопії використовується свій криміналістичний об'єкт – фонограма [36,37]. У зв'язку з цим для фоноскопічної експертизи використовують свої специфічні методи і технічні засоби. Існують два основних напрямки криміналістичного дослі-

дження фонограм: перше – дослідження технічних характеристик звукозапису, друге – дослідження безпосередньо самого усного мовлення людини. В основі експертного криміналістичного дослідження лежить теорія криміналістичної ідентифікації та діагностики. Метою ідентифікації та діагностики є встановлення об'єкта. Ідентифікація має на увазі зіставлення звукових ознак кількох конкретних об'єктів, і результатом цього зіставлення є схожість або відмінність об'єктів. При діагностиці використовують відомі ознаки класу або групи об'єктів і порівнюють їх з ознаками конкретного об'єкта, в результаті чого визначається приналежність його до даного класу або групи. Для досягнення мети експерти вирішують різні види завдань. Це дослідження джерела звуку, середовища поширення (закрите приміщення, відкритий простір), а також приймача (запис, зберігання, відтворення). Основним завданням ідентифікаційних досліджень фонограм є ідентифікація особистості по фонограмі усного мовлення.

Існує багато інструментів дослідження мовних сигналів, заснованих на математичній обробці, серед яких основним є спектральний аналіз. Використання часової залежності і спектрального аналізу дає додаткову інформацію і називається спектрально-часовим аналізом усного мовлення.

Інший вид аналізу, лінгвістичний аналіз усного мовлення, доповнює спектральний додатковою більш високорівневою інформацією. До сих пір в криміналістиці використовується зіставлення спектрограм мовного сигналу шляхом порівняння чисто зовнішньої схожості. Найбільш простий спосіб дослідження усного мовлення з метою ідентифікації – дослідження безпосередньо на слух або по фонограмі. Ознак, які використовуються при дослідженнях усного мовлення, існує велика кількість, але чіткої систематизації ознак немає.

Для прийняття рішення про подібність або відмінність, або належність чи неналежність до відомого класу об'єктів у завданні ідентифікації особистості застосовують статистичні методи. Статистичні методи зазвичай надають відповідь з певною ймовірністю, яка, в загальному випадку, не може бути доказом подібності або приналежності. Обчислювальні потужності сучасних комп'ютерів дозволяють обробляти велику кількість даних, що сприяє створенню більш надійних автоматизованих систем ідентифікації особистості. З іншого боку, необхідність побудови автоматизованих систем ідентифікації обумовлена впливом суб'єктивного фактору при криміналістичній експертизі.

Важливою характеристикою системи голосової ідентифікації є стійкість. Під перешкодами розуміються спотворення, шуми, імпульсні перешкоди тощо. Сучасні методи класифікації, що використовуються в системах голосової ідентифікації, дуже чутливі до шуму, що призводить до зниження надійності при впливі шуму.

Перешкоди можуть надходити від різних джерел, наприклад, від двигуна автомобіля, від звуку вітру, від голосів інших дикторів, від реверберації звуку. В роботі [38] джерела спотворень поділяють на два види: адитивний шум і спотворення каналу. Адитивний шум може бути стаціонарним і нестаціонарним. Стаціонарний шум має спектральну щільність потужності, що не змінюється з плином часу, наприклад, шум двигуна автомобіля. Нестаціонарний шум має статистичні властивості, які змінюються з плином часу, наприклад, голоси інших дикторів. Прикладом спотворення каналу може бути реверберація звуку, тобто звук багато разів відбивається від предметів в приміщенні. Навколишній фон в багатьох випадках є перешкодою і може мати значний вплив на голосову ідентифікацію. Так, щоб оцінити, як сильно впливає навколишній фон, в роботі [39] використовували записи різних реальних зовнішніх умов: в машині, в поїзді, в ресторані, в бесіді. У телефонному каналі перешкодами можуть бути клацання, перевантаження, музичні сигнали, гудки (тональні сигнали). Наприклад, в роботі [40] представлений алгоритм детектування музичних фрагментів, який може використовуватися для боротьби з однією з перешкод телефонного каналу.

У статті [41] представляється новий метод для розпізнавання диктора з дуже обмеженими мовними даними навчання і при наявності фонового шуму. Пропонується використовувати одну з форм косинусної подібності в якості міри відстані між векторами мовних ознак. Кожен мовний кадр моделюється за допомогою піддіапазону ознак, в цих рамках підготовки вводяться навчання з багатьма станами (multicondition training) і оптимальний вибір ознаки, в результаті чого система здатна виконувати розпізнавання диктора в присутності реального шуму, що змінюється в часі, який невідомий під час навчання. Експерименти по ідентифікації диктора проводилися з використанням SPIDRE бази даних [42]. Проводилися також експерименти для тестування нової моделі ідентифікації диктора, з огляду на обмежені дані навчання і з різними рівнями і типами реалістичних фонових шумів. Отримані результати показали надійність нової системи.

Робота [43] присвячена проблемі верифікації диктора, де адитивний шум присутній при реєстрації і тестуванні висловлювання. При реєстрації гучних сигналів мови спостерігалося зниження вірогідності стандартної системи верифікації диктора. У цій роботі використовувався мовний корпус з зашумленими умовами, на основі даних NIST SRE 2008 [44] і NIST SRE 2010 [45], побудований з використанням відкритого вихідного коду і вільно доступних зразків шуму.

В роботі [46] досліджується проблема ідентифікації і верифікації дикторів в зашумлених умовах, припускаючи, що мовні сигнали були схильні до шуму навколишнього середовища, характеристики якого невідомі. Це дослідження частково мотивоване можливим застосуванням технологій розпізнавання диктора на мобільних пристроях або в мережі Інтернет. У цій статті описується метод, який поєднує в собі навчання моделі з багатьма станами (multicondition model training) і теорії неповних даних (missing-feature theory) до моделі шуму з невідомими спектрально-часовими характеристиками. Навчання з багатьма станами проводиться з використанням моделювання зашумлених даних з обмеженою зміною шуму, забезпечуючи певну компенсацію, а теорія неповних даних застосовується для уточнення компенсації, ігноруючи зміну зовнішнього шуму, який забезпечує умови навчання, тим самим знижуючи помилки навчання і тестування. Новий алгоритм був протестований за допомогою двох баз даних: модельованої і реалістичних зашумлених мовних даних. Перша база даних була реконструкцію бази даних TIMIT [47] при перезапису в присутності різних типів шуму, використовується для перевірки моделі при ідентифікації диктора з акцентом на різновиди шуму. Друга база даних була базою даних, зібраних портативним пристроєм в реалістичних зашумлених умовах, які використовуються для подальшої перевірки моделі при реальних верифікаціях диктора. Запропонована в статті модель у порівнянні з вихідними системами володіє більш низьким процентом помилки.

В роботі [48] досліджується використання мікрофонних масивів в системах розпізнавання диктора «вільні руки» (handsfree). Гучний зв'язок краще в багатьох потенційних застосуваннях розпізнавання диктора, проте розпізнавання з одним віддаленим мікрофоном є проблематичним в реальних умовах шуму. Можливим вирішенням цієї проблеми є використання мікрофонних масивів, які мають потенціал для поліпшення якості сигналу, що базується виключно на знанні напрямку його приходу. Використання мікрофонного масиву для підвищення надійності систем розпізнавання мови була добре вивчена останнім часом, однак мало дослі-

джень було проведено в області розпізнавання диктора. У статті обговорюється застосування мікрофонних масивів для додатків розпізнавання диктора, що представляє собою експериментальну оцінку застосування технології «вільні руки» при верифікації диктора в зашумлених умовах.

2.4 Коротка характеристика процедур цифрової обробки голосового сигналу

У задачі голосової ідентифікації застосовують різні математичні, алгоритмічні, технічні методи, починаючи з етапу записи голосу і закінчуючи етапом класифікації. Практично кожна система ідентифікації містить чотири основні етапи: отримання сигналу, попередня обробка сигналу, отримання ознак і класифікація ознак. Розглянемо ці етапи для нашої задачі.

Етап отримання сигналу. Метод отримання або запису голосового сигналу, в більшості випадків, є записом сигналу за допомогою мікрофона і надання сигналу в цифровому вигляді за допомогою аналого-цифрового перетворювача. У якості аналого-цифрового перетворювача зазвичай використовують звукову карту персонального комп'ютера або цифровий диктофон. Цифрові дані кодуються через сигнал PCM і поміщаються в формат файлу-контейнера (Waveform Audio File Format) для зберігання записи оцифрованого аудіопотоку. Параметри звукового запису зазвичай такі: бітність відліків – до 16 біт, частота дискретизації – до 64000 Гц. Так як сучасні цифрові мобільні пристрої зазвичай мають вбудований мікрофон і продуктивні апаратні засоби, то створення система автентифікації по голосу з залученням більш витратних за обчисленнями методів цілком вирішуване завдання для мобільних платформ. Проте, забезпечити мінімальні обчислювальні витрати при збереженні точності, завадостійкості до різних видів перешкод і достатню надійність при поширених апаратних засобах все ж необхідно.

Етап попередньої обробки. Отримані цифрові сигнали, як і аналогові, містять в собі деяку частку спотворень і перешкод. Під спотвореннями розуміються спотворення мовотворчого тракту (наприклад, хвороба горла) і каналу, що передає мову (наприклад, спотворення телефонного каналу).

У роботах [49-51] описуються деякі способи боротьби з спотвореннями такого каналу. Перешкодами виступають не тільки зовнішні шуми, а й мова сторонніх людей. Зовнішні шуми зазвичай пригнічують на етапі попередньої обробки за

допомогою різних фільтрів: смугових, медіанний, адаптивних, на основі вейвлет перетворень та інших.

Наприклад, в роботі [52] для боротьби з шумом використовують метод спектрального віднімання. Також для зменшення впливу негативних факторів (спотворень і перешкод) застосовують різні види нормування [53,54].

Етап отримання ознак. Отримання ознак зазвичай відбувається за допомогою Фур'є-перетворення, вейвлет-перетворень, лінійного передбачення та інших. Коефіцієнти перетворень виступають в якості ознак. В даний час точно не визначені голосові ознаки, за якими можна однозначно ідентифікувати особу людини.

Вибір ознак впливає також на надійність ідентифікації. Існують, методи, які описують інтегральні характеристики людського голосу і служать для вилучення тонів, динаміки мови, просодических характеристик. Такими методами є перетворення Фур'є (амплітудно-частотний розподіл), кепстральні перетворення (амплітудно-часовий розподіл), перетворення лінійного передбачення (амплітудно-частотний розподіл). Існують також формантні методи і методи виділення фонем. Використання вейвлет коефіцієнтів [55,56] не забезпечує значної переваги, до того ж урахування частотної і часової складових вимагає додаткових обчислювальних витрат. Для ефективного використання частотної і часової складових голосових сигналів у завданні ідентифікації особистості по голосу необхідно проводити додаткові дослідження. Переваги тих чи інших голосових ознак виявляються в конкретних випадках при певних умовах і на певній мовній базі даних.

В роботі [57] порівнювалися кілька векторів ознак, які будувалися на основі коефіцієнтів: швидкого перетворення Фур'є, лінійного та кепстральних перетворень. Крім того, до коефіцієнтів додатково застосовувалося мел-перетворення. Для різних методів класифікації ефективніше можуть виявитися набори коефіцієнтів різних перетворень.

З перерахованих вище голосових ознак, в різних комбінаціях, формується вектор ознак у вигляді послідовності чисел. Формування векторів ознак для одного або декількох дикторів входить в етап калібрування або навчання системи ідентифікації диктора. На виході етапу калібрування будується загальна модель одного або декількох дикторів, яка співвідноситься з вихідними мовними даними. На етапі тестування (перевірки) вектор ознак витягується зі звукової хвилі і порівнюється з побудованої моделлю.

Кожен вектор ознак повинен формуватися оптимально, з огляду на обчислювальні витрати, з одного боку, і інформативність ознак з іншого. Так, при високій інформативності вектору ознак, а саме, великій кількості ознак, обчислювальні витрати зростають, і навпаки, при малій кількості ознак обчислювальні витрати зменшуються. Однак при збільшенні кількості ознак у векторі інформативність не завжди збільшується. Деякі оцінки інформативності ознак можна отримати на етапі тестування за допомогою помилки класифікації (ідентифікації), тобто мінімум помилки буде відповідати більшої інформативності. Тому часто проводять окремі дослідження щодо формування оптимального вектору ознак для кожного диктора, тобто досліджують залежність поєднань різних наборів ознак з вихідних даних (коефіцієнтів будь-яких перетворень) від помилки класифікації. Сполучення різних наборів ознак може досліджуватися як простим перебором, так і з використанням методів оптимізації (наприклад, генетичних алгоритмів). З іншого боку, самі ознаки змінюються в деякому діапазоні, що викликано або впливом зовнішніх шумів і перешкод, або впливом внутрішніх викривлень голосового апарату людини. Тому відмінність тестових і калібрувальних умов класифікації може привести до формування різних векторів ознак для одного і того ж диктора, а отже – до зниження достовірності методів класифікації.

Етап класифікації ознак. В цей етап входить застосування математичних методів класифікації, за допомогою яких здійснюється прийняття рішення, а також розрахунок помилок класифікації.

У задачі голосової ідентифікації використовують ті ж методи класифікації, що і в області розпізнавання образів, а саме, методи статистичного моделювання, які будують певні моделі векторів акустичних ознак. Найбільш поширеними з них є моделі гаусових сумішей і приховані марковські моделі. Однак інші моделі, наприклад, багатошарові персептрони або машина опорних векторів, також успішно використовуються в даній задачі. Крім того, останнім часом спостерігається тенденція використання комбінацій декількох моделей.

Моделі гаусових сумішей часто використовують для текстонезалежної верифікації дикторів [58,59] при оцінці щільності ймовірностей мінливості мовних даних. При невисоких обчислювальних витратах і малій чутливості до часової мінливості мови, моделі гаусових сумішей добре проявили себе в умовах близьких до умов тихого оточення, застосування високоякісних мікрофонів та ін. В реаль-

них умовах присутності фонового шуму, використання різних мікрофонів і каналів передачі ефективність моделей гаусових сумішей погіршується.

Через обмеженість даних, доступних для тренування моделі диктора, затребувані технології адаптації: ЕМ-алгоритм (Expectation- Maximization) [60], максимум апостеріорної ймовірності (Maximum a Posteriori Probability) або максимум правдоподібності лінійної регресії (Maximum Likelihood Linear Regression) [61].

Приховані марковські моделі є статистичними моделями, в яких система моделюється як марковський процес з невідомими параметрами [62]. Метою є визначення найбільш ймовірного стану послідовності тестового набору щодо попередньо тренувальних моделей. Для додатків розпізнавання диктора кожне стан прихованої марковської моделі може представлятися різними елементами мови [63]. Часова інформація кодується переходом з одного стану в інший щодо дозволених переходів. У цьому випадку метод ідентифікації на основі прихованих марковських моделей полягає у визначенні для кожного диктора найкращого положення між послідовністю тестового мовного вектору і прихованою марковською моделлю, пов'язаною з певним словом або фразою.

Сучасні текстонезалежні системи верифікації диктора, що використовують моделі гаусових сумішей, не враховують часове впорядкування векторів ознак. Лінгвістична і часова структура мовного сигналу у вигляді рахунків і всіх звуків, які використовуються в поданні, не дає унікальну модель. Приховані марковські моделі мають певні переваги. Часові і лінгвістичні знання можуть бути зареєстровані за допомогою прихованих марковських моделей. У задачі текстозалежного розпізнавання диктора використовуються апріорні знання змісту тексту при цьому приховані марковські моделі [64] точніше, ніж моделі гаусових сумішей. В роботі [65] показано деякі переваги застосування комбінації розпізнавання мови на складової прихованої марковської моделі і дикторів-орієнтованого розпізнавання на основі моделі гаусових сумішей.

Багатошарові персептрони є різновидом нейронних мереж, що підлягають навчанню [66]. Застосування нейромережних методів в завданні верифікації диктора показано в роботах [67,68]. Для систем верифікації диктора багатошарові персептрони можуть бути двійковими класифікаторами, які виділяють класи «свого» і «чужого». Багатошаровий персептрон зазвичай складається з декількох шарів, кожен з яких має кілька вершин. Кожна вершина обчислюється як сума лінійних ваг всіх вхідних з'єднань, де ваги суми є підлаштовуємося параметрами. Не-

лінійна функція переходу застосовується для результату обчислення виходу вершини. Ваги мережі оцінюються через градієнт нахилу, заснований на алгоритмі зворотного поширення. Багатошаровий персептрон буде класифікувати доступ «свого» і «чужого», вважаючи кожен кадр тестового висловлювання.

Недоліками багатошарового персептрона є відносна складність вибору оптимальної конфігурації, а також необхідність великої кількості даних для етапів навчання і перевірки (крос-перевірки).

Машина опорних векторів є двійковим класифікатором [69]. Основним принципом є проекція нелінійно роздільних багатовимірних даних у гіперпростір, де вони можуть бути лінійно нероздільні. З огляду на те, що набір векторів ознак належить до двох класів, які поділяються гіперплощиною, машина опорних векторів буде намагатися знайти гіперплощину з максимальним краєм.

Іншими словами, відстань між найближчими поміченими векторами до гіперплощини буде максимальною. Ця гіперплощина може бути в подальшому бути використана (на етапі перевірки) для визначення, до якого класу належить невідомий вектор ознак. В останні роки машина опорних векторів вважається одним з ефективних методів дискримінації [70]. У задачі верифікації диктора машина опорних векторів може використовуватися окремо або в комбінації з іншими методами класифікації. Наприклад, в роботі [71] використовується комбінування методів на основі моделі гаусових сумішей і машини опорних векторів. У ній робиться спроба використовувати компенсації мінливості диктора і каналу, а саме, за допомогою моделі гаусових сумішей сформувати певний середній вектор, який далі обробляється машиною опорних векторів. Комбінування методів призводить до підвищення точності класифікації.

Суттєвою проблемою для систем голосової ідентифікації, заснованих на перерахованих вище методах, є сильний вплив зовнішнього навколишнього шуму на вихідні голосові записи, з яких виділяються інформативні ознаки. Обумовлене шумом спотворення цих ознак викликає високий рівень помилок ідентифікації.

2.5 Коротка характеристика методів формування ознак голосового сигналу

Біометричні технології є перспективним напрямком в області інформаційної безпеки. Голосова біометрія на сьогодні є широко поширеною, і роботи над підвищенням якості голосових систем не втрачають своєї актуальності. Вибір методу

вилучення речових ознак – один з ключових етапів проектування голосових автоматичних систем.

Нижче розглядаються акустичні параметри, обумовлені фізіологічними властивостями мовного тракту людини: частота основного тону, огинаюча спектра, форманти і антіформанти. Основна увага зосередимо на аналізі методів вилучення ознак.

Велику частину складають різні варіанти кепстрального аналізу, оскільки саме вони найбільш часто зустрічаються в сучасних розробках, як у вигляді використання популярних мел-частотних кепстральних коефіцієнтів, так і в нових модифікаціях.

Параметризація мовних характеристик входить в розпізнавання мови, емоцій, мови, гендеру. Розглянемо основні підходи вилучення акустичних ознак мови з метою автентифікації користувача (розпізнавання диктора), матеріал може бути корисний і в задачах обробки мовних сигналів.

Біометричні технології активно впроваджуються в життя суспільства. Про це свідчить існуючий і прогнозоване зростання ринку біометрії як на світовому, так і на вітчизняному рівні. Розпізнавання по голосу завдяки широкій доступності обладнання, можливості дистанційної ідентифікації, простому процесу навчання і використання для споживача є популярною біометрикою, застосовуваної в області інформаційної безпеки.

Підвищення якості розпізнавання особистості в різних умовах і протидія спуфінговим атакам залишаються актуальними проблемами мовної обробки сигналів. Значний компонент автоматичних систем голосової біометрії – отримання інформативних параметрів мовного сигналу.

Спочатку зупинимося на мовних характеристиках. Індивідуальність голосу забезпечується поєднанням поведінкових і фізіологічних ознак. До поведінкових відносять семантику, дикцію, вимову, ритм, інтонації та ін. Вони обумовлені соціальними факторами і можуть бути досить мінливими залежно від ситуації.

Більш надійними є анатомічні особливості мовного тракту, тому для роботи автоматичного розпізнавання найбільш адаптовані алгоритми вимірювання акустичних характеристик. Акустична теорія мови розглядає мовну хвилю як результат роботи джерела звуку і фільтрів.

Детальний виклад про фізіологічні процеси мовотворення і моделі мовного тракту можна знайти в книгах [1, 10]. В цій роботі коротко наведені тільки ті па-

раметри, які беруть участь в автоматичному розпізнаванні дикторів. Характерні риси голосу конкретної людини в цифровій обробці сигналів отримують через спектральний аналіз мовної хвилі.

Частота першої гармоніки спектра є частотою основного тону (основний частотою голосу). Частота основного тону F_0 – зворотна величина тривалості T_0 одного циклу роботи голосових зв'язок:

$$F_0 = \frac{1}{T_0}. \tag{2.2}$$

Основна частота визначає висоту голосу – відчуття, пов'язане з впливом тони на слухову систему людини. Індивідуальність даного параметра пояснюється тим, що тривалість T_0 залежить від маси і пружності голосових зв'язок, а також від перепаду тиску над і під зв'язками. Тому стать і вік диктора впливають на значення основної частоти. Кожна людина має свій діапазон змін частоти основного тону. Як правило, для дорослого він становить від півтора до двох октав. У задачі розпізнавання особистості по голосу необхідно визначати базову основну частоту, тобто звичний і зручний для ідентифікованого людини режим роботи голосових зв'язок.

Важливу роль у визначенні індивідуальних голосових особливостей грають і інші гармоніки, звані обертонами (формантного частотами). Частота конкретного обертони виражається як

$$F_n = n \cdot F_0, \tag{2.3}$$

де n – порядковий номер обертони в спектрі. У сукупності огинаюча спектра (лінія, що з'єднує вершини амплітуд обертонів) відображає регістр, тембр, основну частоту і гучність мови. Її форма визначається розмірами і конфігураціями порожнин рота, гортані і носа, взаємним розташуванням зубів, мови і губ. Спектральна огинаюча показує відносний внесок гармонік в загальну енергію мовного сигналу. При аналізі мови можуть враховуватися нахил і швидкість спаду спектральної огинаючої.

Акустичні резонанси в голосовому тракті створюють піки в огинаючій спектра звуку. Такі піки називаються формантами. За частотам формант можна аналі-

зувати положення артикуляційних органів, що активно використовується в фонетичному аналізі сказаного.

Однак частоти формант залежать не тільки від відтворюваних фонем, але і від мовця: графіки спектра відтворення одного і того ж звуку від двох дикторів мають відмінності. Варіативність формантних частот для різних фонем і контексту досить широкий, але для певної людини у визначеному фонематичному контексті відмінності між звуками відповідають свої відмінності в спектральній картині.

Крім частоти, форманти характеризуються шириною. Ширина (смуга) форманти обмежує діапазон частот по обидві сторони від частоти формант, посилення яких становить не менше 70,7% від максимального резонансного посилення на частоті форманти і служить мірою частотної вибірковості мовного тракту при резонансі.

Іноді вимірювання формант доповнюється перебуванням антіформант – глибоких мінімумів спектру сигналу, що виникають при проголошенні деяких звуків мови.

Наступні ознаки базуються на коефіцієнтах лінійного передбачення (Linear Prediction Coding Coefficients, LPC). Лінійне передбачення вже тривалий час залишається одним з основних підходів до завдань цифрової обробки мови. Воно може використовуватися для оцінки періоду основного тону, формант та інших основних параметрів мови.

Принцип методу лінійного передбачення полягає в тому, що ділянку мовного сигналу можна апроксимувати лінійною комбінацією попередніх ділянок сигналу. Передбачається, що мова створюється порушенням лінійного фільтра, що змінюється в часі, (мовного тракту) випадковим шумом для невокалізованих мовних сегментів або послідовністю імпульсів для голосової мови.

Спрощений процес мовотворення описується лінійною системою зі змінними параметрами. Як приклад – алгоритми лінійного передбачення з мультікодовим управлінням (Code Excited Linear Prediction, CELP), які широко використовуються в сучасних мережах стільникового зв'язку.

Питанням в методі LPC залишається сигнал збудження x (n), який змінюється не суттєво, чи, замінюється на розрахований раніше (наприклад, як CELP).

Сигнал, який розраховується, на виході мовного тракту в момент часу n має такий вигляд

$$\upsilon(n) = Gg(n) + \sum_{k=1}^{p} a_k \upsilon(n-k), \qquad (2.4)$$

де G – деяка константа (коефіцієнт посилення);

g (n) – вхідний голосовий сигнал фільтра в момент часу n;

a_k – деякі коефіцієнти;

$\upsilon(n-k)$ – попередні вихідні сигнали фільтру в моменти часу n-1, n-2,… n-p.

Означене рівняння може бути реалізовано за допомогою цифрової схеми, яка показана на рис. 2.5.

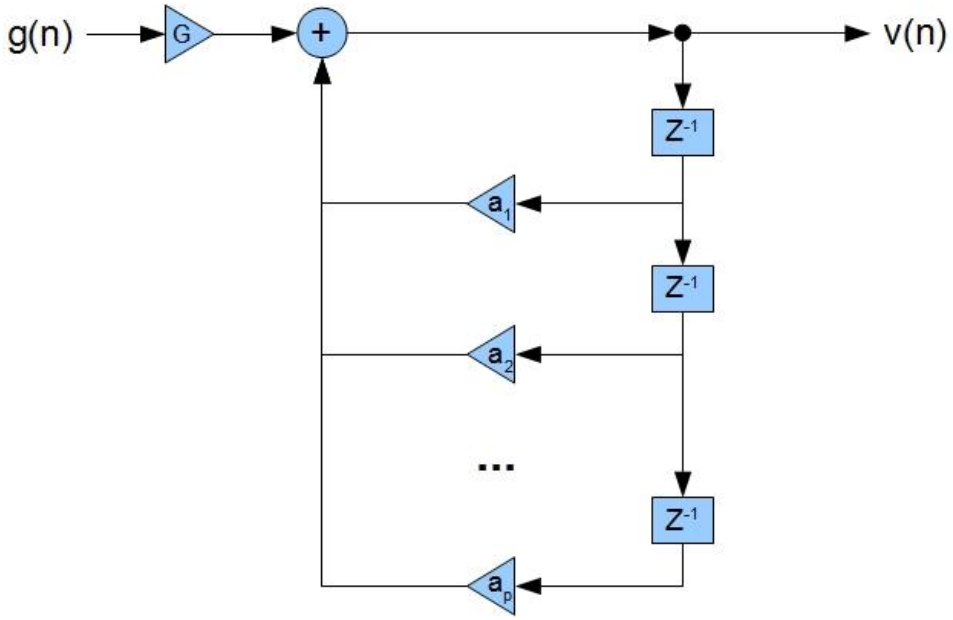

Рисунок 2.5 – Схема апроксимації мовного тракту

Таким чином, в якості шаблону LPC-коефіцієнтів виступають наступні величини: константа G і вектор коефіцієнтів $a_1, a_2,\ldots,a_p$.

Суть обчислень полягає в знаходженні лінійних коефіцієнтів передбачення a_k по мовному сигналу з мінімізацією похибки передбачення. Похибка передбачення *e (n)* визначається як

$$e(n) = \upsilon(n) - \sum_{k=1}^{n} a_k \upsilon(n-k). \qquad (2.5)$$

Існує три базових алгоритми розрахунку коефіцієнтів лінійного передбачення: коваріаційний, автокореляційний і сходовий. Їх докладний опис можна знайти в літературі по мовотворенню.

Домінуючим алгоритмом обробки голосових сигналів в автоматичних системах є знаходження кепстральних коефіцієнтів. Кепстром називається спектр логарифма спектру часової хвилі, який визначається як

$$c(n) = F^{-1}\{\log|F[x(n)]|\} \tag{2.6}$$

де F и F^{-1} – пряме і зворотнє дискретне перетворення Фур'є (ДПФ).

Доцільність використання кепстрального аналізу в задачах ідентифікації диктора полягає в тому, що кепстр описує огинаючу спектра сигналу в стислому вигляді. Більш детально порядок формування кепстральних та мел-частотних кепстральних коефіцієнтів (mel-frequency cepstral coefficients, MFCC) розглянуто в наступному розділі. Тут лише зазначимо, що згідно систематичного огляду, виконаному Сорокіним, серед наукових публікацій з розпізнавання диктора за 2011-2016 роки роботи із застосуванням методів MFCC склали 97%.

Окремі питання захисту локальних мереж розглянути в [72, 73].

3 МЕТОДИКА І РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ ГОЛОСОВОГО СИГНАЛУ КОРИСТУВАЧА ПРИ РОЗРАХУНКУ МЕЛ-ЧАСТОТНИХ КЕПСТРАЛЬНИХ КОЕФІЦІЄНТІВ

Питання розрахунку мел-частотних кепстральних коефіцієнтів займають значне місце, як в наукових дослідженнях, так і в використовуваних системах голосової автентифікації. При цьому слід зазначити, що в процесі розрахунку мел-частотних кепстральних коефіцієнтів використовується амплітудно-частотна інформація голосового сигналу користувача, а фазові дані не беруть участь в цифровій обробці.

Обумовлено це рядом чинників, серед яких виділимо основні:

- неможливість реєстрації фазових даних голосового сигналу;
- відсутність обчислювального ресурсу і процедур для цифрової обробки фазових даних.

Тому в даному розділі основна увага зосереджена на формуванні фазових даних і їх обліку при оцінці мел-частотних кепстральних коефіцієнтів.

3.1 Методика проведення досліджень голосового сигналу користувача системи автентифікації

В основу методики проведення досліджень в рамках магістерської роботи застосовуються такі наукові методи: аналіз, вимір, математичне моделювання та експериментальні дослідження.

Для цього була розроблена і активно використовувалася експериментальна установка, яка представлена на рис. 3.1.

Ядро експериментальної установки – ноутбук з операційною системою і звуковою картою. До звукової карти підключався мікрофон. При цьому була можливість управляти частотою дискретизації, кількість біт квантування амплітуди і часу проведення запису голосового сигналу.

Всі процедури запису голосового сигналу і його цифрової обробки здійснювалися в системі комп'ютерної математики (СКМ) MatLab. Голосовий сигнал зберігався в форматі wav.

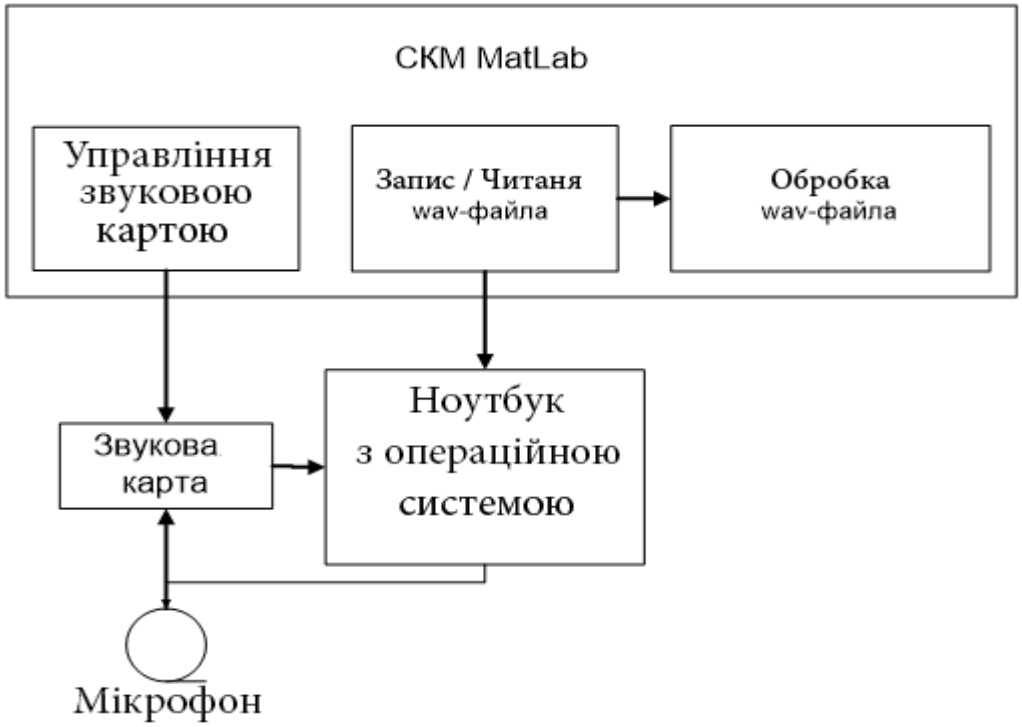

Рисунок 3.1 – Експериментальна установка для проведення досліджень

Цифрова обробка голосових даних передбачала використання як стандартних функцій, таких як

- розрахунок спектра аналізованого сигналу;
- перетворення Гільберта;
- формування графічного представлення результатів досліджень.

Одночасно використовувався і алгоритмічна мова СКМ MatLab для розробки m-файлів.

Тепер зупинимося на процедурах експериментальних досліджень, які будуть розглянуті нижче.

Аналізований голосовий сигнал включав цифри від 0 до 9, які можуть, наприклад, використовуватися для введення пін коду. Уникнути підробки голосового сигналу користувача – видача змісту пін коду на екран системи голосової автентифікації.

Частота дискретизації використовуваного голосового сигналу – 64 кГц, а кількість біт квантування його амплітуди становило 8 біт.

З голосового сигналу виділялися окремі цифри, які вимовив користувач системи автентифікації.

На основі виділеного сигналу можемо сформувати ознаки (шаблон) користувача, який або заноситься в базу, або порівнюється з шаблонами бази даних си-

стеми автентифікації. В якості ознак будемо використовувати кепстральні і мел-частотні кепстральні коефіцієнти.

Сигнал з мікрофона – речова частина, так званого аналітичного сигналу, який буде розглянуто докладно нижче.

На основі виділеного сигналу формувалися фазові дані аналізованого сигналу, які також будуть піддаватися цифровій обробці.

В подальшому можна порівняти результати формування коефіцієнтів, які отримані на основі амплітудно-частотних і фазо-частотних даних, а також зробити висновки.

3.2 Модель аналітичного сигналу

В радиолокации и радиосвязи давно, плодотворно и эффективно используется модель аналитического сигнала. К сожалению, у голосового сигнала достаточно сложно зарегистрировать фазовую составляющую, это одна из причин ограниченного использования фазовых данных в системах голосовой аутентификации.

Аналитический сигнал – некоторая математическая модель, которая позволяет ввести в рассмотрение фазовые данные голосового сигнала и тем самым расширить число используемых информационных параметров анализируемых данных. А это путь повышения эффективности систем голосовой аутентификации.

Таким образом, аналитический сигнал (аналитическое представление сигнала) – используемое в теории обработки сигналов математическое представление аналогового сигнала в виде комплекснозначной аналитической функции времени. Обычный, действительный сигнал (ток, напряжение) x является при этом действительной (вещественной) частью аналитического представления x_a.

Любой аналитический сигнал – комплексное число, которое можно представить в виде вектора, который исходит из начала координат до некоторой точки в комплексом пространстве. Этот вектор во времени описывает траекторию на комплексной плоскости, что показано на рис. 3.2.

Заметим, что модель аналитического сигнала была введена Д. Габором в 1946 году, которая позволила ввести понятие мгновенной амплитуды, частоты и

фазы сигналу. Введенные параметры позволили существенно упростить изучение и повысить эффективность цифровой обработки сигналов.

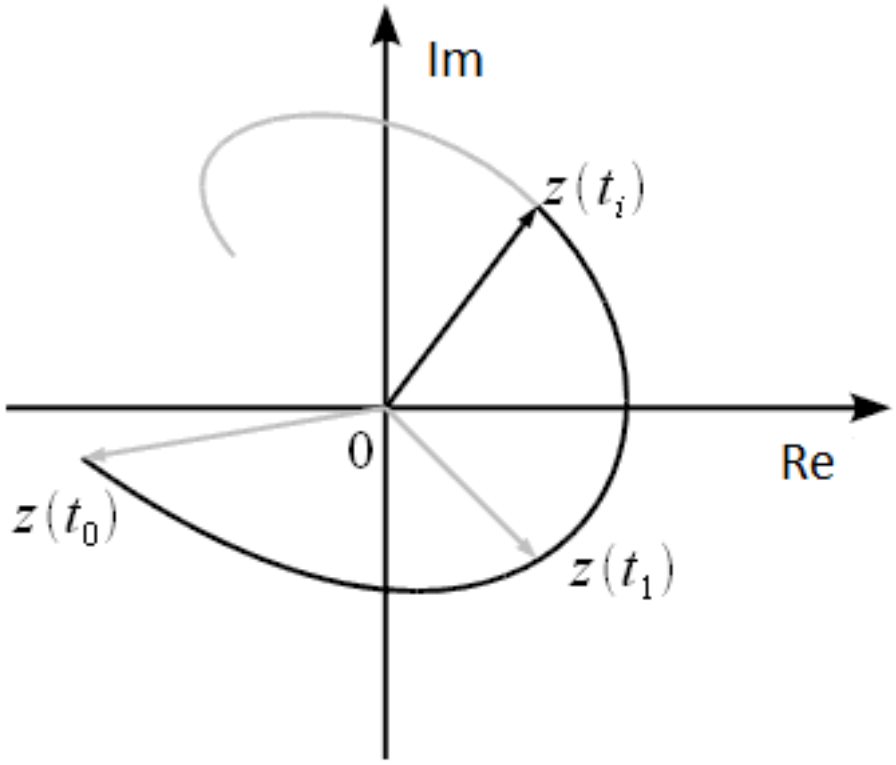

Рисунок 3.2 – Векторне подання комплексного сигналу

Непосредственно с рассмотренной моделью связаны и известные формулы Эйлера. Для рассмотрения комплексного изображения гармонической величины формула Эйлера имеет вид

$$e^{j\cdot(\omega\cdot t+\varphi)} = \cos(\omega\cdot t+\varphi) + j\cdot\sin(\omega\cdot t+\varphi), \qquad (3.1)$$

де ω – кругова частота;

φ – початкова фаза;

t – змінна часу.

Поэтому для неизвестной амплитуды ($A(t)$) и произвольного гармонического сигнала $x(t)$ имеем следующее выражение

$$x(t) = A(t)\cdot e^{j\cdot(\omega\cdot t+\varphi)} = A(t)\cdot\cos(\omega\cdot t+\varphi) + j\cdot A(t)\cdot\sin(\omega\cdot t+\varphi)\ . \qquad (3.2)$$

Векторное представление аналитического сигнала связано с двумя величинами, которые определяются косинусом (a) и относится к вещественной оси, а синусная составляющая (b) – к мнимой оси (див. рис.3.3).

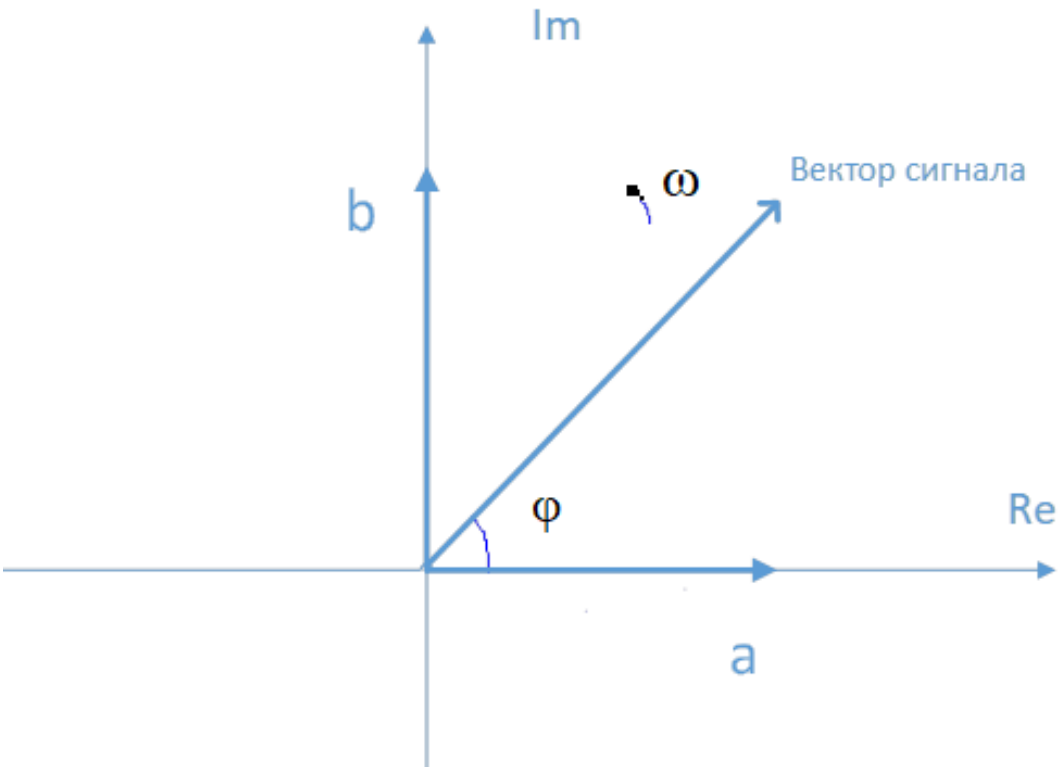

Рисунок 3.3 – Векторне подання аналітичного сигналу

После этого можем определить понятия мгновенной амплитуды, частоты и фазы сигнала. Вновь обратимся к рисунку 3.3 на котором отражен некоторый гармонический сигнал для заданного момента времени. Вектор сигнала может быть представлен в виде двух составляющих a i b.

Известно, что a имеет название синфазной (вещественной) составляющей, которая определяется следующим соотношением

$$a = A(t) \cdot \cos(\omega \cdot t + \varphi), \tag{3.3}$$

a b відповідно

$$b = A(t) \cdot \sin(\omega \cdot t + \varphi), \tag{3.4}$$

яка має назву квадратурної складової.

В розглянутому випадку (див. 3.2 – 3.4) φ це початкова фаза сигналу. Вектор аналітичного сигналу обертається проти годинникової стрілки навколо початку координат зі кутової (кругової) швидкістю ω.

Значення фази сигналу (φ) можна розраховувати для кожного моменту часу (t). Для цього використовують наступний вираз

$$\varphi(t) = \text{arctg}(b/a). \tag{3.5}$$

Таким чином, для визначення поточної фази необхідно визначити синфазну і квадратурну складові аналітичного сигналу.

В математичної моделі голосового сигналу застосовується функціональна залежність, в якої аргументом є циклічна ω або кругова (кутова) частота. Таким чином, голосовий сигнал можна розглядати як функція частоти. Вказана функціональна залежність, яка є спектральним поданням голосового сигналу та має назву спектру сигналу. Означене уявлення сигналу частіше розглядають не як власне сигнал, а як характеристику сигналу в частотній області.

В даний час знайшли широке застосування дискретні та цифрові системи що привело до необхідності використовувати дискретизовані сигнали. Зараз розрізняють сигнали: дискретні за часом; квантовані за рівнем амплитуди; цифрові (дискретні за часом і квантовані за рівнем).

В цифровому вигляді зареєстрований голосовий сигнал, а саме його синфазну складову, можна представити в вигляді

$$u_i = A_i \cdot \exp\left\{j \cdot \left[2 \cdot \pi \cdot f_o \cdot (i-1)/f_д + \varphi_i\right]\right\}, \tag{3.6}$$

де A_i – поточна амплітуда сигналу;

f_o – частота несучого коливання аналізованого сигналу;

$f_д$ – частота дискретизації сигналу;

φ_i – фазові дані сигналу, який аналізується;

$i = 1,..., N$ – номер відліку аналізованого сигналу;

N – кількість відліків цифрової послідовності, яка аналізується.

Зв'язок кругової (кутової) частоти і частоти несучого коливання пов'язані співвідношенням

$$f_o = \omega/(2 \cdot \pi). \quad (3.7)$$

Таким чином, в теорії сигналів геометричні методи базуються на представленні аналізованого сигналу як вектору в просторі можливих векторів, які повинні відповідати певним умовам, а саме лінійності та ортогональності. Зауважимо, що можливе використання поняття лінійного простору дійсних або комплексних сигналів з властивостями лінійного простору векторів.

Наявність загальних властивостей, що задовольняють принципам лінійності, є причиною об'єднання сигналів в безліч, що утворить простір сигналів. При цьому є можливість одні елементи безлічі висловити через інші елементи. Сучасні дослідження властивостей сигналів, які відповідають принципу суперпозиції, в рамках векторного уявлення виявляється корисним для синтезу пристроїв.

Періодичний аналітичний сигнал доцільно представити у вигляді суми нескінченного числа гармонійних складових (синусоїдальної і косинусоїдальної), які характеризується своєю амплітудою і частотою. Сукупність означених складових називають спектром сигналу, а сукупність їх амплітуд – амплітудно-частотним спектром сигналу, який будемо використовувати при розрахунку шаблону голосового сигналу користивуча.

Таким чином, з аналізу наведених співвідношень можна виділити наступні інформаційні параметри голосового сигналу: амплітуда, частота і фаза. На жаль, інший інформаційний параметр радіосигналів – поляризація, пов'язаний зі значними складнощами реєстрації для голосових сигналів. В сучасних системах голосової автентифікації широко використовується амплітуда і частота матеріалів реєстрації.

Далі зупинимося на методиці формування та застосування фазової інформації для голосових даних користувача.

У системах радіолокації і радіозв'язку, де широко і ефективно використовується фазові данні, для формування фази сигналу спочатку широко використовувалися фазові вращатели, які неможливо застосовувати в системах голосової автентифікації. Це було однією з причин не використання фазової інформації в системах голосової автентифікації.

Зараз стан справ у формуванні фазової інформації істотно змінилося, вона швидко та ефективно обчислюються програмно, за допомогою сучасних цифрових сигнальних процесорів або спеціалізованих мікросхем. В основу процедур формування фазових даних покладено перетворення Гільберту, яке має вигляд [35]

$$u_m(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{u(\tau)}{t-\tau} d\tau, \tag{3.8}$$

де τ – змінна інтегрування;

$u(\tau)$ – синфазна складова аналізованого сигналу.

В результаті перетворення отримаємо квадратурну (уявну) складову аналітичного сигналу.

Розрахунок фазової інформації в цифровому варіанті має такий вигляд

$$\varphi(i) = \operatorname{arctg}((u_m(i)/u(i)). \tag{3.9}$$

Перетворення Гільберту в СКМ MatLab реалізується за допомогою функції heming.

Функція arctg, на жаль, розраховує кути в діапазоні від $-\pi/2$ до $\pi/2$. У голосового сигналу змінюється фазовий кут в межах від 0 до 2π. Тому для визначення правильного значення фазового кута, необхідно $\varphi(t)$ відповідним чином відкоригувати з урахуванням знаків чисельника і знаменника в співвідношенні функції arctg.

Після корекції фазовий кут буде мати форму пилоподібного сигналу невідомої тривалості. Крім того, після формування фазової інформації доцільно виконати процедури попередньої обробки. Це обумовлено наступними факторами, серед яких виділимо:

- полігармонійний характер голосового сигналу, над яким здійснюється перетворенням Гільберту. Означене перетворення орієнтоване на роботу з гармонійними та стаціонарними даними;

- некоректні фазові дані при рівності нулю складової $u(i)$ у функції arctg;

- при малих значеннях елементів $u_m(i)$ або $u(i)$, останні можуть губитися в шумах округлення.

Вказані фактори призводять до того, що в фазових даних мають місце випадкові помилки і аномальні вимірювання. Цим обумовлена попередня обробка голосового сигналу та фазових даних.

Попередня обробка може базуватися на апріорних даних щодо формі фазової інформації голосового сигналу та дозволяє значно підвищити якість формування характеристик існуючих і перспективних складових шаблонів.

Фазова інформація в голосовій автентифікації може використовуватися за такими напрямками:

- підвищення співвідношення сигнал/шум матеріалів реєстрації (в радіолокації і радіозв'язку це відомий напрямок використання фази);

- підвищення якості формування окремих елементів традиційних шаблонів, наприклад, частоти основного тону, формантної інформації і т.д.;

- розробка нових елементів та процедур формування шаблонів за рахунок використання фазових даних.

3.3 Методика оцінки кепстральних коефіцієнтів голосового сигналу

В експериментах з розпізнавання голосу людини встановлено, що огинаюча амплітудного спектру значно впливає на його пізнаванність. В зв’язку з цим використання різних способів аналізу огинаючої амплітудного спектру з метою розпізнавання людини виправдано.

Тому в якості унікального вектору ознак людини використовують одновимірний вектор кепстральних коефіцієнтів, а також вектор, який складений з його похідних.

Спочатку визначимо, що таке кепстр [35].

Перехід з часової області в частотну дозволяє отримати більш наочні, докладні і компактні данні, тобто, відбувається стиснення інформації. При цьому, для більш "простих" часових сигналів відбувається більшою мірою його стиснення в частотної області. Спектр дозволяє отримати наочне, компактне, чисельне представлення періодичностей, які присутні в сигналах в часовій області.

З переваг спектрального представлення сигналів прийшла ідея кепстрального аналізу. На практиці не часто зустрічаються вібрації гармонійного характеру, в

багатьох випадках спектр вимагає тривалого і вдумливого аналізу. Спрощення аналізу спектру можна досягти за рахунок заміни в спектрі вісь частоти на вісь часу, а саме представити, що спектр є просто сигналом.

Це дозволяє легше визначити невидимі періодичності в означеному "сигналі". Зазначимо, що присутні у спектрі періодичності є гармонійні ряди. Значить, з'являється добра можливість представити спектральну інформацію більш компактна, коли кожний гармонійний ряд спектру буде представлений в ідеалі однією складової в кепстрі.

Розглянемо означене на прикладі. На рис. 3.4 представлений спектр аналізованого сигналу. Цей спектр визначений в діапазоні частот від 0 до 80 Гц з кроком 1 Гц. Таким чином, спектр складається з 80-ти складових. Основні дискретні складові в аналізованому спектрі з двох частотних рядів: гармоніки з частотою 10 Гц (ряд r1) і з частотою 18 Гц (ряд r2).

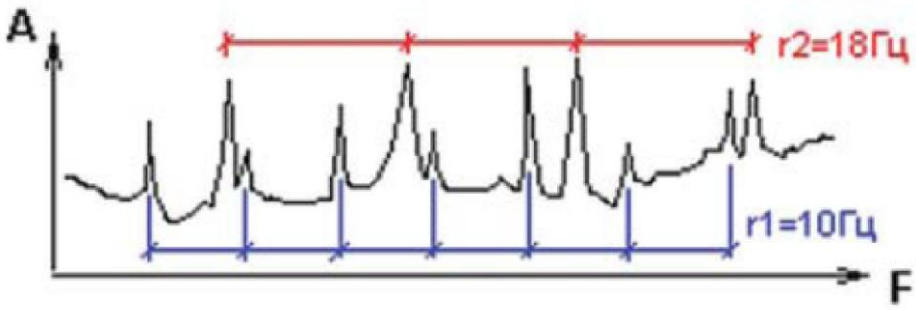

Рисунок 3.4 – Спектр аналізованого сигналу

Для розрахунку кепстру необхідно замінити вісь частот на вісь часу. Таким чином, в кепстрі доцільно чекати дві частотних складових (це показано на рис. 3.5), які характеризують ряди 10 Гц і 18 Гц аналізованого спектру. Дійсно на рис 3.5 можна виділити дві складових з періодом 0,055с і 0,1 с, що відповідає частотам 10 і 18 Гц.

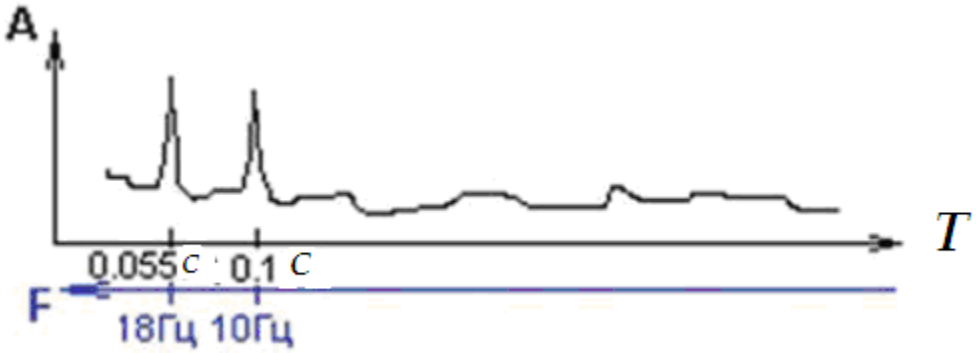

Рисунок 3.5 – Кепстр аналізованого сигналу

Цей приклад визначає, що кепстр дає можливість значно спростити взначення гармонійних складових з спектру, який аналізується.

Необхідно мати на увагу, що спектр може бути сильно «порізаним», а саме мати значну кількість піків в спектрі. Тому для зменшення кількості максимумів в кепстрі спектр спочатку логарифмують. Процедура логарифмування спрямована тільки на те, щоб хоч якось згладити від спектру, який аналізується і, відповідно, зменшити кількість паразитних піків в кепстрі.

Зменшення кількості високочастотних складових в спектрі, можна досягти за рахунок використання перетворення Гільберта-Хуанга та виключення перших модових функцій.

Аналізований голосовий сигнал спочатку ділиться на фрейми - ділянки по 25-30 мс з перекриттям фреймів рівним приблизно10 мс.

У ряді робот застосовується додаткова операція попередньої обробки сигналу. Наприклад, для кожного фрейми сигналу застосовується деяка вагова функція (вікно Хеммінгу)

$$w_n = 0.54 - 0.46 \cdot \cos(2\pi \cdot \frac{n}{N-1}), \quad n = 0,1,\ldots,N-1, \qquad (3.10)$$

де N – довжина вікна (фрейму), яка визначена в відліках. В якості вагової функції може бути застасоване вікно Хеммінга (див. співвідношення (3.10)). Означена процедура, в першу чергу, призначена для зменшення спотворень в перетворенні Фур'є на краях аналізованого фрейму. Математично це може мати такий вигляд

$$X_k = \sum_{n=0}^{N-1} w_n \cdot x[n] \cdot \exp(-\frac{2\pi i}{N} kn), \quad (3.11)$$

де $x[n]$ – n-й елемент вибірки. А значення індексів k відповідають частотам

$$f_k = \frac{F_S}{N} k, \quad (3.12)$$

де F_S – частота дискретизації аналізованого сигналу.

Таким чином, реалізація цього способу обробки така: на інтервалі часу в 25 - 30 мс розраховується поточний спектр потужності сигналу, а до логарифму результату застосовується зворотне перетворення Фур'є.

Тепер проаналізуємо вплив вікна Хеммінга на спектр фрагменту аналізованого сигналу. Фрагмент аналізованого сигналу цифри «один» (чорний колір) та вікно Хеммінга (червоний колір) при вихідних даних: $N = 256$, $F_S = 16$ *кГц* наведені на рис. 3.6. На рис. 3.7 наведені два спектри: чорний колір без урахування вікна Хеммінга; червоний колір – з урахуванням вікна Хеммінга.

Аналіз вказаних графіків свідчить про вплив на краях аналізованого фрагменту вікна Хеммінга.

На рисунку 3.8 наведена структурна схема за допомогою, якої визначаються кепстральні коефіцієнти. На рис. 3.8 використовуються наступні позначення:

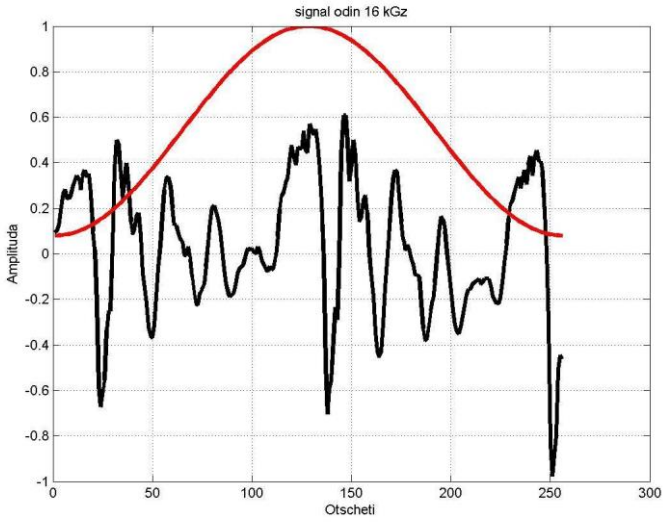

Рисунок 3.6 - Фрейм голосового сигналу та вікно Хеммінга

FFT – блок швидкого перетворення (БШП) Фур'є сигналу; LOG – процедури логарифмування спектру; IFFT – блок зворотного БШП Фур'є.

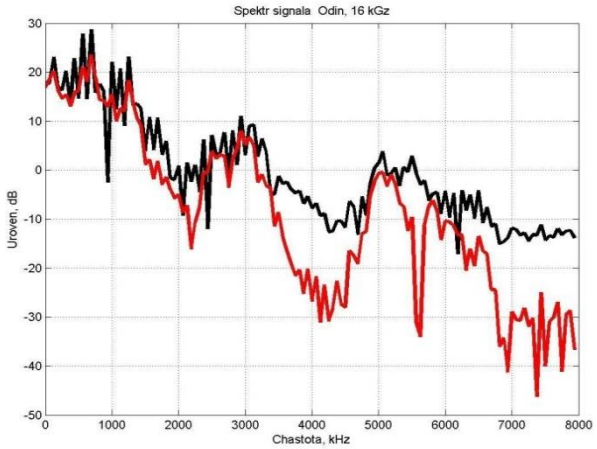

Рисунок 3.7 – Спектр аналізованого фрейму сигналу

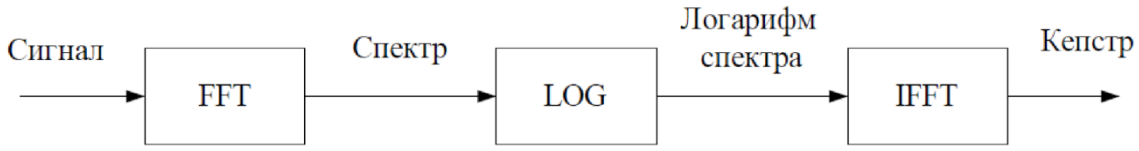

Рисунок 3.8 – Структурнана схема розрахунку кепстру

Тепер визначмо порядок розрахунку кепстральних коефіцієнтів. Він наступний

$$c_n = \int_0^{\Theta} \ln|S(jw,t)|^2 e^{-jn\Omega w} dw, \qquad (3.13)$$

де $\Omega = 2\pi/\Theta$, Θ – верхня частота в спектрі сигналу,

$|S(jw,t)|^2$ – спектр потужності сигналу.

Число кепстральних коефіцієнтів n визначається необхідною якістю згладжування спектру, та, як правіло, знаходиться в межах декількох десятків (від 20 до 40).

В разі застосування гребінки фільтрів коефіцієнти кепстрального перетворення визначаються з використанням співвідношення

$$c_n = \sum_{m=1}^{K} \ln[Y(m)]^2 \cos\left[\frac{\pi n}{M}(m-\frac{1}{2})\right], \qquad (3.14)$$

де $Y(m)$ – аналізований сигнал m-го фільтру;

c_n – n-й кепстральний коефіцієнт;

K – число фільтрів, які застосовуються.

В співвідношенні (3.14) $\cos\left[\frac{\pi n}{M}(m-\frac{1}{2})\right]$ – дискретне косинусне перетворення, яке виконує функцію зворотного БШП Фур'є.

Таким чином, розраховані кепстральні коефіцієнти – є результат застосування зворотного БШП Фур'є до логарифму енергетичного спектру аналізованого сигналу.

3.4 Методика розрахунку мел-кепстральних коефіцієнтів

Означений метод виділення ознак, що указано вище, є одним з найпоширеніших як в системах розпізнавання дикторів та мови, так і в системах голосової автентифікації. На вхід процедур поступає послідовність відліків аналізованого фрейму сигналу, який досліджується на даній ітерації, x0, ..., xN-1. Як й при розрахунку кепстральних коефіцієнтів, до означеної послідовності застосовується, так звана вагова функція і потім дискретне швидке перетворення Фур'є.

Розглянімо інший метод застосування кепстральних коефіцієнтів, що враховують особливості слуху людини. Для цього застосовується спеціалізована шкала. Тому виконується аналіз фреймів сигналу за рахунок виділення кепстральних коефіцієнтів за мел-шкалою (Mel Frequency Cepstral Coefficient - MFCC) [36].

В означеному методі аналізу використовується модель функціонування органів людського слуху і застосовує частотну шкалу мел, яка враховує частотну чутливість вуха людіни. Відомо, що мел-шкала лінійна до 1 кГц та логарифмічна на більших частотах. MFCC – представлення сигналу реалізовано як дійсний кепстр фрейму сигналу, який оброблений з використанням процедур швидкого перетворення Фур'є та є відображенням енергетичного спектру сигналу на мел-шкалу.

На мел-шкалу відображення виконується за допомогою блоку, так званих, трикутних фільтрів (смугасто-пропускаючи фільтри), які лінійно розташовані за мел-шкалою. Число MFCC-коефіцієнтів – від 10 до 30.

Властивості людського слуху, як правіло, враховуються за рахунок нелінійного перетворення шкали частот в мел-шкалу . Означена шкала розраховується виходячи з присутності в людському слуху, так званих критичних смуг, таких, що сигнали будь-якої частоти в межах критичної смуги, невиразні.

Мел - шкала розраховується за наступним співвідношенням

$$m(f) = 1127 \cdot \ln(1 + f / 700), \tag{3.15}$$

де f – частота в Гц;

m – частота в мелах.

Для зворотного перерахунку мел-шкали в частоту застосовується наступне співвідношення

$$f = 700 \cdot (e^{m/1127} - 1). \tag{3.16}$$

На рисунку 3.9 наведена залежність психофізичної одиниці висоти звуку мел від частоти.

Є інша шкала , яка має назву барк. В шкалі барк різниця між двома частотами, які дорівнюють критичній смузі, рівна 1 барк. Частота в барках розраховується таким чином

$$B = 13 \cdot \operatorname{arctg}(0.00076f) + 3.5 \cdot \operatorname{arctg}(\frac{f}{7500}). \tag{3.17}$$

Рорахунок мел кепстральних коефіцієнтів здійснюється за допомогою співвідношення

$$c(n) = \sum_{k=1}^{U} \ln(S(k)) \cdot \cos(\frac{\pi n}{U}(k - \frac{1}{2})), \quad n = 0,1,\ldots,N, \tag{3.18}$$

де $S(k)$ – середня спектральна потужність k-го фільтру;

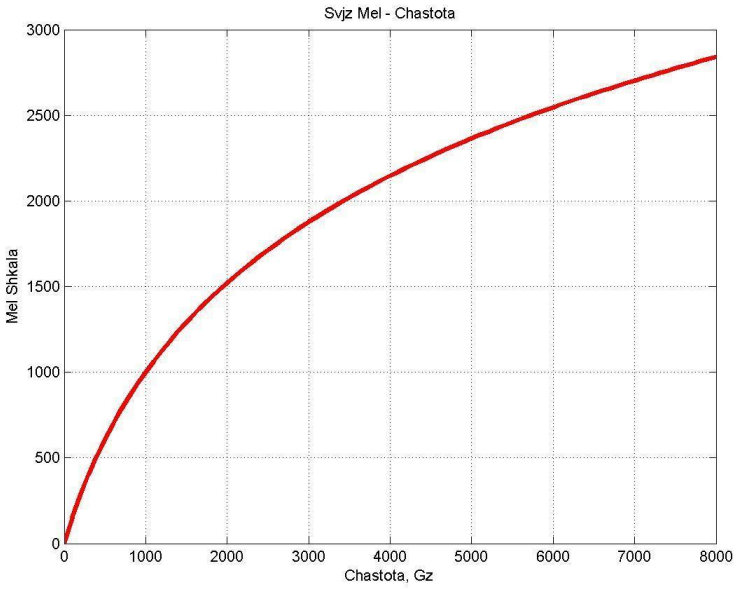

Рисунок 3.9 – Зв'язок мел-шкали з частотою

U – число фільтрів, які застосовуються;

N - число кепстральних коефіцієнтів, які розраховуються.

Середня спектральна потужність *k*-го фільтру розраховується за допомогою співвідношення

$$S(k) = \frac{1}{N(k)} \sum_{i=M(k)}^{M(k)+N(k)} w(k,i) \cdot |X(i)|, \qquad (3.19)$$

де *k* – номер фільтру (від 1 до 20), який розраховується;

M(k) – початкова частота *k*-го фільтру;

N(k) – ширина *k*-го фільтру;

$w(k,i)$ – деяка вагова функція трикутної форми;

X(i)-я амплітуда дискретного перетворення Фур'є (відлік спектру) для аналізованого фільтру.

За рахунок множення функції на фільтр, здійснюється усереднення її на деякій ділянці.

Визначимо методику розрахунку трикутних фільтрів. Число трикутних фільтрів, як правіло, відповідає кількості кепстральних коефіцієнтів, які розраховуються. Методику розрахунку характеристик трикутних фільтрів розглянемо на прикладі. Потрібно розрахувати десять трикутних фільтрів в діапазоні частот від

40 Гц до 8 кГц. Вказаний діапазон частот в мел-шкалі має наступні межі: від 62.6 до 2840 мел. Відомо, що людина краще розрізняє звуки на низьких частотах. Тому кількість фільтрів на низьких частотах має бути більшою. Для врахування цього факту доцільно визначений діапазон (від 62.6 до 2840 мел) мел-шкали, які займають десять трикутних фільтрів, розбити на однакові інтервали. В зв'язку з цим маємо наступні значення мел-шкали:

$$f_m = 62.6,\ 315.2,\ 567.6,\ 820.1,\ 1072.6,\ 1325.1,\ 1577.6,\ 1830.1,\ 2082.6,\ 2335.1,\ 2587.5,\ 2840. \quad (3.20)$$

Далі здійснюється перерахунок кордонів фільтрів в частоту за допомогою співвідношення (3.16). В цього випадку будуть такі значення:

$$f_r = 40,\ 225.8,\ 458.3,\ 749.2,\ 1113.1,\ 1568.4,\ 2138.1,\ \ldots\ldots\ldots\ldots\ldots\ldots 2850.8,\ 3742.5,\ 4858,\ 6253.8,\ 8000. \quad (3.21)$$

Аналіз (3.21) свідчить про різну частотну ширину трикутних фільтрів, а саме, в області низьких частот вони займають смугу близько 200 Гц, а в кінці аналізованого діапазону – близько 1500 Гц. Таким чином, маємо більшу кількість трикутних фільтрів в області низьких частот.

Черговий шаг пов'язаний з визначенням опорних точок для трикутних фільтрів. Визначений діапазон частот в спектральної області включає опорні точки, за якими будемо розраховувати трикутні фільтри. Визначення опорних точок в спектральний області виконується з застосування співвідношення

$$\mathrm{f_o(i)} = \left[\frac{\mathrm{K \cdot f_r(i)}}{\mathrm{F_S}} \right], \quad (3.22)$$

де K – число відліків аналізованого спектру;

$\mathrm{F_S}$ – частота дискретизації сигналу.

Для прикладу $K = 256$ и $F_S = 16$ кГц маємо наступний масив опорних точок

$$f_o(i) = 1,\ 4,\ 8,\ 12,\ 18,\ 26,\ 35,\ 46,\ 60,\ 78,\ 101,\ 128. \quad (3.23)$$

Слід зауважити, що $f_o(i)$ – номери відліків на осі частот спектру сигналу, число яких дорівнює половині обсягу аналізованого фрейму. Співвідношення поділу осі частот спектру сигналу на аналізовані фільтри (реєстрація масиву опорних частот) має вигляд:

$$f_o(i) = \left[\frac{K}{F_S} m^{-1}(m(f_{min}) + i\frac{m(f_{max}) - m(f_{min})}{M+1}) \right] \ldots\ldots\ldots\ldots\ldots\ldots (3.24)$$

де M – кількість кепстральних коефіцієнтів, які розраховуються (оброблюваних семплів);

$m(f)$ – функція перерахунку з частоти з одиниць в герц в одиниці в мелах;

m^{-1} – співвідношення перерахунку частоти з мел-шкали в шкалу герц;

f_{min} – мінімальна частота в герцах, яка приймає участь в розрахунку;

f_{max} – максимальна частота в герцах, яка приймає участь в розрахунку.

Визначено співвідношення (3.24) краще, оскільки відсутний етап розрахунку та розбиття аналізованого діапазону в мелах і герцах, а відразу визначає опорні частоти. Кількість елементів масиву опорних частот повинна дорівнювати $M+2$.

Визначимо алгоритм розрахунку аналізованих трикутних фільтрів. Алгоритм цього розрахунку передбачає участь в означеної процедури 3 опорних частоти (ліва, центральна і права). Визначимо ліва частота – початок трикутного фільтру. Центральна частота, відповідно, – це максимальне значення (вершина) аналізованого фільтру, права визначена частота – це кінцеве значення фільтру.

Для m-го фільтру використовується наступна формула

$$H_{m-1}(n) = \begin{cases} 0, & \text{если} \quad n < f_o(m-1); \\ \dfrac{n - f_o(m-1)}{f_o(m) - f_o(m-1)}, & \text{если} \quad f_o(m-1) < n < f_o(m); \\ \dfrac{f_o(m+1) - n}{f_o(m+1) - f_o(m)}, & \text{если} \quad f_o(m) < n < f_o(m+1); \\ 0, & \text{если} \quad n > f_o(m+1); \end{cases} \qquad (3.25)$$

Визначимо, що аргумент n змінюється з кроком один в межах $f(m-1)-const < n < f(m+1)+const$. Величина const має декілька одиниць. Результати розрахунку першого фільтру $m=2$, а десятого $m=11$ (див. рис. 3.10).

На рис. 3.10 представлений зовнішній вигляд та границі 1 та 10 фільтрів.

Енергію для кожного аналізованого фільтру (вікна) можна розрахувати таким чином

$$S(m)=\ln\left(\sum_{k=0}^{K_m}\left|X_a(k)\right|^2\cdot H_m(k)\right),\quad 0\le m<M, \tag{3.26}$$

де K_m – число відліків в m-ом аналізованому фільтрі;

$X_a(k)$ – k-я складова щільності амплітудного спектру;

m – номер аналізованого трикутного фільтру;

M – число фільтрів (кепстральних коефіцієнтів), які розраховуються.

Як відомо, функція ln надає можливість згладити окремі піки амплітудного спектру. Крайній крок розрахунку пов'язаний зі зворотним перетворенням Фур'є та визначенням мел-кепстральних коефіцієнтів. Цю задачу виконує дискретне косинусное перетворення. Визначене перетворення – властивість компактності енергії: більшої енергії сигналу відповідає менша кількість коефіцієнтів. Для виконання розрахунків використовується наступне співвідношення

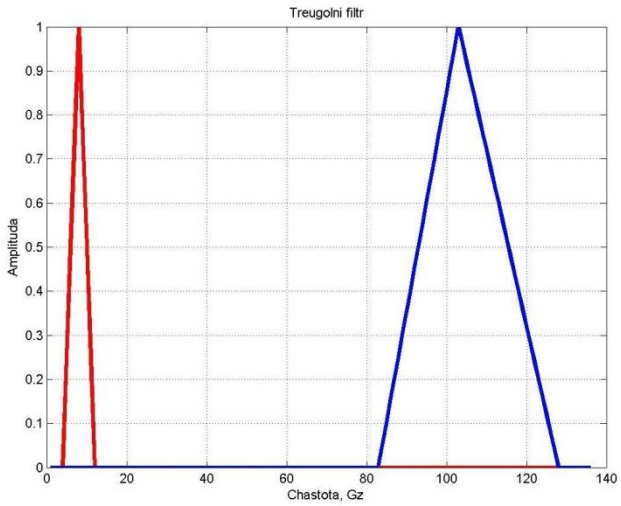

Рисунок 3.10 – Результат розрахунку першого і десятого трикутних фільтрів

$$c_o = \sum_{k=0}^{M-1} S(k) \cdot \cos\left(\frac{\pi n}{M}(k + \frac{1}{2})\right), \quad n = 0,1,2,..., M-1, \tag{3.27}$$

Коефіцієнт c_o не використовується. Цей коефіцієнт включає енергію сигналу. Загальне число коефіцієнтів M, зазвичай, вибирають від 12 до 30.

Заключний етап – розрахунок дельта значень та подвійних дельта значень коефіцієнтів, які отримані на попередніх етапах формування ознак. Дельта значення кепстральних коефіцієнтів обчисліваються за формулою

$$d(t) = \frac{c(t+1) - c(t-1)}{2}. \tag{3.28}$$

Подвійні дельта значення розраховуються так само, тільки замість кепстральних коефіцієнтів застосовують обчислені перші дельта значення. Перші та другі різниці кепстральних коефіцієнтів втричі збільшують розмірність простору прийняття рішень (ускладняється та збільшуються обчислювання), однак покращує ефективність, наприклад, розпізнавання диктора або процедур автентифікації користувача.

3.5 Результати експериментального дослідження голосового сигналу користувача системи автентифікації

При формуванні ознак шаблону широке застосування знайшли такі ознаки голосового сигналу: частота основного тону, формантна інформація, спектральні і кепстральних коефіцієнти. Серед ознак особливе місце займають мел-частотні кепстральні коефіцієнти, які дозволяють додатково вирішувати такі завдання: розпізнавання емоцій, визначення статі, сегментації аудіо з декількома голосами і поділу мови на фрази, визначати патологічні характеристики голосу в медицині та ін.

Аналізу піддавався голосовий сигнал користувача, який вимовляв цифри від 0 до 9. Частота дискретизації сигналу становила 64 кГц. Відношення сигнал / шум аналізованої послідовності становило понад 25 дБ.

При цьому основна увага буде приділятися аналізу діапазону спектра до 8 кГц, що обумовлено наявністю характерних ознак користувача в його сигналі (у діапазоні від 0,1 кГц до 8 кГц).

Методика проведених досліджень включала, як зазначено вище, використання перетворення Гільберту для формування квадратурної складової голосового сигналу. На основі квадратурної складової формувалися фазові дані.

Наступні процедури пов'язані з використанням співвідношень розрахунку MFCC по амплітудної і фазової інформації голосового сигналу. Процедури формування MFCC на основі амплітудної інформації розглянуті вище. Відомо, що розраховується до сорока цих коефіцієнтів.

Для цього використовуються семпли голосового сигналу в кілька десятків мілісекунд з перекриттям. Для обраних семплів виконується перетворення Фур'є, спектр якого логарифмується, а потім виконується зворотне перетворення Фур'є, на основі якого з певних співвідношеннях розраховуються MFCC. Таким чином, отримаємо до сорока оцінок MFCC по амплітудному і фазового сигналу. Далі виконувалася статистична обробка отриманих результатів та порівняльна характеристика оцінок за допомогою коефіцієнта кореляції.

Результати формування MFCC розглянемо на прикладі обробки голосового сигналу цифри «один», який представлений на рис. 3.11.

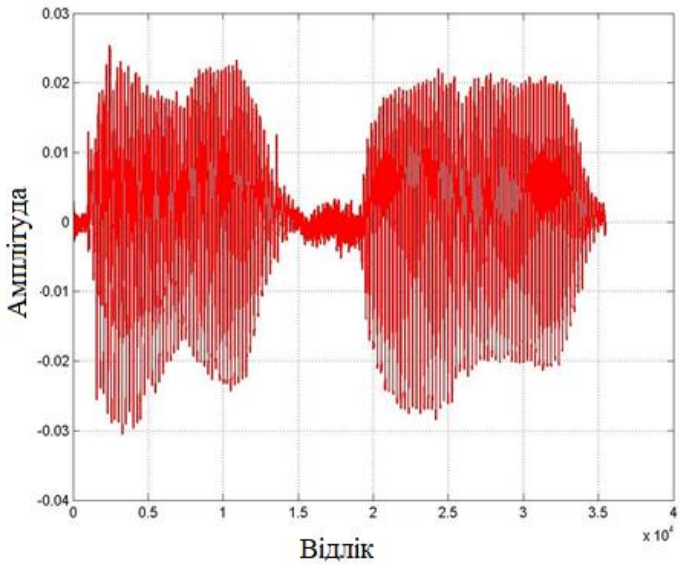

Рисунок 3.11 –Голосовий сигнал цифри «один»

Далі необхідно за матеріалами реєстрації за допомогою мікрофона сформувати квадратурну складову голосового сигналу. Для цього використовувалося перетворення Гільберту і на цій основі сформувати фазові дані голосового сигналу. Фрагмент розглянутого сигналу з двома складовими представлений на рис. 3.12 у верхній частині, а в нижній частині представлені відповідні фазові дані.

Аналіз представлених результатів свідчить, що при формуванні квадратурної складової і фазових даних можу мати місце помилки.

Тому після формування фазових даних повинен мати місце етап попередньої обробки матеріалів реєстрації, розрахована квадратурна складова і фазові дані.

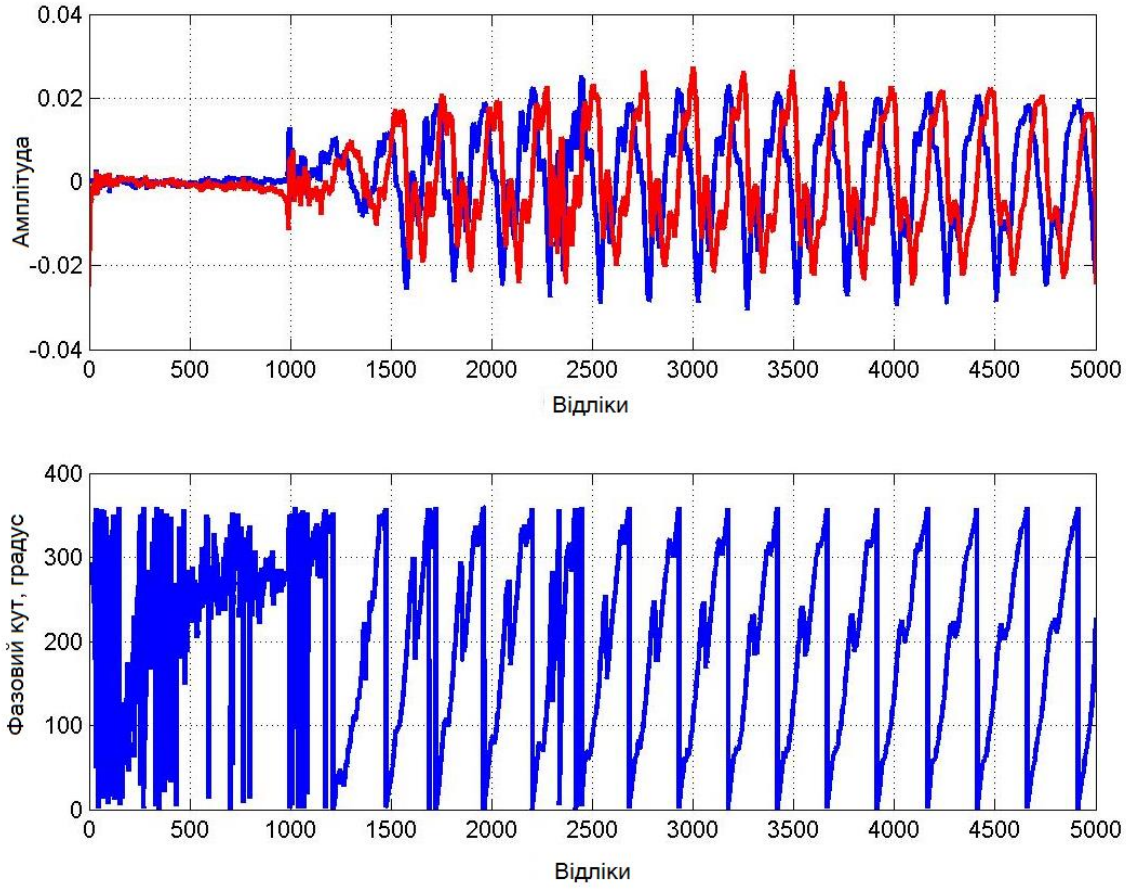

Рисунок 3.12 – Фрагмент реєстрації та обробки голосового сигналу цифри «один»

Основу для зазначеної попередньої обробки можуть бути апріорні дані про форму фазових даних, які повинні мати форму пилоподібного сигналу невідомої тривалості (див. рис. 3.13 – зліва).

На цьому ж малюнку праворуч показаний фазовий сигнал, який має помилки. На цій апріорної інформації можна скорегувати, як результати реєстрації з мі-

крофона, так і квадратурну складову і фазові дані. Для цього необхідно виконати процедури виділення фазових даних (пилоподібного сигналу), а потім провести аналіз причин спотворення фазових даних і програмно їх усунути.

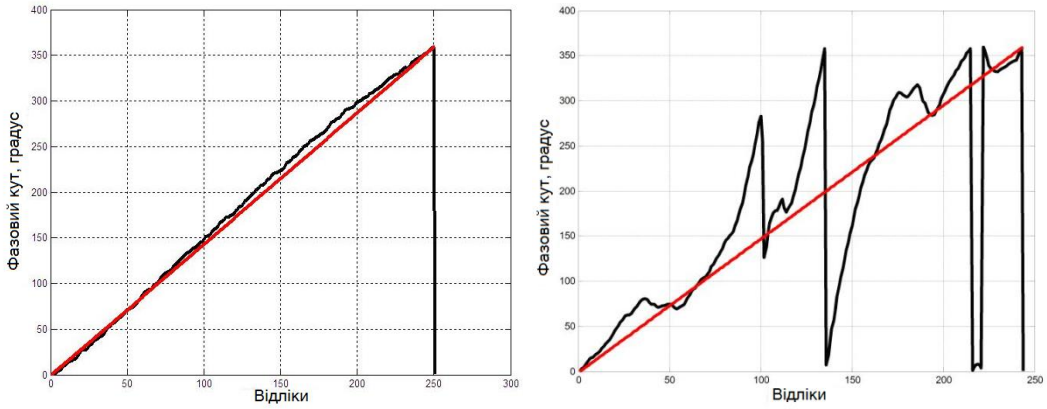

Рисунок 3.13 – Фазові дані матеріалів реєстрації

При цьому на рис. 3.13 червоним кольором показаний очікуваний фазовий сигнал, а чорним – результати розрахунку. Для виявлення помилкових фазових даних можна використовувати в якості критерію квадрат різниці між очікуваними і отриманими значеннями фазового кута.

Спочатку розраховуються трикутні фільтри, співвідношення для розрахунку яких розглянуті вище.

Далі амплітудні дані розбиваються на семпли. У розглянутому прикладі довжина семплу становить 1024 відліків, що дорівнює для даної частоти дискретизації 16 мс. Перекриття семплів становило 0.75. Перед обробкою як амплітудні, так і фазові дані піддаються нормалізації, а саме, поділу на відповідне максимальне значення.

Подальша обробка здійснювалася відповідно до розглянутими вище процедурами. А саме, розрахунок спектру, логарифмування, зворотне перетворення Фур'є у вигляді дискретного косинусного перетворення і розрахунок MFCC. У розглянутому прикладі розраховувалося 10 коефіцієнтів. Далі вибирався другий семпл і розрахунки повторювалися.

Результати обробки експериментальних амплітудно-частотних даних голосового сигналу цифри «один» представлені на рис. 3.14. При цьому оцінювалася 10 MFCC в 35 семплах.

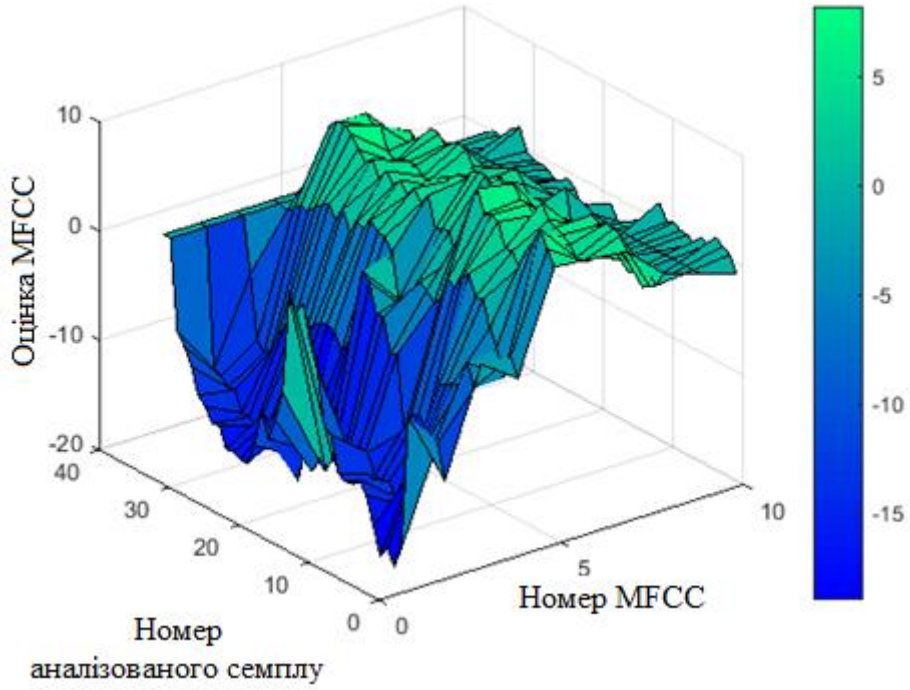

Рисунок 3.14 – Оцінки MFCC по амплітудної інформації цифри «один»

Аналогічні процедури виконувалися і для фазової інформації розглянутого голосового сигналу, а результати наведені на рис. 3.15.

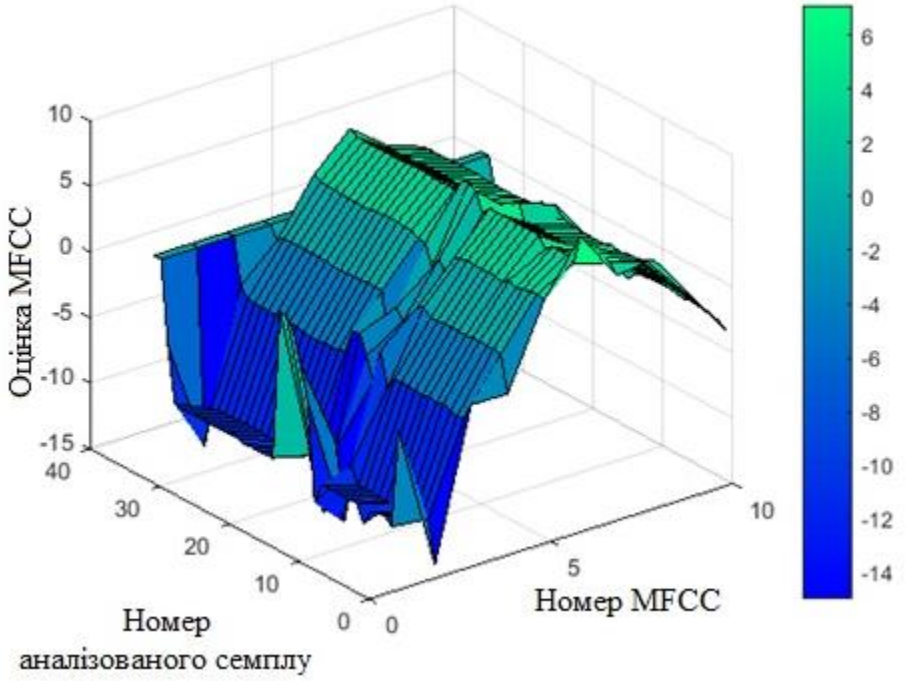

Рисунок 3.15 – Оцінки MFCC по фазової інформації цифри «один»

При цьому еталонними вважалися MFCC, розраховані за амплітудною інформацією (див. рис.3.14), які широко використовуються в процедурах автентифікації користувачів.

Порівняльний аналіз рис. 3.14 і 3.15 свідчить про схожість змін MFCC для амплітудної і фазової інформації. Більш того MFCC, розраховані по фазовим даними мають більш стійкий характер (менше «порізаний» графік). Для статистичної оцінки результатів (оцінка ступеня схожості) на рис. 3.16 представлена залежність нормованого коефіцієнту кореляції MFCC для різних семплів.

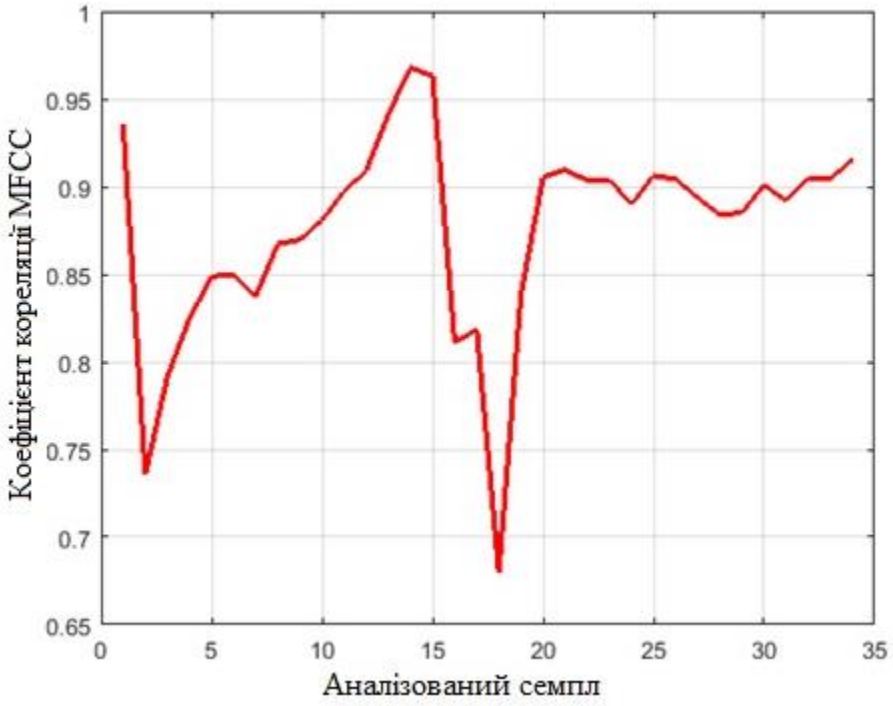

Рисунок 3.16 – Залежність коефіцієнта кореляції MFCC від номера семплу

Для більшості семплів коефіцієнт кореляції перевищує значення 0.8, що свідчить про високий збігу розрахованих даних. Останнє дає підставу використовувати оцінки MFCC, розраховані за фазовою інформацією, в шаблонах користувача системи автентифікації.

ВИСНОВКИ

У магістерській роботі розглянуто актуальне наукове завдання підвищення ефективності систем голосової автентифікації.

Основнім напрямком вирішення зазначеного завдання запропоновано в процесі цифрової обробки застосовувати фазові дані голосового сигналу, який обробляється.

Адекватність та достовірність зазначеного варіанту вирішення вказаної задачі і аналіз ефективності застосування фазових даних голосового сигналу досліджується в процесі експериментальної оцінки мел-частотних кепстральних коефіцієнтів, які входять до більшості шаблонів користувача системи автентифікації в якості обов'язкових параметрів.

Крім цього, мел-частотні кепстральні коефіцієнти дозволяють додатково вирішувати такі завдання: сегментація аудіо з декількома голосами і поділі мови на фрази, розпізнавання емоцій, визначення статі людини, яка говорить та ін.

Таким чином, в роботі вирішується актуальне наукове завдання розробки та дослідження нових процедур для уточнення оцінок мел-частотних кепстральних коефіцієнтів. Уточнення означених оцінок проводилося на основі застосування фазових даних голосового сигналу.

Результати роботи отримані в процесі статистичного аналізу результатів моделювання з використанням експериментальних голосових сигналів користувача системи автентифікації.

Як показали результати досліджень, фазові дані голосового сигналу дозволяють отримувати адекватні і достовірні оцінки в процесі оцінки мел-частотних кепстральних коефіцієнтів. Розроблено та досліджено методика отримання адекватних та достовірних оцінок мел-частотних кепстральних коефіцієнтів на основі застосування фазових даних голосового сигналу. Достовірність методики підтверджена в процесі модельного експерименту.

Результати наукових досліджень оприлюднені в п'яти наукових працях.

Подальші дослідження доцільно проводити в напрямку оцінки якості формування ознак для традиційно використовуваних шаблонів з урахуванням фази голосового сигналу, а також розробки нових процедур формування елементів шаблонів на основі фазових даних.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Beigi H. Fundamentals of Speaker Recognition. / H. Beigi. – NY: Springer, 2011.–1029 p.

2. ISO/IEC 2382-37:2012 Information technology – Vocabulary – Part 37: Biometrics.

3. Zaika M. "Investigation of Informativeness and Stability of Mel-Frequency Cepstral Coefficients Estimates based on Voice Signal Phase Data of Authentication System User,"/ M.Pastushenko, Ya.Krasnozheniuk, M. Zaika // International Conference "Problems of Infocommunications. Science and Technology" (PICS&T′2020), 2020, pp. 1-5.

4. Заіка М.В. Особливості обробки фазової інформації голосового сигналу користивуча системи голосової автентифікації / М.С. Пастушенко, Я.О.Красноженюк, М.В. Заіка // Матеріали шостої Міжнародної науково-технічної конференції «Проблеми електромагнітної сумісності перспективних бепроводових мереж зв'язку ЕМС- 2020». – Харків, ХНУРЕ. Том 4. – 2020. - с. 73-75.

5. Заіка М.В. Анализ направлений повышения безопасности голосовой аутентификации пользователей систем доступа / М.С. Пастушенко, Я.О.Красноженюк, М.В. Заіка // Матеріали шостої Міжнародної науково-технічної конференції «Проблеми електромагнітної сумісності перспективних бепроводових мереж зв'язку ЕМС- 2020». – Харків, ХНУРЕ. Том 4. – 2020. - с. 42-44.

6. Биометрическая_идентификация_(мировой_рынок) [Электронный ресурс] – Режим доступа: http://www. tadviser.ru/index.php – Загл. с экрана.

7. Пастушенко О.Н. Анализ качественных показателей биометрических систем аутентификации пользователей. / О.Н. Пастушенко, И.Ш. Невлюдов. // Проблеми телекомуникацый. – 2012. – №4(9). – С.96–103.

8. Иконин С.Ю. Система автоматического распознавания речи SPIRIT ASR Engine / С.Ю. Иконин, Д.В. Сарана // Цифровая обработка сигналов. – 2003. – №4. – С. 5–13.

9. Mariethoz J. Speaker Verification Based on User-Customized Password / J. Mariethoz, B. Herve, M.F. BenZeghiba // IDIAP Research Report 01-13. – Martigny, 2001. – 22 p.

10. Фант Г. Акустическая теория речеобразования / Г. Фант; Пер. с англ. – М.: Наука, 1964. – 284 с.

11. Фланаган Дж.Л. Анализ, синтез и восприятие речи / Дж.Л. Фланаган; пер. с англ. А.А. Пирогова. – М. : Связь, 1968. – 396 с.

12. Pellandini F. GSM Speech Coding And Speaker Recognition / F. Pellandini,

M. Ansorge A. Dufaux [at al.] // International Conference on Acoustics, Speech, and Signal Processing (ICASSP): Book of abstracts. – Istanbul, 2000. – vol. 2. – pp. 1085–1088.

13. Amrouche A. Effect of GSM speech coding on the performance of Speaker Recognition System / A. Amrouche, A. Krobba, M. Debyeche // 10th International Conference on Information Sciences Signal Processing and their Applications (ISSPA): Book of abstracts. – Kuala Lumpur, 2010. – pp. 137–140.

14. Дворянкин С.В. О необходимости новых подходов к оценке эффективности технических средств акустозащиты / С.В. Дворянкин // Информация и безопасность. – 2002. – №2. – С. 244–245.

15. Андреев Б.В. Расследование преступлений в сфере компьютерной информации/ Б.В. Андреев, П.Н. Пак, В.Н. Хорст. – М.: Юрлитинформ, 2001.– 152 с.

16. Вехов В.Б. Компьютерные преступления: способы совершения и раскрытия / В.Б. Вехов. – М. : Право и Закон, 1996. – 182 с.

17. Усов А.И. Судебно-экспертное исследование компьютерных средств и систем: основы методического обеспечения / А.И. Усов. – М. : Изд-во «Экзамен», изд-во «Право и закон», 2003. – 368 с.

18. Сорокин В.Н. Определение пола диктора по голосу / В.Н. Сорокин, И.С. Макаров // Акустический журнал. – 2008. – Т. 54. – № 4. – С. 659–668.

19. Галунов В.И. О возможности определения эмоционального состояния говорящего по речи / В.И. Галунов // Речевые технологии. – 2008. – № 1. – С. 60–66.

20. Ромашкин Ю.Н. Распознавание пола диктора на основе gmm-модели голоса / Ю.Н. Ромашкин, Ю.О. Петров // Речевые технологии. – 2009. – № 2. – С. 31–38.

21. Neustein A. Forensic Speaker Recognition: Law Enforcement and Counter- Terrorism / A. Neustein, H.A. Patil. – New York : Springer, 2012. – 540 p.

22. Каганов А.Ш. Криминалистическая идентификация личности по голо-

су и звучащей речи / А.Ш. Каганов. – М. : Юрлитинформ, 2009. – 291 с.

23. Сорокин В.Н. Фундаментальные исследования речи и прикладные задачи речевых технологий / В.Н. Сорокин // Речевые технологии. – 2008. – № 1. – С. 18–48.

24. Гребнов С.В. Разработка и реализация двухуровневого метода голосового управления на основе скрытых марковских моделей / С.В. Гребнов // Информационные технологии. – 2009. – № 9. – С. 40–46.

25. Кривнова О.Ф. Области применения речевых корпусов и опыт их разработки / О.Ф. Кривнова // Труды xviii сессии Российского акустического общества РАО. – Таганрог, 2006. – С. 81–84.

26. Дшхунян В.Л. Электронная идентификация. Бесконтактные электронные идентификаторы и смарт-карты / В.Л. Дшхунян, В.Ф. Шаньгин. – М. : ООО «Издательство АСТ», Издательство «НТ Пресс», 2004. – 695 с.

27. Pastushenko M. "Specifics of Receiving and Processing Phase Information in Voice Authentication Systems" / M. Pastushenko, V. Pastushenko, O. Pastushenko // International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kyiv, Ukraine. – 2019 – P.621–624.

28. Pastushenko M. Analysis of voice signal phase data informativity of authentication system / M. Pastushenko, Ya. Krasnozheniuk, O. Lemeshko // Zaporizhzhia, Ukraine, April 27-May 1, 2020. Proceedings of The Third International Workshop on Computer Modeling and Intelligent Systems (CMIS-2020). PP 1040-1053

29. Пастушенко М.С. Оцінка частоти основного тону голосового сигналу користувача системи автентифікації. [Електронний ресурс] / Є.Є.Куценко, М.С. Пастушенко. // Проблеми телекомунікацій. – 2019. – № 2(25). – С. 97–103. Режим доступу до ресурсу; https://nure.ua/ru/branch/elektronnoe-nauchnoe-spetsializirovannoe-izdanie-problemyi-telekommunikatsiy

30. Пастушенко М.С. Оцінка частоти основного тону голосового сигналу користувача системи автентифікації / Є. Є.Куценко, М. С. Пастушенко. // Харків, НАНГУ, Міжнародна науково-практична конференція “Застосування інформаційних технологій у підготовці та діяльності сил охорони правопорядку”. Збірник тез доповідей. –2020. – С. 24.

31. Пастушенко Н.С. Экспериментальное исследование информативности амплитудного спектра голосового сигнала для аутентификации пользователя

[Электронный ресурс] / Н.С. Пастушенко, Б.Д. Малонга, О.Н. Файзулаева // Проблеми телекомунікацій. – 2015. – № 2 (17). – С. 3-11.

32. Иванов А.И. Биометрическая идентификация личности по динамике подсознательных движений / А.И. Иванов. – Пенза: Изд-во Пенз.гос.ун-та, 2000. – 187 с.

33. Дворянкин С.В. О практическом применении технологии речевой подписи / С.В. Дворянкин // Конференция «Телекоммуникационные и вычислительные системы» в рамках международного конгресса Коммуникационные технологии и сети» (CTN-2002), международного форума информатизации (МФИ 2002): матер. конф. – М., 2002. – С. 159.

34. Huang Q. Speaker Detection In Broadcast Speech Databases / Q. Huang, S. Parthasarathy, I. Magrin-Chagnolleau, A. E. Rosenberg // International Conference on Spoken Language Processing: Book of abstracts. – Sydney, 1998. – vol. 4. – pp. 1339–1342.

35. Torres-Carrasquillo P. Approaches and applications of audio diarization / P. Torres-Carrasquillo, D. A. Reynolds // International Conference on Acoustics, Speech, and Signal Processing (ICASSP): Book of abstracts. – Philadelphia, 2005.

– vol. 5. – pp. 953–956.

36. Галяшина Е.И. Идентификация дикторов по цифровым фонограммам /Е.И. Галяшина // Речевые технологии. – 2010. – № 3. – С. 23–28.

37. Женило В.Р. Компьютерная фоноскопия / В.Р. Женило. – М. : Изд-во Академия МВД России, 1995. – 206 с.

38. Huang X. Spoken language processing: a guide to theory, algorithm, and system development / X. Huang, A. Acero, H.W. Hon. – New Jersey : Prentice Hall, 2001. – 1008 p.

39. Kim H. Noise-Robust Speaker Recognition Using Subband Likelihoods and Reliable-Feature Selection / H. Kim, M. Ji, S. Kim // ETRI Journal. – 2008. – vol. 30 (1). – pp. 89–100.

40. Козлов А.В. Алгоритм детектирования музыкальных фрагментов в задачах речевой обработки / А.В. Козлов, А.И. Лоханова, К.К. Симончик //«Научно-технические ведомости СПбГПУ».(103),–2010.– № 4 (103). – С. 7–11.

41. Crookes D. Speaker recognition in noisy conditions with limited training data / D. Crookes, J. Ming, N. McLaughlin // 19th European Signal Processing Conference (EUSIPCO 2011): Book of abstracts. – Barcelona, 2011. – pp. 1294–

1298.

42. Reynolds D.A. The effects of handset variability on speaker recognition performance: experiments on the switchboard corpus / D.A. Reynolds // ICASSP- 96. – 1996. – vol. 1. – pp. 113.

43. Scheffer N. Towards noise-robust speaker recognition using probabilistic linear discriminant analysis / N. Scheffer, M. Graciarena, L. Ferrer, L. Burget, Y. Lei //International Conference on Acoustics, Speech, and Signal Processing (ICASSP): Book of abstracts. – Kyoto, 2012. – pp. 4253–4256.

44. The NIST Year 2008 Speaker Recognition Evaluation Plan [Электронный ресурс] – Режим доступа: http://www.itl.nist.gov/iad/mig/tests/sre/2008/ sre08_evalplan_release4.pdf – Загл. с экрана.

45. The NIST Year 2010 Speaker Recognition Evaluation Plan [Электронный ресурс] – Режим доступа:http://www.nist.gov/itl/iad/mig/upload/nist_sre10_eval plan-r6.pdf – Загл. с экрана.

46. Reynolds D.A. Robust Speaker Recognition in Noisy Conditions / D.A. Reynolds, J.R. Glass, T.J. Hazen, J. Ming // Audio, Speech, and Language Processing. july – 2007. – vol. 15 (5). – pp. 1711–1723.

47. Zue V. Speech database development at MIT: TIMIT and beyond / V. Zue, S. Seneff, J. Glass // Speech Communication. – 1990. – vol. 9. – pp. 351–356.

48. Sridharan, S. Robust Speaker Recognition using Microphone Arrays / S. Sridharan, J. Pelecanos, I. McCowan // A Speaker Odyssey – The Speaker Recognition Workshop: Book of abstracts. – Crete, 2001. – pp. 101–106.

49. Reynolds D.A. Channel robust speaker verification via feature mapping / D.A. Reynolds // International Conference on Acoustics, Speech and Signal Processing (ICASSP), Hong Kong, China, 06 Apr – 10 Apr 2003. – 2003. – vol. 2. – pp. 53– 56.

50. Laface P. Channel factors compensation in model and feature domain for speaker recognition / P. Laface, E. Dalmasso, F. Castaldo [at al.] // Odyssey: Speaker and Language Recognition Workshop, San Juan, 28-30 June 2006 .

51. Sridharan S. Modeling session variability in text-independent speaker verification / Sridharan S., Baker B., Vogt R. // 9th European Conference on Speech Communication and Technology, September 4-8 in Lisbon, Portugal. – 2005. – pp. 3117-3120.

52. Крак Ю.В. Система распределённого автоматизированного докумен-

тирования речевых сигналов / Ю.В. Крак, А.С. Загваздин // Речевые технологии. – 2012. – № 2. – С. 43-53.

53. Gauvain J. Feature and score normalization for speaker verification of cellular data / J. Gauvain, C. Barras // International Conference on Acoustics, Speech, and Signal Processing (ICASSP): Book of abstracts. – Hong Kong, 2003. – vol. 2. – pp. 49–52.

54. Morgan N. RASTA processing of speech / H. Hermansky, N. Morgan // Transactions on Speech and Audio Processing. – 1994. – vol. 2 (4). – pp. 578–589.

55. Tiwary U.S. Text independent speaker identification using wavelet transform / U.S. Tiwary, G.K. Verma // International Conference on Computer and Communication Technology (ICCCT): Book of abstracts. – Allahaba, 2010. – pp. 130–134.

56. Osman R. Development of a speaker recognition system using wavelets and artificial neural networks / R. Osman, C.P. Lim, S.C. Woo // International Symposium on Intelligent Multimedia, Video and Speech Processing: Book of abstracts. – Hong Kong, 2001. – pp. 413–416.

57. Ахмад Х.М. Сравнительное исследование эффективности различных методов кепстрального описания речевых сигналов в задачах распознавания / Х.М. Ахмад // Вестник ТГТУ. – 2007. – Том 13, № 4. – С. 887–891.

58. Reynolds D.A. Speaker Verification Using Adapted Gaussian Mixture Models / D. A. Reynolds, T. F. Quatieri, R. B. Dunn // Digital Signal Processing. – 2000. – vol. 10. – pp. 19–41.

59. Tutorial on Text-Independent Speaker Verification / D.A. Reynolds, D. Petrovska-Delacr´ etaz, J. Ortega-Garcá [at al.] // EURASIP Journal on Applied Signal Processing. – 2004. – vol. 4. – pp. 430–451.

60. Zhang Y. Optimization of GMM Training For Speaker Verification / Y. Zhang, M. Scordilis // Odyssey 2004 -- The Speaker and Language Recognition Workshop, Toledo, Spain May 31 -- June 3, 2004. -- Toledo, 2004. -- pp. 231--236.

61. MLLR techniques for speaker recognition / M. Ferras, C. C. Leung, C. Barras, J.-L. Gauvain // Odyssey 2008 -- The Speaker and Language Recognition Workshop, Stellenbosch, South Africa January 21--24, 2008. -- Stellenbosch, 2008.– pp. 21--24.

62. Rabiner L. R. A tutorial on hidden markov models and selected applications in speech recognition / L. R. Rabiner // Proceedings of the IEEE. 1989. – vol. 77

(2). – pp. 257–286.

63. Carey M. Speaker Recognition using a Trajectory-Based Segmental HMM / M. Carey, M. Russell, Y. Liu // Proceedings of Odyssey – Speaker and Language Recognition Workshop: Book of abstracts. – Toledo, 2004. – pp. 45–50.

64. Arslan L.M. HMM-based text-dependent speaker recognition with handset- channel recognition / L.M. Arslan, O. Bvyük // Signal Processing and Communications Applications Conference: Book of abstracts. – Diyarbakir, 2010. – pp. 383–386.

65. Nakagawa S. Text-Independent/Text-Prompted Speaker Recognition by Combining Speaker-Specific GMM with Speaker Adapted Syllable-Based HMM /

S. Nakagawa W. Zhang, M. Takahashi // IEICE Transactions on Information and Systems. March – 2006. – vol. 89 (3). – pp. 1058–1065.

66. Хайкин С. Нейронные сети: полный курс / С. Хайкин; пер. с англ. Н.Н. Куссуль, А.Ю. Шелестова, под ред. Н.Н. Куссуль. – М. : Издательский дом «Вильямс», 2006. – 1104 с.

67. Геппенер В.В. Разработка систем автоматической верификации дикторов с использованием нейронных сетей / В. В. Геппенер, К. К. Симончик // Нейрокомпьютеры: разработка, применение. – 2006. – № 7. – С. 14–23

68. Novakovic J. Speaker identification in smart environments with multilayer perceptron / J. Novakovic // Telecommunications Forum (TELFOR): Book of abstracts. – Belgrade, 2011. – pp. 1418–1421.

69. Вапник В.Н. Теория распознавания образов (статистические проблемы обучения) / В.Н. Вапник, А.Я. Червоненкис. – М. : Издательство «Наука», 1974. – 416 с.

70. Renals S. Speaker verification using sequence discriminant support vector machines / V. Wan, S. Renals // Speech and Audio Processing. – 2005. – vol. 13 (2). – pp. 203-210.

71. Reynolds D.A. Support vector machines using GMM supervectors for speaker verification / D.A. Reynolds, D.E. Sturim, W.M. Campbell // Signal Processing Letters. – 2006. – vol. 13 (5). – pp. 308–311.

72. Заіка М. В. Статистический анализ трафика для обнаружения DDoS-атак / М.В. Заіка, Т. А. Радівілова, М.Х. Тавалбех, Д.Я. Глушаєв // Матеріали 3-ї міжнародної науково-технічної конференції «Комп'ютерні та інформаційні системи та технології», Харків, 2019, с.137.

73. Zaika M. Statistical analysis of traffic protocols to detect DDoS attacks / T. Radivilova, M, Tawalbeh, M. Zaika // Матеріали XIV Міжнародній науково-технічній конференції «АВІА-2019», Київ. Сс.8.15-8.17