

КВАЛІФІКАЦІЙНА РОБОТА

бакалавр  
Освітній рівень

Пристрій та алгоритми обробки мовних команд  
підсистем голосового управління керуючих систем  
Назва теми

КвРКІ. 180243.18.02.16 ПЗ  
Шифр

Галузь знань 12 «Інформаційні технології»

Шифр, назва

Спеціальність 123 «Комп'ютерна інженерія»

Шифр, назва

Освітня програма «Комп'ютерна інженерія»

Назва

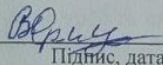
Виконав: студент IV курсу, група KI-18-2



Підпис

Стецюк М.В.

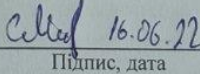
Керівник доц., к. т. н, доцент кафедри Кб



Підпис, дата

Орленко В.С.

Нормоконтролер ст. викладач кафедри Кб

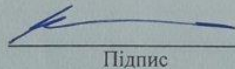
 16.06.22

Підпис, дата

Мостовий С.В.

До захисту допускаю:

Зав. кафедри Кб, к.т.н., доц



Підпис

Кльоц Ю.П.

«16» червня 2022р.

Хмельницький, 2022

Зов.	Поз.	Позначення	Найменування	Кільк.	Прим.
	1		Завдання на кваліфікаційну роботу	1	
	2		Анотація	1	
	3	КвРКІ.180243.18.02.16 ПЗ	Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем	1	
	4	КвРКІ.180243.18.02.16 Е8	Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем	1	
	5	КвРКІ.180243.18.02.16 Е8	Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем	1	

Зем.	Арх.	№ Докум.	Підп.	Дата	КвРКІ.180243.18.02.16 ВП			
розробив		Стецюк М.В.		16.06	Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем	Літера	Аркуш	Аркушів
перев.		Орленко В.С.		16.06		н	1	2
монтр.		Мостовий С.В.		16.06	Відомість проекту	ХНУ зр. КІ-18-2		
ств.		Кльон Ю.П.		16.06				



ХМЕЛЬНИЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

Факультет ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

Кафедра КІБЕРБЕЗПЕКИ

Освітній рівень БАКАЛАВР

Галузь знань 12 ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

Спеціальність 123 КОМП'ЮТЕРНА ІНЖЕНЕРІЯ

Освітня програма ОСВІТНЬО-ПРОФЕСІЙНА ПРОГРАМА ПІДГОТОВКИ БАКАЛАВРІВ

**ЗАТВЕРДЖУЮ**

Завідувач кафедри кібербезпеки

к.т.н. доцент Кльоц Ю.П.

"16" серпня 2022 року

**ЗАВДАННЯ  
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

Стецюку Миколі Валерійовичу

(прізвище, ім'я, по батькові)

1. Тема проекту (роботи) Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем

Науковий керівник Орленко Вікторія Сергіївна, к.т.н., доцент

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджена наказом № ректора університету додаток №18 від 1.05.2022

2. Строк подання студентом проекту (роботи) на кафедру 01.06.2022.



3. Вихідні дані до проекту (роботи) Розробити алгоритми та пристрої обробки мовних команд підсистем голосового управління керуючих систем.

4. Зміст пояснювальної записки (перелік питань, які потрібно розробити)

Аналіз та дослідження систем голосового управління і алгоритмів обробки мовних команд. Розробка алгоритмів і пристроїв обробки мовних команд підсистеми голосового управління керуючих систем. Реалізація підсистеми голосового управління керуючих систем.

5. Перелік графічного матеріалу (із зазначенням обов'язкових креслень) 1. Технологічний процес обробки мовних команд (E8) 2. Алгоритм фільтрації нестационарних фонових шумів (E8) 3. Алгоритм сегментації сигнал/пауза (E8) 4. Схема електрична функціональна (E2) 5. Схема електрична принципова (E3)

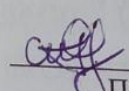
6. Консультанти розділів дипломного проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання виконав
Нормоконтроль	Мостовий С.В., ст. викладач кафедри КБКСМ	-	
Антиплагіат	Мостовий С.В., ст. викладач кафедри КБКСМ	-	

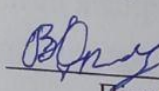
7. Дата видачі завдання: « 1 » лютого 2022р.

**КАЛЕНДАРНИЙ ПЛАН**

№ з/п	Назва етапів дипломної роботи	Строк виконання етапів роботи	Проміжок виконання
1	Грунтовне ознайомлення з предметною галуззю	Лютий - 1 декада	Виконано
2	Визначення структури кваліфікаційної роботи	Лютий - 2 декада	Виконано
3	Робота над першим розділом роботи	Березень - 1 декада	Виконано
4	Робота над другим розділом роботи	Березень - 2 декада	Виконано
5	Робота над третім розділом роботи	Березень - 3 декада	Виконано
6	Оформлення графічного матеріалу	Травень - 1 декада	Виконано
7	Оформлення пояснювальної записки	Травень - 2 декада	Виконано
8	Попередній захист кваліфікаційної роботи	Травень - 2 декада	Виконано
9	Доопрацювання кваліфікаційної роботи	Травень - 3 декада	Виконано
10	Подання роботи для перевірки на плагіат	Травень - 3 декада	Виконано
11	Захист кваліфікаційної роботи	Червень - 1 декада	Виконано

Студент  Підпис

М.В. Стецюк  
Ініціали, прізвище

Керівник проекту (роботи)  Підпис

В.С. Орленко  
Ініціали, прізвище

## АНОТАЦІЯ

Тема дипломної роботи: «Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем».

Автор роботи: студент групи КІ – 18 – 2 Стецюк М.В.

Керівник роботи: ктн. доц. Орленко В.С.


Пояснювальна записка: 64 с., 27 рисунків, 7 таблиць, 3 дод., 18 джерел.

Графічна частина: 5 плакатів.

Перелік ключових слів: мовна команда, голосове управління, фоновий шум, частота основного тону, сегментація, фільтрація.

В першому розділі проведено аналіз та дослідження розв'язуваних завдань та стану предметної області, огляд сучасного стану систем голосового керування. Досліджено особливості обробки мовних команд у підсистемах голосового управління керуючих систем, проведено аналіз методів обробки мовних команд у підсистемі голосового управління керуючих систем.

У кваліфікаційній роботі вирішенні наступні задачі: запропоновано підхід до обробки мовних сигналів підсистеми голосового управління керуючих систем, заснований на алгоритмах і пристроях обробки та оптимізовану послідовність дій: фільтрацію нестационарних фонових шумів, сегментацію сигнал/пауза, вимірювання чистот основного тону. Розроблено структурну та електричну схеми з'єднання підсистеми голосового управління керуючої системи домашньої автоматизації за бездротовим протоколом передачі даних Z-Wave. Апаратно реалізована підсистема голосового управління керуючої системи на базі модуля обробки – модуля візуалізації – TFT – дисплей 2.8" USB TFT display for Raspberry Pi; мінікомп'ютера Raspberry Pi Model B; модуля реєстрації - мікрофона Міні USB 2.0 мікрофон MIC; модуля аудіо - динамічної головки 26CR08FE; модуля живлення - AC/DC конвертора TRACO POWER TML 05105.

  
Підпис студента

16.06.22  
Дата

## ЗМІСТ

	стор.
ВСТУП .....	4
АНАЛІЗ ТА ДОСЛІДЖЕННЯ СИСТЕМ ГОЛОСОВОГО УПРАВЛІННЯ І АЛГОРИТМІВ ОБРОБКИ МОВНИХ КОМАНД .....	6
1.1 Аналіз предметної області сучасного стану систем голосового управління .....	6
1.2 Аналіз методів обробки мовних команд у підсистемі голосового управління керуючих систем .....	12
1.3 Дослідження особливостей обробки мовних команд у підсистемі голосового управління керуючих систем .....	19
1.4 Постановка задачі .....	24
РОЗРОБКА АЛГОРИТМІВ І ПРИСТРОЇВ ОБРОБКИ МОВНИХ КОМАНД ПІДСИСТЕМИ ГОЛОСОВОГО УПРАВЛІННЯ КЕРУЮЧИХ СИСТЕМ .....	25
2.1 Технологічний процес підсистеми голосового управління обробки мовних команд .....	25
2.2 Алгоритм та пристрій фільтрації нестационарних фонових шумів на основі дослідження шумових та енергетичних характеристик .....	27
2.3 Алгоритм та пристрій вимірювання частоти основного тону на основі дослідження частотно-часових характеристик .....	32
2.4 Алгоритм та пристрій сегментації сигнал/пауза на основі дослідження логарифму енергетичних характеристик .....	37
2.5 Висновки.....	43
3 РЕАЛІЗАЦІЯ ПІДСИСТЕМИ ГОЛОСОВОГО УПРАВЛІННЯ КЕРУЮЧИХ СИСТЕМ .....	44
3.1 Підсистема голосового управління у складі керуючих систем .....	44

<i>КвРКІ.180243.18.02.16 ПЗ</i>									
Зм.	Арк.	№ докум.	Підпис	Дата	Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем Пояснювальна записка	Літера	Аркуш	Аркушів	
Розробив		Стешок М.В.	<i>Стешок</i>	16.06		н		2	68
Перевірив		Орленко В.С.	<i>Орленко</i>	16.06					
Н. контр.		Мостовий С.В.	<i>Мостовий</i>	16.06					
Затвердив		Кльоц Ю.П.	<i>Кльоц</i>	16.06					
						<i>ХНУ зр. КІ-18-2</i>			

3.2 Апаратна реалізація підсистеми голосового управління керуючої системи .....	50
3.3 Висновки .....	60
ВИСНОВКИ .....	62
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	63
ДОДАТОК А Лістинг програми «Визначення частоти основного тону мовного сигналу».....	65
ДОДАТОК Б Лістинг програми «Визначення перешкод у мовних сигналах»...	66
ДОДАТОК В Копія графічної частини.....	68

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						3
Зм	Арк	№ докум.	Підпис	Дата		

## ВСТУП

Процес взаємодії з електронно-обчислювальною машиною тривалий час був доступний лише фахівцям, які спілкувалися з комп'ютером у вигляді діалогового інтерфейсу. Використання графічного інтерфейсу призвело до поширення персональних комп'ютерів, однак людина завжди прагнула більш природного і універсального способу взаємодії з електронно-обчислювальною машиною - голосового управління. Як інтерфейс взаємодії використання голосового управління набуло широкої популярності і в керуючих системах. Голосове управління засноване на використанні технології розпізнавання мовлення, що є обробкою надходжених мовних сигналів з метою придушення та фільтрації шуму, визначення інформативних параметрів, сегментації на інформативні ділянки, що відображають унікальні відповідні властивості мови та подальшого розпізнавання. Коректне розпізнавання мови - одне з найскладніших і найцікавіших задач голосового управління.

Дослідження систем голосового управління, показало що вони вирішують безліч прикладних задач у області обробки мовних сигналів. Проте це не знижує інтерес у напрямку обробки мовних сигналів.

На теперішній час існує достатня кількість підсистем голосового управління. Для управління комп'ютеризованою технікою 95 процентів систем голосового управління представлені програмними продуктами, і лише 3% з них це пристрої управління високоінтелектуальними комплексами. Серед виробників систем голосового управління можна відзначити наступні зарубіжні компанії Apple, Google, Home Sapiens, Linguatex, Microsoft, Nuance.

Найбільший інтерес у систем голосового управління представляють пристрої та алгоритми обробки мовних сигналів. Традиційні методи обробки нестационарних та нелінійних сигналів, через їх низьку ефективність, стають малоефективними для використання існуючих пристроїв та алгоритмів у спеціалізованих системах голосового управління. Підвищення ефективності систем голосового управління вимагає на сьогодні, вдосконалення існуючих пристроїв, елементів та алгоритмів

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						4
Зм	Арк	№ докум.	Підпис	Дата		

керуючих систем. Це надає актуальності задачі розробки пристроїв та алгоритмів обробки мовних сигналів, які в свою чергу підвищують ефективність сегментації, фільтрації, розпізнавання мовних сигналів, виділення інформативних параметрів системи голосового управління.

Метою дослідження є розробка пристроїв та алгоритмів обробки нестационарних та нелінійних мовних сигналів на основі методу множинної декомпозиції на емпіричні моди, що дозволить підвищити ефективність підсистем голосового управління керуючої системи.

Для досягнення поставленої мети необхідно провести: аналіз та дослідження предметної області; огляд сучасного стану систем голосового керування; дослідження особливості обробки мовних команд у підсистемах голосового управління керуючих систем; аналіз методів обробки мовних команд у підсистемі голосового управління керуючих систем.

У роботі використовувалися алгоритми цифрової обробки мовних сигналів, статистичного та спектрального аналізу. Для проведення аналізу та досліджень, відображення результатів були використані середовище програмування Microsoft Excel, Matlab, LabVIEW, C#.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						5
Зм	Арк	№ докум.	Підпис	Дата		

# 1 АНАЛІЗ ТА ДОСЛІДЖЕННЯ СИСТЕМ ГОЛОСОВОГО УПРАВЛІННЯ І АЛГОРИТМІВ ОБРОБКИ МОВНИХ КОМАНД

## 1.1 Аналіз предметної області сучасного стану систем голосового управління

Мовний сигнал - акустичний сигнал, який утворюється артикуляційним відділом мовного апарату людини з метою мовного спілкування (обміну інформацією), за допомогою використання мовних конструкцій, створюваних відповідно до певних правил [2]. Фонема є мінімальною сенсорною одиницею мовного сигналу, яка служить для ототожнення та розрізнення слів і звуків [3, 4].

Мовні сигнали є нелінійними та нестационарними сигналами складної форми, характеристики та параметри яких швидко змінюються протягом короткого проміжку часу. На рис. 1.1 наведено приклад слова «Наперсток» - мовний сигнал.

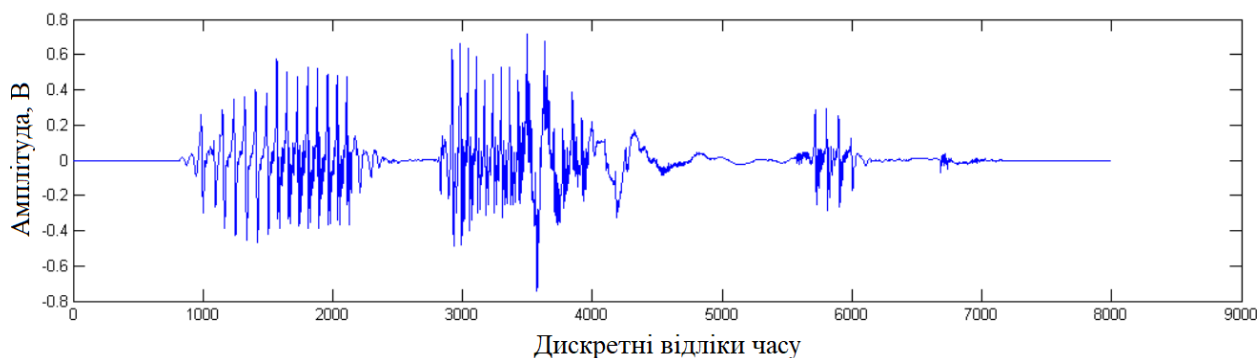


Рисунок 1.1 - Слово «Наперсток» – мовний сигнал

Мовний сигнал, як будь-який вимірювальний сигнал, характеризується набором вимірювальних параметрів: частота основного тону, форманти, звуковий тиск [6, 7]. Частота основного тону - при вимові вокалізованої мови, частота коливань голосових зв'язок. Особливістю частоти основного тону - інтервали безперервно змінюються у значних межах проходження імпульсів голосових зв'язок. У більшість випадків тривалість вокалізованих ділянок мови проходження імпульсів невелика, значну частину займають, при цьому, перехідні процеси, тому виміряти точне значення частоти основного тону є трудомісткою задачею. Форманти представляють частотні області мовного сигналу максимальної концентрації енергії

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		6

[4]. Форманта це область, у якій посилюється декілька гармонік основного тону, внаслідок резонансу, виробленого голосовими зв'язками людини. Форманта, у спектрі звуку чітко виділяється областю посилених гармонік частот основного тону. Форманти характеризуються шириною смуги частот, амплітудою та частотою. При проведенні дослідження мовних сигналів за амплітуду форманти приймають відповідно рівень чи інтенсивність форманти. Ширина смуги частот приймає значення критичної смуги слуху людини або з деякою допустимою помилкою дорівнює смузі частот формантою області.

Звуковий тиск - надлишковий тиск, виникає, при проходженні через неї звукової хвилі, в пружному середовищі. В точці середовища значення звукового тиску змінюється при переході до інших точок середовища так і за часом, практичний інтерес представляє середньоквадратичне значення звукового тиску, пов'язане з інтенсивністю звуку.

Рівень звукового тиску - вимірне значення звукового тиску за відносною шкалою, відповідно до опорного тиску  $P_{SPL} = 20 \text{ мкПа}$ , відповідає порогу з частотою 1000 Гц синусоїдальної звукової хвилі.

Результати проведених дослідження [7-10] показали, що коливання голосових зв'язок людини характеризуються нерегулярністю, яка проявляється у вигляді флуктуацій сусідніх періодів тону та у значних змінах тривалості періодів основного тону (10% – 30%). Нерегулярності виникають в результаті неповного змикання голосових зв'язок людини в кінці і на початку тональних ділянок, при цьому періоди з повним змиканням чергуються з неповним змиканням голосових зв'язок.

Наведена інформація про мовні сигнали людини та їх параметри дозволяє зробити наступний висновок: процес управління інформаційним об'єктом з використанням мовних сигналів є важливим і складним, що необхідно враховувати при розробці пристроїв та алгоритмів обробки мовних сигналів (команд) для систем голосового управління.

Сучасні системи голосового управління - складні технічні комплекси, призначені для управління процесами різного призначення та об'єктами [12, 13]. Будь-яку керуючу систему, структурно можна представити у вигляді

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						7
Зм	Арк	№ докум.	Підпис	Дата		

взаємопов'язаної сукупності керуючого органу (керуючої підсистеми) і об'єкта управління (керованої підсистеми) за допомогою використання зворотного зв'язку та керуючого впливу (рис. 1.2) [11, 12].

У зв'язку з успіхами на сучасному етапі розвитку інформаційних технологій та обчислювальної техніки визначилася тенденція до наростання складності керуючих систем, а також всіх інших видів інформаційних систем взаємодії «людина-машина». Таким чином, Важливим є можливість взаємодії людини з керуючою системою мовою спілкування, максимально наближеною до природної мови. Застосування системи голосового управління в інформаційно-керуючих системах, дозволяє організувати зручну та ефективну взаємодію користувача з керуючою системою, як інтерфейс взаємодії «людина-машина» [13, 14].

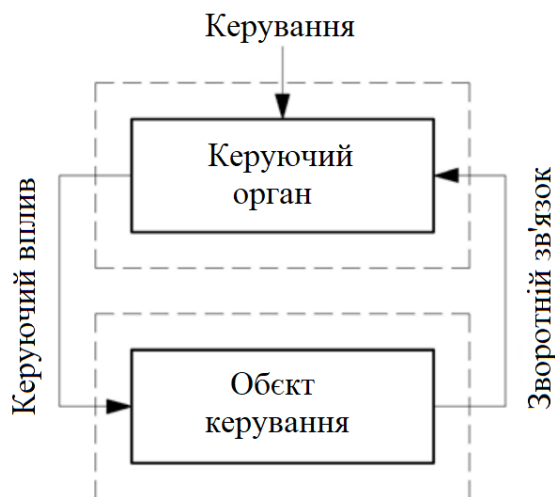


Рисунок 1.2 – Структурна схема керуючої інформаційної системи

Голосове управління - спосіб взаємодії за допомогою мовних команд, з керуючою системою. Задача зводиться до розпізнавання, проте на відміну від інформаційних систем розпізнавання мови, системи голосового управління призначені для розпізнавання окремих керуючих мовних команд. На рис. 1.3 наведено функціональну схему системи голосового управління, складається з функціональних вузлів [14, 15]: блоку обробки, блоку розпізнавання та мікрофона. З урахуванням апаратного та програмного забезпечення внутрішній пристрій блоків системи голосового управління набагато складніший. Мова фіксується мікрофоном у вигляді звукових хвиль, який перетворює звукові хвилі на аналоговий мовний

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						8
Зм	Арк	№ докум.	Підпис	Дата		

сигнал. Блоці обробки перетворює аналоговий сигнал в цифровий вигляд, сегментується на фрагменти, проходить фільтрацію, відбувається виділення інформативних параметрів для подальшого проведення аналізу. Блок розпізнавання порівнює мовну команду, що надійшла в керуючу систему, з шаблоном з бази даних, які формуються під час проведення навчання системи. База даних шаблонів формується на етапі проектування та розробки системи, може коригуватися і доповнюватися в процесі експлуатації та налаштуванні під конкретного користувача. Процес реалізації бази даних проходить в інтерактивному режимі.

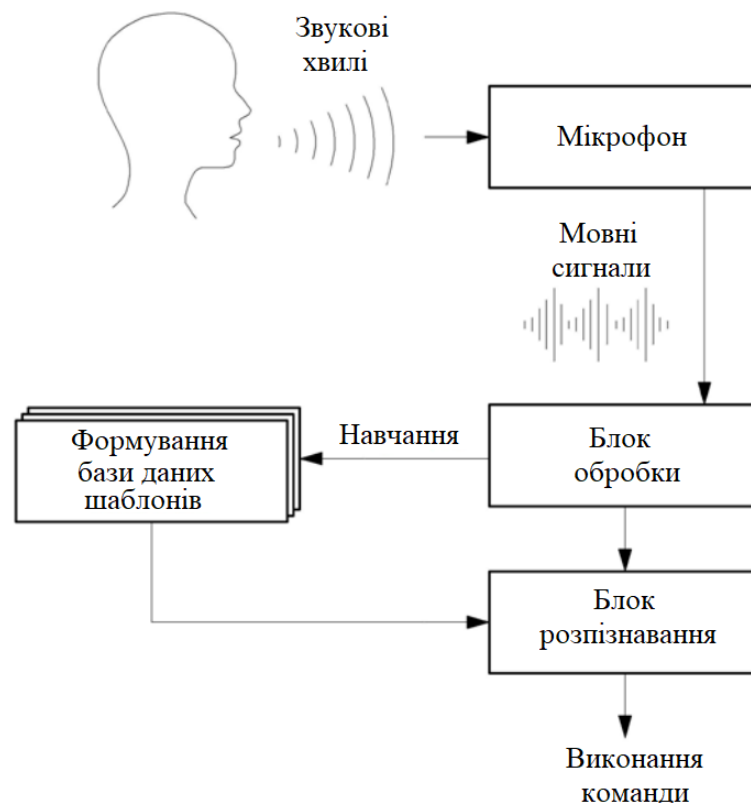


Рисунок 1.3 - Функціональна схема системи голосового управління

Окремо розпізнавання мовних команд більш простіше, ніж розпізнавання мовних сигналів злитого мовлення і вимагає, при цьому великих обчислювальних затрат. На сьогоднішній день, з цієї причини існує, величезний вибір апаратного та програмного забезпечення, що має, при цьому низьку вартість при задовільній якості розпізнавання. Проведене дослідження показує, що системи розпізнавання не подолали рівень точності у 86%, у людини даний показник - 97% [12]. Для проведення оцінки сучасного стану систем голосового управління, на рис.1.4

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						9
Зм	Арк	№ докум.	Підпис	Дата		

наведено за основними параметрами класифікацію: призначення, технічне виконання, тип мовлення, персоналізація, тип структурної одиниці, розмір словника.

Системи голосового управління можна класифікувати за параметрами технічної реалізації, за споживчими властивості систем. Системи голосового управління з технічною реалізацією включають програмно-апаратні засоби і програмні продукти.

Призначення голосової системи визначає якість розпізнавання, яку можливо змінювати з використанням додаткових засобів обробки. Системи голосового управління поділяються на три види: системи розпізнавання, системи диктування, командні системи. Персоналізації системи голосового управління діляться на дикторонезалежні та дикторозалежні [12]. Перші призначені для роботи з будь-яким користувачем другі для роботи з одним користувачем.

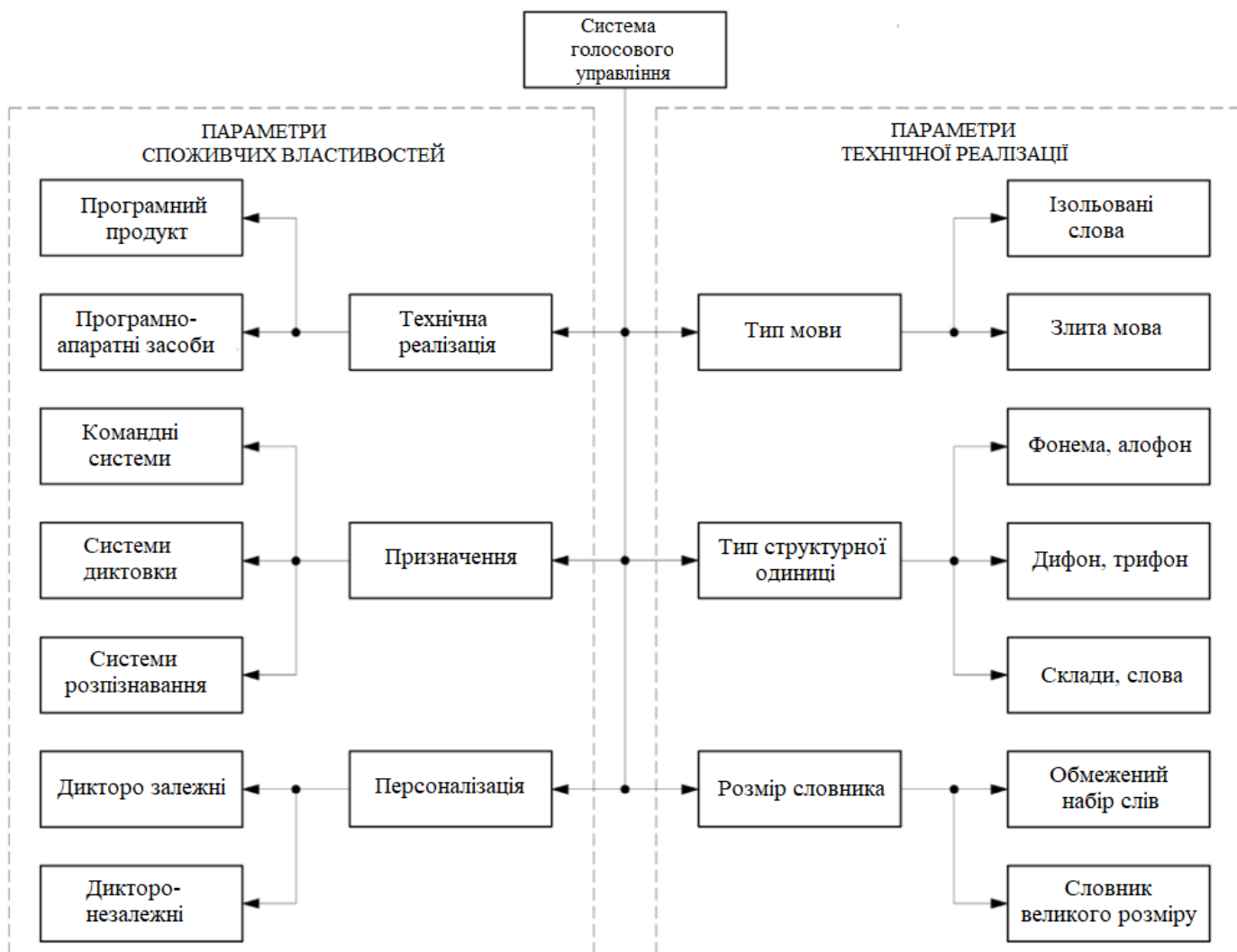


Рисунок 1.4 - Класифікація систем голосового управління

За типом мови системи голосового розпізнавання поділяються на системи: розпізнають зливу мову; розпізнають роздільну мову [13].

Розмір словника голосових систем - кількість слів, яку може розпізнати система. Розмір словника впливає на точність розпізнавання, чим більший розмір, тим більша ймовірність появи помилок. Якщо словник системи складається з десяти тисяч слів, ймовірність помилки до 45%, якщо з десяти слів, розпізнавання практично без помилок. Голосові системи розпізнавання поділяються на системи зі словником великого розміру та з обмеженим словником [15].

При розпізнаванні мовних сигналів можуть бути обрані частини слів, трифони, фонемі, дифони, алофони чи окремі слова [14]. На теперішній час існує достатньо систем голосового управління. 96% з них є програмними продуктами для управління планшетним і персональним комп'ютером, смартфоном, комунікатором: "Ttype Premium"; «Dragon Naturally Speaking»; "IBM ViaVoice"; "Linguatec Voice Pro"; "Voice Commander"; "Диктограф"; модуль розпізнавання мови ОС Windows [32]; модуль розпізнавання мови ОС Mac OS; модуль розпізнавання мови мобільної ОС Android «S-Voice» фірм Samsung та Google; модуль розпізнавання мобільної ОС iOS "Siri" фірми Apple. 4% голосових систем є пристрої голосового управління, вбудовані в інтелектуальні комплекси: система голосового управління «Розумним» будинком фірми Note Sapiens; система голосового управління «Розумним» будинком фірми Insyte; модуль розпізнавання мовних сигналів у телевізорах SMART фірми Samsung.

Серед широкого набору технічних можливостей голосових керуючих систем інтерес представляють пристрої та алгоритми обробки, які використовують керуючі системи. Виробники не надають відповідної інформації, у силу комерційної таємниці.

Точність розпізнавання є основним параметром систем голосового управління - безрозмірна величина, дорівнює відношенню безпомилко розпізнаних мовних сигналів до загальної кількості мовних команд. Точність розпізнавання, як показали дослідження, потребує підвищення. Практично всі системи голосового управління працюють за наявності доступу до інтернет мережі, що також, в деяких ситуаціях, є

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		11

недоліком. Організація роботи систем голосового управління без доступу до мережі Інтернет можлива лише за обмеженням розміру словника та за наявності обчислювальних потужностей.

Таким чином, підвищення ефективності роботи систем голосового управління може досягнуто за рахунок розробки відповідних пристроїв та алгоритмів обробки мовних команд, які забезпечать, в даному випадку, підвищення точності розпізнавання, виділення інформативних ознак, сегментації, фільтрації, що дозволить знизити рівень помилок, підвищити точність розпізнавання, відкрити нові області ефективного застосування систем голосового управління.

## 1.2 Аналіз методів обробки мовних команд у підсистемі голосового управління керуючих систем

Аналіз методів обробки мовних команд, на сучасному етапі, у системах голосового управління є важливим етапом дослідження у кваліфікаційній роботі. На рис. 1.5 наведена класифікація використовуваних методів обробки в системах голосового управління керуючих систем [15].

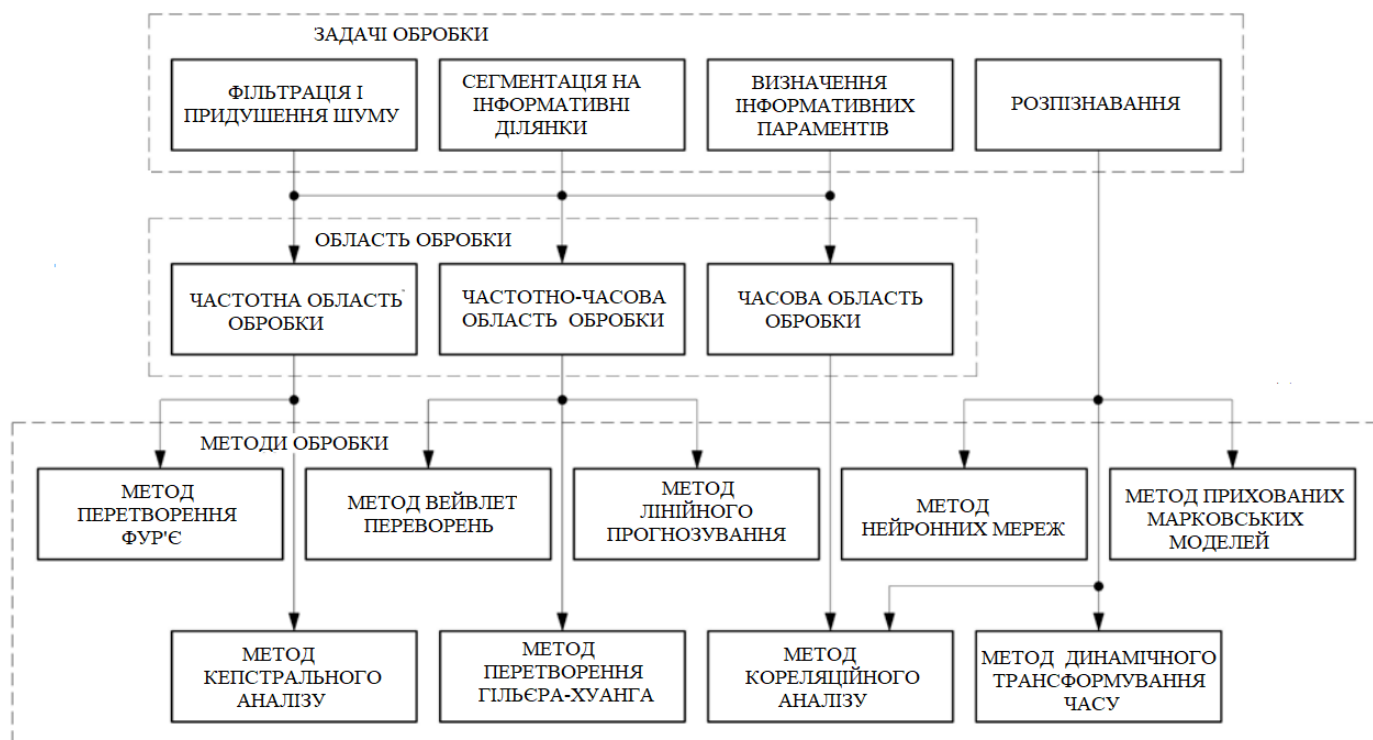


Рисунок 1.5 - Класифікація методів обробки мовних команд в СГУ

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						12
Зм	Арк	№ докум.	Підпис	Дата		

Наведена на рис.1.5 класифікація методів обробки мовних команд виконана з урахуванням задач обробки сигналів в підсистемах голосового управління: розпізнавання, визначення інформативних параметрів, придушення шуму та фільтрація, сегментація на інформативні ділянки.

Відповідну задачу обробки мовних команд можливо реалізувати з використанням конкретних методів. Залежно від області обробки мовних команд, методи поділяються на три групи: часова, частотна, частотно-часова.

Часовий методи обробки мовних команду полягає у визначенні характерних точок мовної команди з подальшим використанням для аналізу. Як характерні точки, з погляду технічної реалізації, можуть бути моменти перетину нульової осі функцією сигналу, мінімуми (максимуми). Недолік методів обробки мовних сигналів у часовій області, в виділенні характерних точок полягає в неоднозначності, викликаній шумами сигналу та усуненнями нульового рівня.

Частотний методи обробки мовного сигналу полягає у використанні зареєстрованих відліків даних у мовній команді. Деякі мовні команди займають характерні спектральні області, мають специфічний частотний склад. Використання частотних методів дозволяє обробляти мовні команди з високою точністю. Недолік обробки в частотній області - порівняно високі обчислювальні витрати, недостатньо висока спектральна роздільна здатність, низька адаптивність до локальних властивостей мовних сигналів.

Частотно-часові методи обробки мовних сигналів включають переваги частотного і часового аналізів з мінімальними недоліками.

Проведений аналіз сучасних методів обробки мовних команд показав, що залежно від обробки сигналів, методи розподіляються на групи, в залежності від виду аналізу: метод вейвлет перетворення; метод кепстрального аналізу; метод перетворення Фур'є; метод лінійного передбачення; метод перетворення Гільберта-Хуанга; метод кореляційного аналізу; метод динамічного трансформування часу; метод нейронних мереж; метод прихованих марківських моделей.

Метод перетворення сигналів Фур'є використовується в багатьох областях так і в мовних технологіях. Обробка мовних сигналів з використанням перетворення

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						13
Зм	Арк	№ докум.	Підпис	Дата		

Фур'є розглядається як перетворення мовного сигналу з часової в частотну область і розкладанням команди на частотні складові. Для цифрової обробки сигналів використовують дискретне перетворення Фур'є, мовний сигнал часто представляють як суму гармонійних складових у дискретному вигляді. Побудова спектра, на основі використання дискретного перетворення Фур'є, дозволяє наочно та компактно подати необхідну інформацію про мовну команду. У спектральному аналізі неможливо детально провести обробку короткочасних локальних особливостей, є недоліком дискретне перетворення Фур'є.

Більшість задач обробки мовних команд реалізуються з використанням методу вейлет перетворення. Вейвлет - добре локалізована функція (зосереджена в невеликій області деякої точки і спадає до нуля в міру віддалення) у частотній і часовій області. До вейвлет перетворення застосовуються дві операції: масштабування (стиснення або розтяг, переміщення області локалізації за частотою), зсув (переміщення області локалізації в часі).

Сутність вейвлет перетворення полягає в розбиття мовного сигналу на зсунуті та масштабовані по осі часу вейвлета та обчислення коефіцієнтів кореляції ділянок мовного сигналу на заданому масштабі. Результатом є набір коефіцієнтів, які показують, наскільки поведінка мовного сигналу відповідає поведінці вейвлета в даному масштабі. Коефіцієнти показують близькість мовного сигналу до вейвлета даного масштабу. Чим ближче мовний сигнал даного часу до виду вейвлета, тим більше значення має коефіцієнт.

Використання масштабування і зсуву в частотно-часовій області надає можливість проводити аналіз мовних команд використовуючи різні масштаби і точно визначати характерних особливостей у часі. Найчастіше, на практиці, зустрічаються вейвлет перетворення функції обробки мовних сигналів: вейвлет Добеши, вейвлет Хаара, вейвлет Марлета, вейвлет «Мексиканський капелюх». Метод вейвлет перетворень має переваги в порівнянні з Фур'є перетворенням – надає можливість проводити аналіз короткочасних локальних особливостей мовних сигналів, короткі провали, сплески, розриви.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						14
Зм	Арк	№ докум.	Підпис	Дата		

Для вирішення проблеми адаптивності використовується метод перетворення Гільберта-Хуанга. Перевага методу Гільберта-Хуанга - висока адаптивність, базисні функції, що використовуються при розкладанні мовного сигналу, витягуються безпосередньо з команди і дозволяють, при цьому враховувати тільки властиві сигналу особливості. Як результат використання методу Гільберта-Хуанга мовний сигнал подається в частотно-енергетично-часовій області, дозволяє виявити області концентрації енергії та приховані модуляції, які дозволяють проводити аналіз як локальних так і глобальних властивостей мовних сигналів, вимагають, при цьому менших обчислювальних витрат.

В області обробки мовних сигналів метод кепстрального аналізу набув широкої популярності, завдяки перевагою стиснення інформації мовного сигналу під час переходу в частотну область обробки. При перетворенні мовного сигналу з часової в частотну область, отримана інформація є більш компактною, наочною, докладною. З'явилася, таким чином можливість представити спектральну інформацію компактніше, коли гармонійний ряд спектра буде представлений однією складовою в кепстрі. Загальноприйнято вважати, кепстр це спектр логарифма діапазону вихідного мовного сигналу, первісний спектр представлений у логарифмічному масштабі.

В основі кепстрального аналізу, в задачах обробки мовних команд, лежить виділення кепстральних коефіцієнтів на мел-шкалі. Метод отримання кепстральних коефіцієнтів заснований на моделі органів слуху людини, моделює частотну чутливість людського вуха, використовує частотну шкалу в мелах.

Метод лінійного прогнозування - найбільш використовуваний метод у задачах обробки мовних команд.

На теперішій час існує два основних підходи до визначення коефіцієнтів лінійного прогнозування: коваріаційний та автокореляційний. Дані підходи використовують представлення сигналу у часовій області. Отриманні коефіцієнти передбачення визначають характеристику фільтра частотну, яка характеризує у певний момент часу стан голосового тракту. Таким чином, з одного боку, точність залежить від стаціонарності досліджуваного мовного сигналу, з іншого даний

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						15
Зм	Арк	№ докум.	Підпис	Дата		

момент може бути локалізований. Дані піходи обчислення коефіцієнтів забезпечують отримання середньої оцінки ділянки мовного сигналу частотно-часової області.

Для визначення взаємозв'язку декількох статичних величин застосовується метод кореляційного аналізу. Мірою кореляції статичних величин служить коефіцієнт кореляції. Метод кореляційного аналізу статистичних даних при обробці мовних команд досить популярний. Популярність обумовлена наступними моментами: коефіцієнти кореляції при підрахунку прості, застосування методу не потребує математичної спеціальної підготовки. Ключовими поняттями, відносно задач обробки мовних команд кореляційного аналізу є взаємно-кореляційна та автокореляційна функції.

Взаємнокореляційна функція визначає відповідну ступінь кореляції послідовностей значень мовних команд, розкладених у ряди, взятих зі зсувом.

Автокореляційна функція визначає статистичний взаємозв'язок між значенням величин з одної мовної команди, розкладеної в ряд, взятої зі зсувом.

Найбільш ефективний метод розпізнавання мовних команд - метод з використанням нейронних мереж. Мережа складаються з нейронів і зв'язками організованими між ними. Нейрон є клітиною нейронної мережі, який може бути в двох станах: загальмованості та збудження. Нейрони звані між собою синапсами - односпрямовані зв'язки, аксонами - вихідні зв'язки нейрона, сигнали гальмування чи збудження надходять на синапси нейронів. На рис. 1.6 наведено загальний вигляд нейрона мережі. Односпрямований зв'язок характеризується вагою  $w$  - величина синаптичного зв'язку, (еквівалентний електричної провідності). Негативні та позитивні значення ваг відповідають загальмованому чи збудженому стану синапсів. Сума входів визначає поточний стан нейрона мережі. При використанні нейронної мережі необхідно побудувати відповідну мережу в задачі розпізнавання мовних команд. Також необхідно навчити її мовних сигналів - підібрати відповідні вагові коефіцієнти синапсів задля мінімізації кількості помилок.

Найбільш ефективний метод розпізнавання мовних сигналів - метод прихованих марківських моделей – статистична модель, імітує процес, аналогічний

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						16
Зм	Арк	№ докум.	Підпис	Дата		

марківському з невідомими параметрами. При використанні методу прихованих марківських моделей головною задачею визначення невідомих параметрів. Отримані, таким чином, параметри використовуються при подальшому аналізі, для розпізнавання образів.

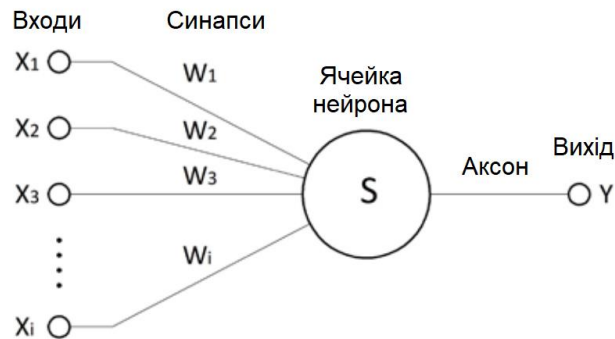


Рисунок 1.6 – Загальний вигляд нейрона нейронної мережі

Метод прихованих марківських моделей при розпізнаванні ґрунтується на наступних припущеннях: мовний сигнал сегментований на стани (фрагменти), усередині яких мовний сигнал розглядається як стаціонарний; ймовірність появи сигналу, що породжується методом, не залежить від попередніх породжених сигналів і залежить від поточного стану моделі.

На даний час існує кілька типів прихованих марківських моделей, різняться за топологією. На рис. 1.7 представлена топологія системи з трьома станами. Система є кінцевим автоматом, змінює стан у кожен дискретний момент. Перехід зі стану  $S_i$  в стан  $S_j$  з ймовірністю  $a_{ij}$  здійснюється випадковим чином. Модель породжує вектор спостережень  $O_n$  з ймовірністю  $b_j(O_n)$  кожного дискретного моменту часу.

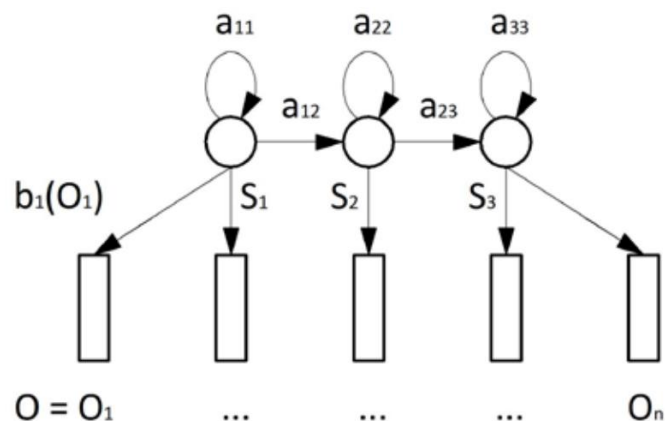


Рисунок 1.7 – Топологія прихованої марківської моделі із трьома станами.

Для того щоб отримати оцінку розбіжності між мовними сигналами, представленими, в свою чергу як вектори, необхідно виконати вирівнювання за часом, що можливо реалізувати з використанням методу динамічного трансформування часу. Метод динамічного трансформування - еластичне порівняння із шаблоном вектора спостережень. Шаблон та вектор спостережень лежать на осях сітки (рис. 1.8). Для кожної комірки сітки вираховується різниця між фрагментами шаблону та вектора спостережень. Оптимальне вирівнювання між шаблоном та вектором спостережень показано маршрутом на сітці.

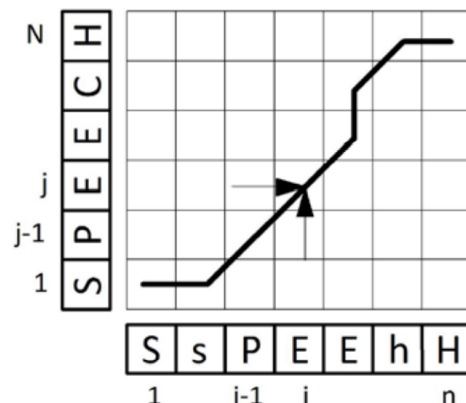


Рисунок 1.8 - Метод динамічного програмування

Метод працює із фрагментами, проведений аналіз ознак складається з обробки у регулярних інтервалах вектора ознак. Вектор ознак може мати множину фрагментів, то необхідно втконати відповідні розрахунки локальної оцінки відстані. Застосування для обчислення Евклідової відстані забезпечує найкращі результати розпізнавання. Результуюча оцінка розбіжності маршруту –сума локальних відстаней між фрагментами шаблону та сигналу.

Наведена класифікація медодів демонструє їх різноманіття, недоліки та переваги видів аналізу для вирішення задач обробки мовних сигналів в системах голосового управління.

### 1.3 Дослідження особливостей обробки мовних команд у підсистемі голосового управління керуючих систем

Обробка мовних команд - дослідження, в якому здійснюється відновлення та стиснення мовних сигналів, видалення інформації, придушення шумів та фільтрація, посилення, розподіл інформаційних потоків, кодування. [12]. Підсистеми голосового управління керуючих систем обробка мовних команд зводиться до проведення аналізу аналізу словосполучення чи окремого слова з чітко вираженими границями, включає наступні задачі: розпізнавання; реєстрація; фільтрація; попередня корекція; визначення інформативних параметрів; сегментація сигнал/пауза. На рис. 1.9 наведені етапи обробки мовних сигналів у системі голосового управління.

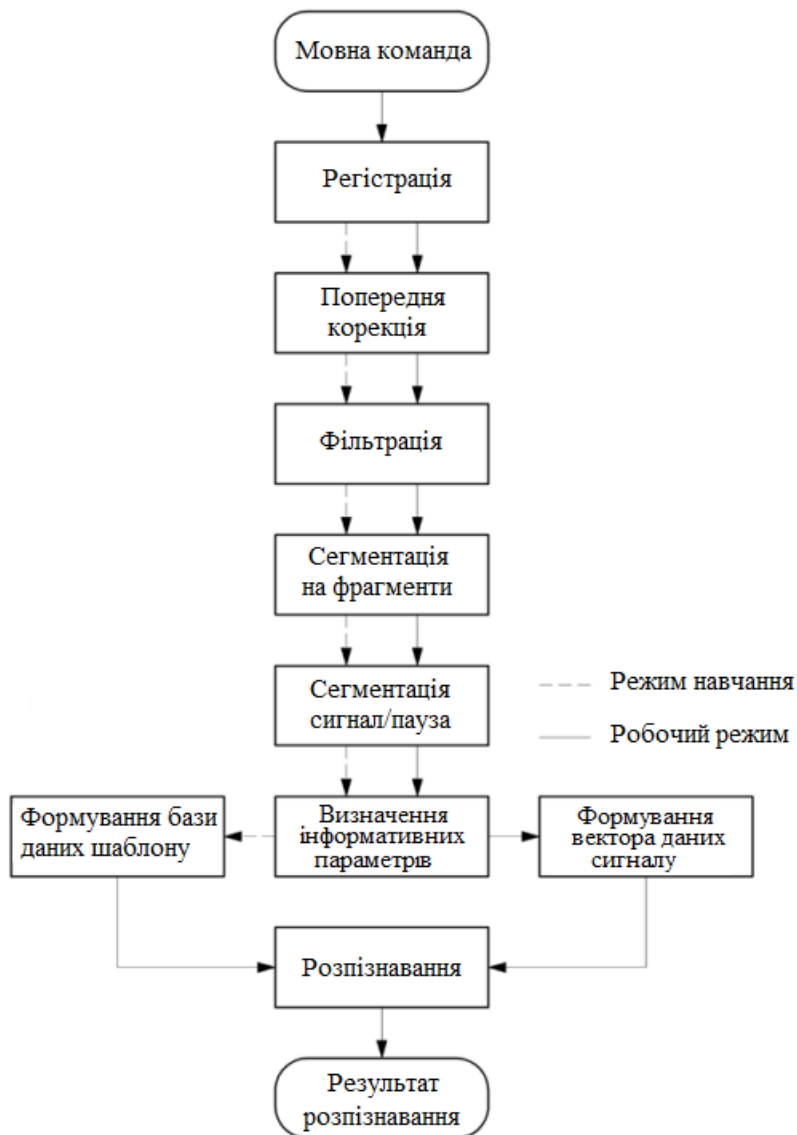


Рисунок 1.9 - Етапи обробки мовних сигналів у системі голосового управління

Реєстрація - запис мовного сигналу в реальному часі. Тиск звукової хвилі, на даному етапі сприймається мікрофоном і перетворюється на аналоговий сигнал. Мовний сигнал з аналогового сигналу перетворюється на цифровий, з використанням, на даному етапі аналого-цифрового перетворювача, який здійснює квантування та дискретизацію.

Реєстрація включає наступні додаткові операції: збільшення типів реєстрованої інформації, кількості каналів запису; групове та індивідуальне переналаштування параметрів реєстрації та без переривання процесу перетворення запису; автоматичне регулювання вирівнювання та посилення рівня віддаленого та ближнього користувача, забезпечується якісний відповідний запис сильних як так і слабких мовних команд.

Для усунення спотворень спектру звуку використовується попередня корекція, які виникають під час промови у мовному апараті. Мовний сигнал пропускають з відповідною функцією передавання через коригуючий фільтр. Даний етап не обов'язковий, в більшості голосових системах не передбачена корекція.

Фільтрування дозволяє зменшити частку шумів, підвищити розбірливість, спричинених технологічними причинами та акустичними. Шум - безпорядні коливання хвилі різної фізичної природи, відрізняються складністю спектральною часовою структурою. Відносно до мовних команд – шум, в даному випадку сукупність аперіодичних звуків різної частоти та інтенсивності, які змінюють, при цьому інформативні ознаки мовного сигналу.

По характеру взаємодії з мовним сигналом шум поділяється на мультиплікативні та адитивні. Адитивні шуми накладаються з мовним сигналом та вносять незначну помилку. Мультиплікативні шуми перемножуються з мовним сигналом, вносять, при цьому, найбільшу помилку, дані шуми можуть змінювати інформативні параметри мовних сигналів. В загальному виді, комбінація мовного сигналу та шуму, має наступний вид:

$$S(t) = [k_s(t) + k_n(t)] \times e(t) + n(t), \quad (1.1)$$

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						20
Зм	Арк	№ докум.	Підпис	Дата		

де  $S(t)$  - мовний сигнал,  $k_s(t)$  - коефіцієнт, який характеризує мовний сигнал;  $e(t)$  - мовний сигнал;  $n(t)$  адитивний шум;  $k_n(t)$  - коефіцієнт, який характеризує мультиплікативний шум.

Відношення інтенсивності шуму та мовного сигналу називається відношення сигнал/шум, відіграє, в даному випадку, важливу роль у задачах шумоподавлення та фільтрації. Відношення сигнал/шум визначається у логарифмічних одиницях - децибелах:

$$N = 10 \cdot \lg I_s / I_n, \quad (1.2)$$

де  $I_n, I_s$  - інтенсивності шуму та сигналу відповідно.

Шуми в мовних сигналах, за походженням можна поділити на антропогенні та фізіологічні. До фізіологічних шумів відносяться комплекс звуків різної частоти та інтенсивності, перебувають, при цьому у безпорядному поєднанні з мовними сигналами. Фізіологічні шуми безпосередньо походять з порушеннями мови. До фізіологічних шумів, відносять звуки, структура та форма яких безпосередньо пов'язана з порушенням звуковимовляння: розпад мовних сигналів (афазія); порушення голосу (дисфонії, ринофонія, афонії); порушення ритму та темпу мовних сигналів (заїкання, брадилалія, спотикання, тахілалія).

До антропогенних шумів відноситься решта видів шумів. Антропогенні шуми виникли внаслідок діяльності людини, створювані людиною. Дані шуми називають виробничими чи промисловими шумами. Джерела антропогенних шумів: транспортні засоби - залізничні поїзди, автомобілі, літаки, офісна та побутова техніка, промислові підприємства, ремонтні та будівельні роботи.

Шуми, за постійністю параметрів поділяються на нестационарні та стаціонарні. Стаціонарний шум – характеризується сталістю параметрів: автокореляційної функції, розподілу інтенсивності за спектром, інтенсивності. Прикладом стаціонарного шуму - білий шум, по всьому діапазону частот спектральні складові рівномірно розподілені. Короткі проміжки часу триває нестационарний шум. Нестационарні шуми діляться на коливальні, імпульсні, переривчасті. Імпульсний шум - складається з декількох звукових сигналів, тривалістю, при цьому менше одної секунди, рівні відрізняються не менше ніж на 6 дБ. Переривчастий шум -

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						21
Зм	Арк	№ докум.	Підпис	Дата		

рівень поступово змінюється (на 4 дБ), тривалість інтервалів становить більше секунди, на протязі яких рівень шуму залишається постійним.

До нестационарних шумів у мовних сигналах відносяться: рідкісні імпульсні перешкоди в радіотехніці, окремі стуки у виробничих умовах, вуличний шум транспорту.

По ширині спектра виділяють вузькосмуговий та широкосмуговий шум. Тональний (вузькосмуговий) шум - у якому певної частоти прослуховується звук. Широкасмуговий шум - з безперервним спектром включає набір багатьох звукових частот шириною смуги більше однієї октави.

Шуми по частотному діапазону поділяються на: високочастотний (1000 Гц), середньочастотний шум (400 - 1000 Гц), низькочастотний шум (400 Гц).

Сегментація мовного сигналу - лінійний розподіл сигналу на складові відрізки, фрагменти. Мовні команди є складної форми нестационарними сигналами, характеристики та параметри яких змінюються швидко протягом часу. Однак в основі методів обробки мовних команд лежить припущення, що властивості сигналу з часом повільно змінюються. Це приводить до використання методів короткочасного аналізу, в яких фрагменти мовного сигналу обробляються і виділяються, як короткі ділянки окремих звуків. Для того щоб отримати відповідні набори однакової довжини інформативних параметрів, необхідно сегментувати мовну команду на рівні фрагменти, вважаючи при цьому, що мовний сигнал на такому відрізку стаціонарний (рис. 1.10). Для запобігання втраті інформації необхідно використовувати перекриття фрагментів мовного сигналу на границі.

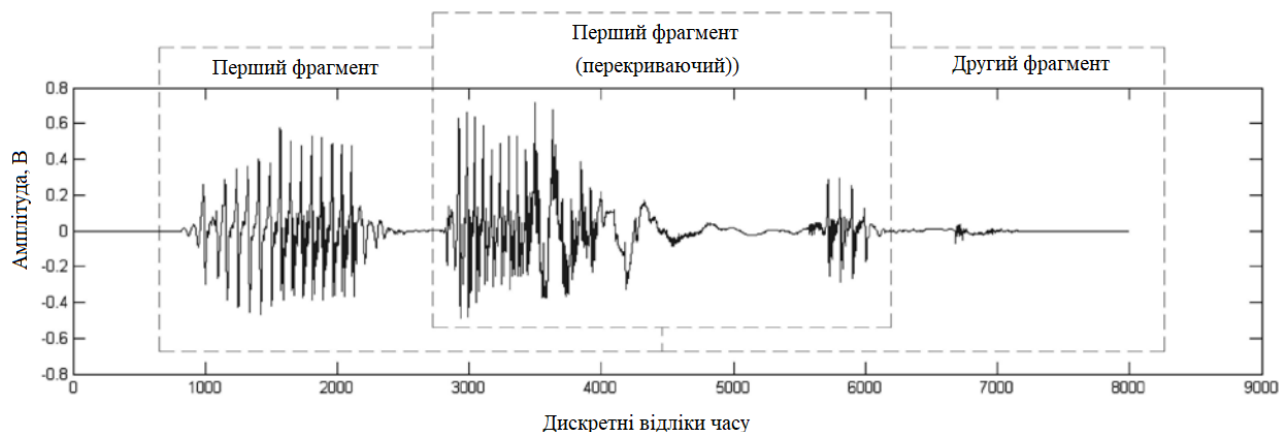


Рисунок 1.10 - Сегментація мовної команди на кадри

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		22

Сегментація сигнал/пауза - визначення моментів закінчення та початку фрази. Дана задача є однією з найважливіших, за наявності шуму в області обробки мовних сигналів, при використанні системи голосового управління необхідно визначити моменти закінчення та початку сигналу.

Визначення інформаційних параметрів - визначення унікальних характеристик та властивостей мовних команд [11, 12]. Поняття, що характеризують інформативні параметри мовної команди, пов'язані з розмірами, динамікою зміни мовного апарату, формою, описуються емоційним станом людини.

Таким чином, на основі проведеного аналізу досягнень в області виділення інформативних ознак мовних команд, інформативні параметри можна розділити на три групи, які дозволяють розпізнати мовні команди: кепстральні, амплітудно-спектрально-часові, частотні.

Амплітудно-частотні параметри: амплітуда, формантні частоти, інтенсивність, часота основного тону, енергія. Амплітудно-частотні параметри несуть достатню та необхідну інформацію про мовну команду при мінімальному часі сприйняття, використовуються у простих системах голосового управління.

Спектральні параметри: варіація огинаючої спектра мови; відносний час перебування мовного сигналу у смугах спектра; середнє значення спектра мовної команди; медіанне значення спектра мовного сигналу у смугах; відносна потужність спектра мовної команди у смугах.

Часові параметри: коефіцієнт форми сегмента; висота сегмента; тривалість сегмента структурної одиниці мінімального мовлення (трифона, алофона, фонем, дифона). Спектрально-часові параметри характеризують мовну команду, виходячи з наявності компонентів трьох видів у фізико-математичній сутності: тональних (періодичних) ділянок звукової хвилі; ділянок, які містять мовні паузи; неперіодичних ділянок звукової хвилі (вибухових, шумових). Особливості мовного потоку характеризують спектрально-часові параметри, пов'язані з динамікою перебудови органів мовлення, є відповідними інтегральними характеристиками мовного потоку, відображають синхронності руху артикуляційних органів людини, своєрідність взаємозв'язку [12].

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						23
Зм	Арк	№ докум.	Підпис	Дата		

Кепстральні параметри: коефіцієнти кепстру лінійного передбачення та мел-частотні кепстральні коефіцієнти. Більшість систем голосового управління зосереджують зусилля на використанні частотної характеристики мовного тракту, відкидаючи характеристики сигналу збудження. Коефіцієнти першої моделі забезпечують, при цьому найкращий розподіл звуків. Для відділення хибного сигналу від сигналу мовного тракту використовується кепстральному аналіз.

Розпізнавання мовних сигналів є задачею кінцевою обробки в системах голосового управління. Принцип роботи полягає в порівнянні мовного сигналу, який надійшов в систему, з шаблоном бази даних, які формуються в ході навчання системи. Застосування відповідних методів порівняння впливає якість розпізнавання. На теперішній час існує декілька математичних апаратів порівняння, широко використовуваних у системах голосового управління.

Проведене дослідження, допоможе об'єктивно провести аналіз методів обробки мовних сигналів.

#### 1.4 Постановка задачі

З проведеного дослідження слідує: традиційні методи обробки мовних команд, призначені для стаціонарних та лінійних сигналів. Застосування методів для задач обробки мовних сигналів у системах голосового управління є неефективним.

Таким чином, розробка пристроїв та алгоритмів обробки мовних сигналів є одним з найбільш актуальним та важливим завданням при побудові систем голосового управління керуючих систем. Для розв'язання даної задачі включає наступні етапи: аналіз та дослідження предметної області; огляд сучасного стану систем голосового керування; дослідження особливості обробки мовних команд у підсистемах голосового управління керуючих систем; аналіз методів обробки мовних команд у підсистемі голосового управління керуючих систем; реалізація, дослідження та впровадження пристроїв та алгоритмів у вигляді підсистеми голосового управління керуючих систем.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		24

## 2 РОЗРОБКА АЛГОРИТМІВ І ПРИСТРОЇВ ОБРОБКИ МОВНИХ КОМАНД ПІДСИСТЕМИ ГОЛОСОВОГО УПРАВЛІННЯ КЕРУЮЧИХ СИСТЕМ

### 2.1 Технологічний процес підсистеми голосового управління обробки мовних команд

Підвищення точності у підсистемі голосового управління розпізнавання мовних команд може бути забезпечене вдосконаленням всіх етапів обробки мовного сигналу. Для підвищення точності у підсистемі голосового управління розпізнавання мовних команд пропонується використовувати технологічний процес обробки сигналу, схематично представлено на рис. 2.1.

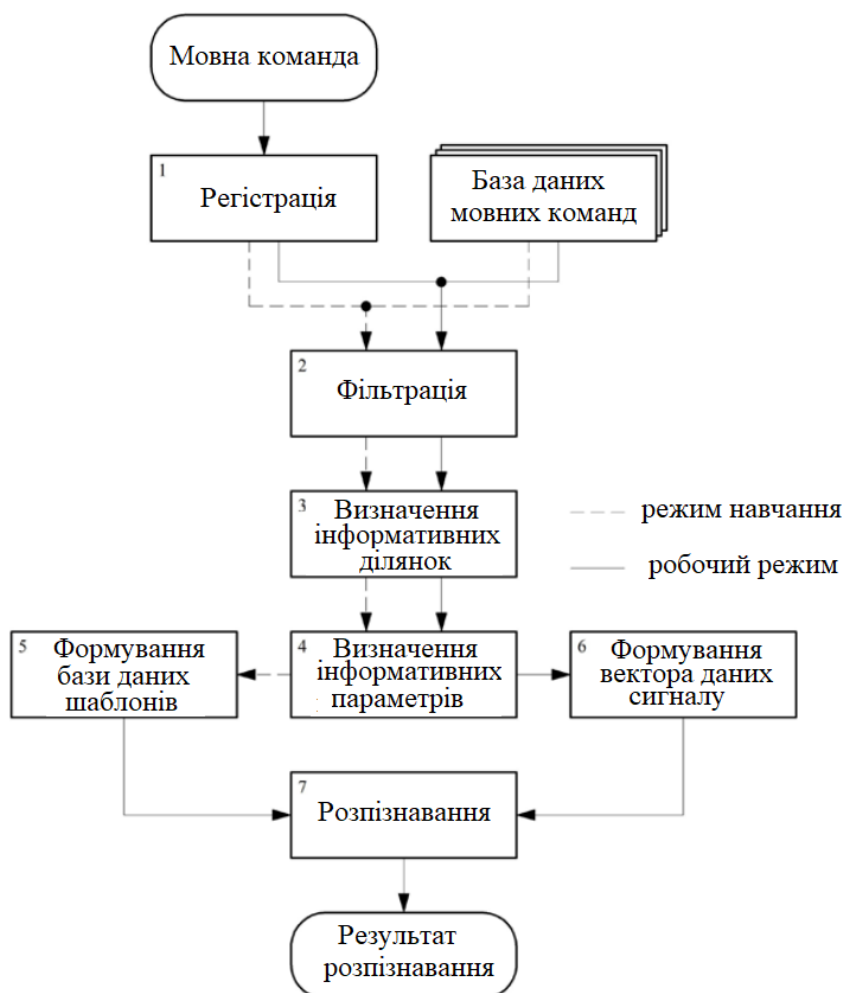


Рисунок 2.1 - Технологічний процес обробки мовних сигналів у підсистемі голосового управління

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						25
Зм	Арк	№ докум.	Підпис	Дата		

У запропонованому технологічному процесі системи обробки мовних команд виконується послідовність наступних етапів:

1. Реєстрація - аудіо захоплення мовного сигналу в режимі реального часу та перетворення команди в цифровий вид з використанням сучасних стандартних засобів: мікрофон, аналого-цифровий перетворювач попередній та основний підсилювач.

2. Фільтрація - важливий етап обробки мовних сигналів, дозволяє зменшити частку шумів у сигналі, підвищити розбірливість. До етапу фільтрації можна віднести: видалення постійної складової, яка виникає в аналого-цифровому перетворювачі (зміщення сигналу на деяку постійну величину відносно нуля) попередню корекцію, для усунення спотворень природних, які виникають при формуванні мови у мовному апараті людини.

3. Визначення інформативних ділянок мовних сигналів - виділення певних частин команди, які містять відповідну інформацію про унікальні властивості мовного сигналу. До даного етапу можна віднести: сегментацію невокалізованих та вокалізованих ділянок мовної команди, сегментацію ділянок пауз та сигналу.

4. Визначення інформативних параметрів - виявлення спектрально-часових, кепстральних характеристик, амплітудно-частотних мовних сигналів, які характеризуються розмірами, динамікою зміни формою мовного апарату.

5. Формування даних сигналу і бази даних шаблонів - об'єднання інформативних параметрів в один вектор.

6. Розпізнавання - процес порівняння шаблону (отриманого в процесі навчання системи) з бази даних з мовним сигналом.

Технологічний процес обробки мовних сигналів у підсистемі голосового управління ґрунтується на поєднанні різних пристроїв та алгоритмів обробки мовних команд. В основі технологічний процесу обробки мовних сигналів на етапах обробки команди використовується метод комплементарної множинної декомпозиція на емпіричні моди (Complementary ensemble empirical mode decomposition, CEEMD) [16].

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		26

На етапах обробки мовних команд, що надходять необхідно розробити пристрої та алгоритми обробки сигналів, що дозволять оцінити якість розпізнавання мовних сигналів підсистеми голосового управління.

## 2.2 Алгоритм та пристрій фільтрації нестационарних фонових шумів на основі дослідження шумових та енергетичних характеристик

Нестационарні фонові шуми - найбільш характерні шуми, негативно впливають на працездатність розпізнавання підсистем голосового управління керуючих систем. Існують фонові шуми різних типів, що впливають на якість мовних сигналів. Можливий варіант, що мовна команда одночасно декількома видами шумів схильна до спотворення.

Фонові шуми погіршують розбірливість мовних команд, можуть призвести до значної похибки між еталонами і зашумленими мовними командами, що надходять в систему розпізнавання. Значні похибки є основною причиною некоректного розпізнавання мовного сигналу, тому методи фільтрації спрямовані на зменшення цієї похибки, які застосовуються у підсистемах голосового управління керуючих систем.

Проведене дослідження фільтрації зашумлених мовних команд показало, що до фільтрації зашумлених мовних команд існує багато різних підходів. Різноманітність підходів обумовлена відсутністю досить ефективних підходів її вирішення, так і важливістю проблеми. Важливо, що з певних завдань ефективність фільтрації оцінюється по-різному. Рівень залишкового шуму для підсистеми голосового управління є критерієм ефективності фільтрації.

Розглянуті підходи до фільтрації залишкового шуму в повному обсязі не вирішують проблему. Причиною є неможливість коректно аналізувати нестационарні та нелінійні мовні сигнали складної форми (шумові та корисні) різних рівнів інтенсивності. Аналіз алгоритмів фільтрації зашумлених мовних команд, що вони засновані на методах множинної декомпозиція на емпіричні моди, використовують інструменти для аналізу нестационарних даних. На рис. 2.2

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						27
Зм	Арк	№ докум.	Підпис	Дата		

представлено алгоритм фільтрації. Алгоритм передбачає декомпозицію зашумленої мовної команди на емпіричні моди (блок 3), визначення моди з основним шумом та аналіз емпіричних мод (блоки 4, 5), фільтрацію з використанням адаптивної компенсації (блок 6).

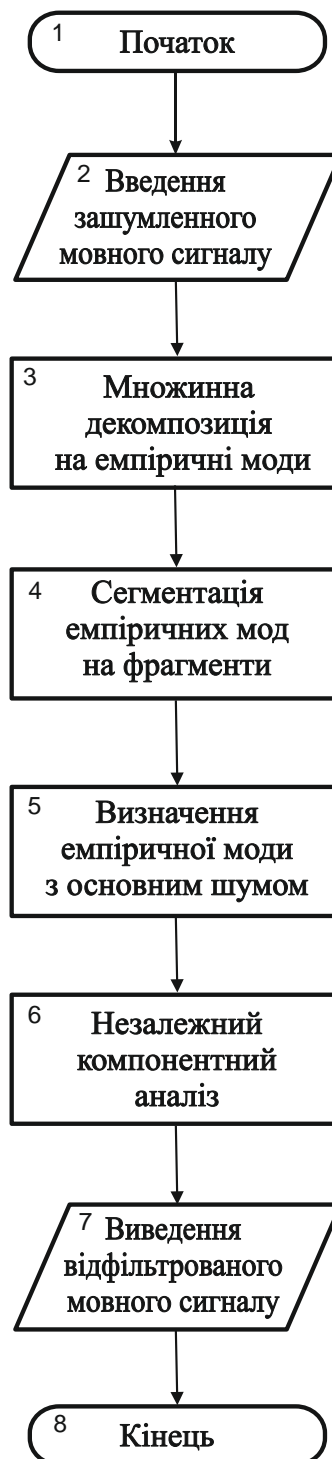


Рисунок 2.2 –Алгоритм фільтрації мовних сигналів на основі МДЕМ

Ефективність фільтрації алгоритму нестационарних фонових шумів залежить від наступних блоків: блок визначення емпіричних мод з основним шумом, який при

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		28

розпізнаванні сигналу повинен забезпечувати коректне виявлення емпіричної моди в умовах змішування мод вираженого ефекту; блок декомпозиції, задача якого забезпечувати розкладання сигналу на емпіричні моди в умовах інтенсивних фонових шумів з мінімальним ефектом змішування мод.

Дослідження алгоритму фільтрації мовних сигналів на основі МДЕМ показали при інтенсивних фонових шумах низьку ефективність фільтрації. Для зменшення спотворень мовного сигналу та рівня залишкових шумів пропонується використовувати алгоритм пристрою фільтрації, наведеного на рис. 2.3.

Наведений алгоритм роботи пристрою (рис. 2.3) забезпечує коректне визначення емпіричних мод з основним шумом з використанням додаткового блоку аналізу емпіричних мод, зменшення ефекту змішування емпіричних мод при наявності інтенсивних фонових шумів за рахунок використання комплементарної множинної декомпозиція на емпіричні моди.

Розглянемо основні етапи роботи алгоритму фільтрації нестационарних шумів фону мовних сигналів.

Блок 2 - Введення зашумленої мовної команди  $x(n)$  ( $n$  - дискретний відлік часу  $0 < n \leq N$ ,  $N$  - число дискретних відліків у мовному сигналі).

Блок 3. Налаштування комплементарної множинної декомпозиція на емпіричні моди, декомпозиція зашумленого мовного сигналу полягає в отриманні усереднених значень ЕМ та залишку.

Блок 4. Після виконання операції декомпозиції зашумлена мовна команда представляє набір емпіричних мод і залишок. Встановлюється номер моди, оскільки подальша робота алгоритму пристрою здійснюється з кожною емпіричних модою окремо.

Блок 5. На перших етапах роботи алгоритму пристрою необхідно встановити пороги вагових коефіцієнтів (шумового та енергетичного), які необхідні на наступних етапах для визначення емпіричних мод з основним шумом. Початкові значення вагового шумового та енергетичного коефіцієнтів виставляються рівними одиниці, в подальшому будуть використовуватися у визначенні емпіричної моди, що містить основний шум.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						29
Зм	Арк	№ докум.	Підпис	Дата		

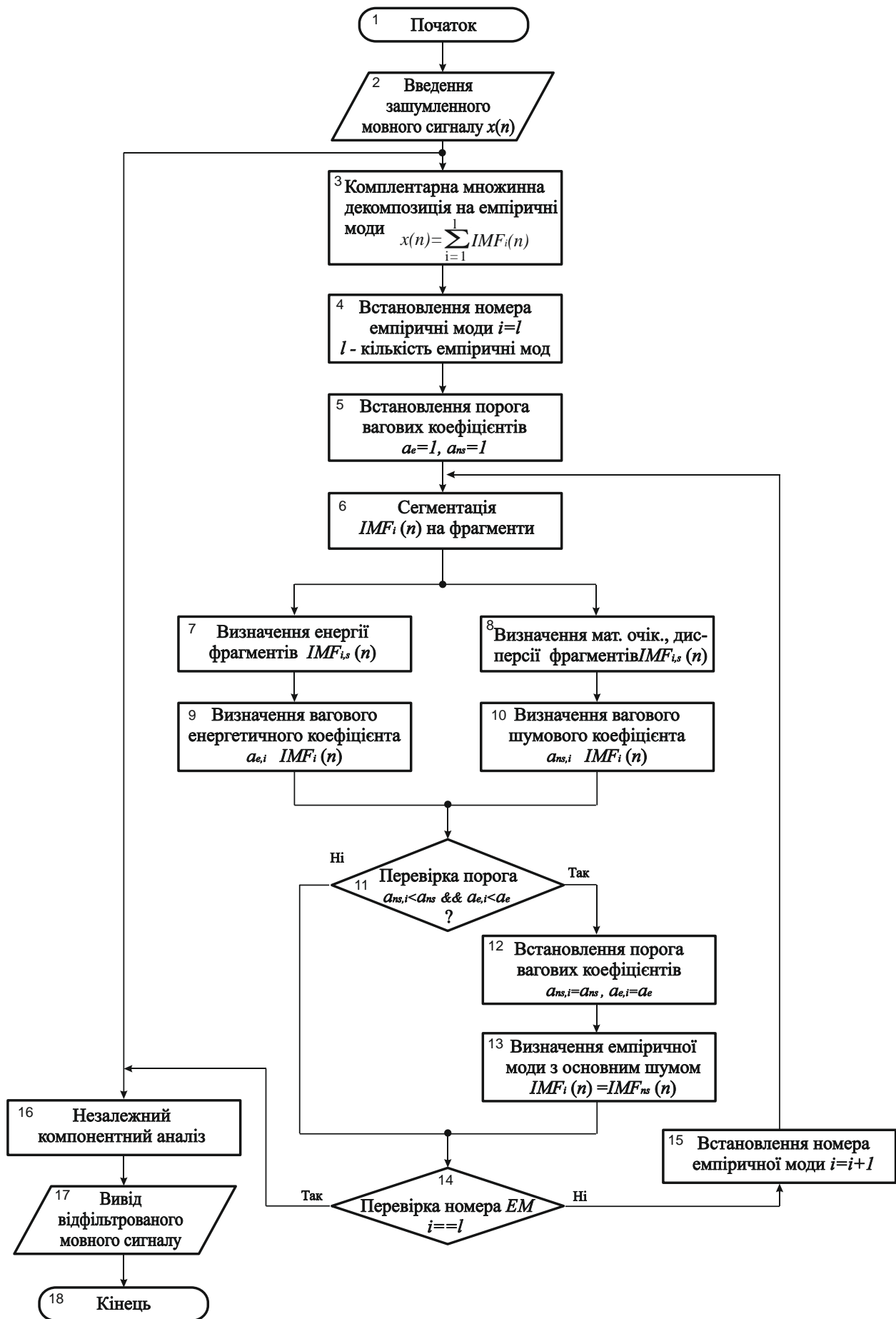


Рисунок 2.3 – Алгоритм пристрою фільтрації нестационарних фонових шумів

Блок 6. Здійснюється сегментація емпіричних мод на фрагменти

Блок 7. Визначення енергії фрагментів. Формування мови, відповідно до фізіологічного аспекту, людина перед реєстрацією голосової команди та її вимовою та її робить короткочасну паузу – зазвичай 200 мс. Дана ділянка паузи відповідає тиші з фоновим шумом і не містить мови та.

Блок 8. Перші 200 мс мовної команди містять фонові шуми, можна обчислити статичні характеристики фонового шуму.

Блок 9. Визначення емпіричних мод з основним шумом здійснюється з використанням вагового енергетичного коефіцієнта. Якщо значення коефіцієнта наближається до мінімального значення, відповідна емпірична мода вважається модою з фоновим шумом. Якщо значення енергетичного коефіцієнта наближається до одиниці, то відповідна емпірична мода вважається корисною модою, містить мовний сигнал.

Блок 10. Визначення емпіричної моди з фоновим шумом здійснюється за допомогою шумового вагового коефіцієнта. Після визначення статусу ділянки необхідно визначити статус фрагмента. Для визначення статусу використовується правило - якщо у фрагменті 70% ділянок мають статус шуму, фрагмент відноситься до шумового. Після визначення статусу фрагментів визначається шумовий ваговий коефіцієнт емпіричної моди. Якщо шумовий коефіцієнт наближається до мінімального значення, то емпірична мода є модою з основним шумом. Якщо шумовий коефіцієнт наближається до одиниці, то емпірична мода є модою з корисним сигналом.

Блоки 11, 12, 13. Після визначення вагових коефіцієнтів емпіричної моди здійснюється порівняння з пороговими значеннями (блок 11). Якщо умова виконується, встановлюються нові значення порогових коефіцієнтів рівні поточним значенням (блок 12) та енергетична мода вважається як мода з основним шумом (блок 13).

Блоки 14, 15. Після завершення аналізу емпіричної моди здійснюється порівняння номера поточної моди зі загальною кількістю. Якщо умова не виконується, то встановлюється наступний номер емпіричної моди (блок 15) і

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						31
Зм	Арк	№ докум.	Підпис	Дата		

перехід до блоку 6 для аналізу наступної емпіричної моди. Якщо умова виконується, перехід до блоку 16.

Блок 16. Після визначення моди з основним шумом та аналізу емпіричних мод здійснюється фільтрація з використанням методу адаптивної компенсації шуму, в результаті виділяються складові відфільтрованої мовної команди та шуму.

### 2.3 Алгоритм та пристрій вимірювання частоти основного тону на основі дослідження частотно-часових характеристик

Мовні сигнали – нелінійні, нестационарні сигнали складної форми, які складаються з невокалізованих та вокалізованих ділянок, утворюються в результаті неперіодичних та періодичних коливань голосових зв'язок людини. Частота коливань голосових зв'язок людини є важливою характеристикою мови – частота основного тону. Визначення частоти основного тону, залежно від ситуації є: маркування періодів основного тону, вимірювання частоти основного тону, виділення контуру основного тону.

Точність частоти основного тону визначається частотним та часовим дозволами виміру. Часовий дозвіл - швидкість реагування зміни частоти, частотний - точність визначення малих змін частоти. Обидва показників чутливі до модуляцій основного тону, рівня зашумленості мовної команди. Таким чином, вибір алгоритму для вимірювання частоти основного тону є компромісом між часовим і частотним дозволами.

На теперішній час розроблено достатньо методів вимірювання частоти основного тону: методи обробки у частотній області, методи обробки у частотно-часовій області, методи обробки у часовій області.

Вимірювання у частотно-часових методах здійснюється проведенням аналізу з урахуванням, при цьому, передбачуваного контуру основного тону, виділяючи в процесі обробки миттєві максимуми гармонік та розподіляючи мовний сигнал на невокалізовані та вокалізовані ділянки. Недоліком є велика можливість отримання

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						32
Зм	Арк	№ докум.	Підпис	Дата		

максимуму енергії миттєвого в невокалізованих ділянках через присутність в них шуму в мовних командах.

Вимірювання у частотних методах здійснюється з використання максимальних значень енергії піків (спектра) та в подальшому зіставлення з частотами, кратними частоті основного тону. Недоліком - наявність частот у поточній смузі, крім частот основного тону наступних гармонік з більшою енергією.

Вимірювання у часових методах здійснюється за проведенням аналізу у часовій області розподілу спектра, кореляції, переходів через нуль. Часові методи найбільш точніші, вимагають налаштування та ретельної фільтрації мовної команди. Недолік - висока чутливість до рівня шуму сигналу.

Популярність наведених алгоритмів обумовлена низьким відсотком грубих помилок, належною функціональністю, наявністю вільно розповсюджуваних програмних реалізацій. Можливості алгоритмів вимірювання частоти основного тону суттєво обмежені нелінійністю та нестационарністю мовних команд. В їх основі лежить стаціонарна модель мовного сигналу, яка використовує точне повторення періоду основного тону, при зміні періоду точність визначення оцінок частоти істотно падає. На рис. 2.4 представлений спрощений алгоритм вимірювання частоти основного тону.

Алгоритм передбачає декомпозицію мовного сигналу емпіричної моди, попередню фільтрацію мовного діапазону (блок 3), методом множинної декомпозиції емпіричних мод (блок 4), визначення, аналіз мод з основного тону (блок 5) та вимірювання частоти основного тону звикористанням визначень перетинів через нуль (блоки 6, 7).

Ефективність алгоритму залежить, в основному, від декомпозиції, яка забезпечує розкладання мовного сигналу на емпіричні моди з мінімальним ефектом, при цьому, від коректного визначення емпіричних мод з основного тону, змішування мод в умовах наявного ефекту змішування мод. Через некоректне визначення емпіричної моди з основного тону та ефект змішування емпіричних мод виникає проблема, в даній ситуації, неточного визначення переходів мовного сигналу через нуль і відповідно некоректного вимірювання частоти основного тону.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						33
Зм	Арк	№ докум.	Підпис	Дата		

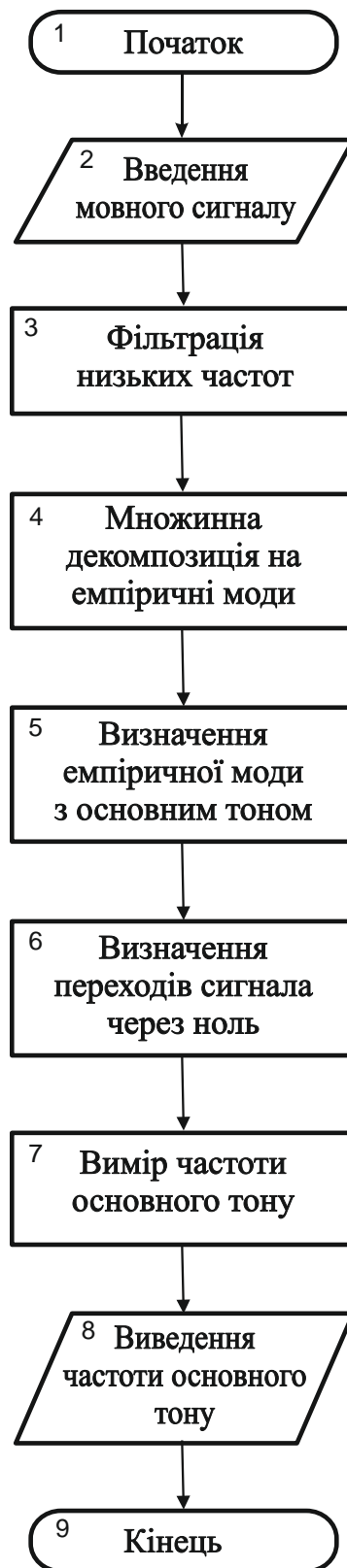


Рисунок 2.4 - Алгоритм вимірювання частоти основного тону

На рис. 2.5 наведено алгоритм пристрою вимірювання частоти основного тону, на основі підходу комплементарної множинної декомпозиції емпіричних мод

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		34

та оцінці енергії емпіричних мод за допомогою оператора ТЕО (Teager Energy Operator,).

Суть алгоритму пристрою вимірювання частоти основного тону полягає в адаптивному розкладі мовної команди на частотні складові емпіричної моди за допомогою методу комплементарної множинної декомпозиції емпіричних мод. Визначається мода з отриманого набору емпіричних мод, яка містить основний тон, з використанням оцінки логарифму енергетичних характеристик. З використанням оператора ТЕО з моди з основного тону визначається частота основного тону.

Алгоритм (рис. 2.5) виконує вимірювання частоти поточного аналізуемого фрагмента. Розглянемо основні етапи роботи алгоритму пристрою виміру частоти основного тону.

Блок 2. Введення мовної команди без пауз  $x(n)$  ( $n$  - дискретний відлік часу  $0 < n \leq N$ ,  $N$  - число дискретних відліків у мовному сигналі).

Блоки 3, 4. Комплементарна множинна декомпозиція емпіричної моди на фрагмента. Подальша робота здійснюється окремо з кожною емпіричною модою. Встановлюється номер моди фрагмента.

Блок 5. Визначення логарифму енергії емпіричної моди.

Блоки 6, 7. Після проведеного аналізу першої емпіричної моди здійснюється порівняння номера поточної моди з їх загальною кількістю (блок 6). Якщо, задана умова не виконується, номер емпіричної моди збільшується на одиницю (блок 7), повернення до блоку 5 для проведення аналізу наступної емпіричної моди. Якщо умова виконується, перехід до блоку 8.

Блоки 8. Необхідно визначити моду з основного тону, після обчислення всіх логарифмів енергії емпіричних мод. Визначення емпіричних мод з основного тону засноване на тому, що моди мають більшу енергію ніж невокалізовані, містять вокалізовану мову. Для подальшого аналізу використовуються моди, що містять більшу енергію інформативних мод з номерами чотири і сім. Встановлюється номер першої інформативної емпіричної моди, з номером  $i = 4$ .

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		35

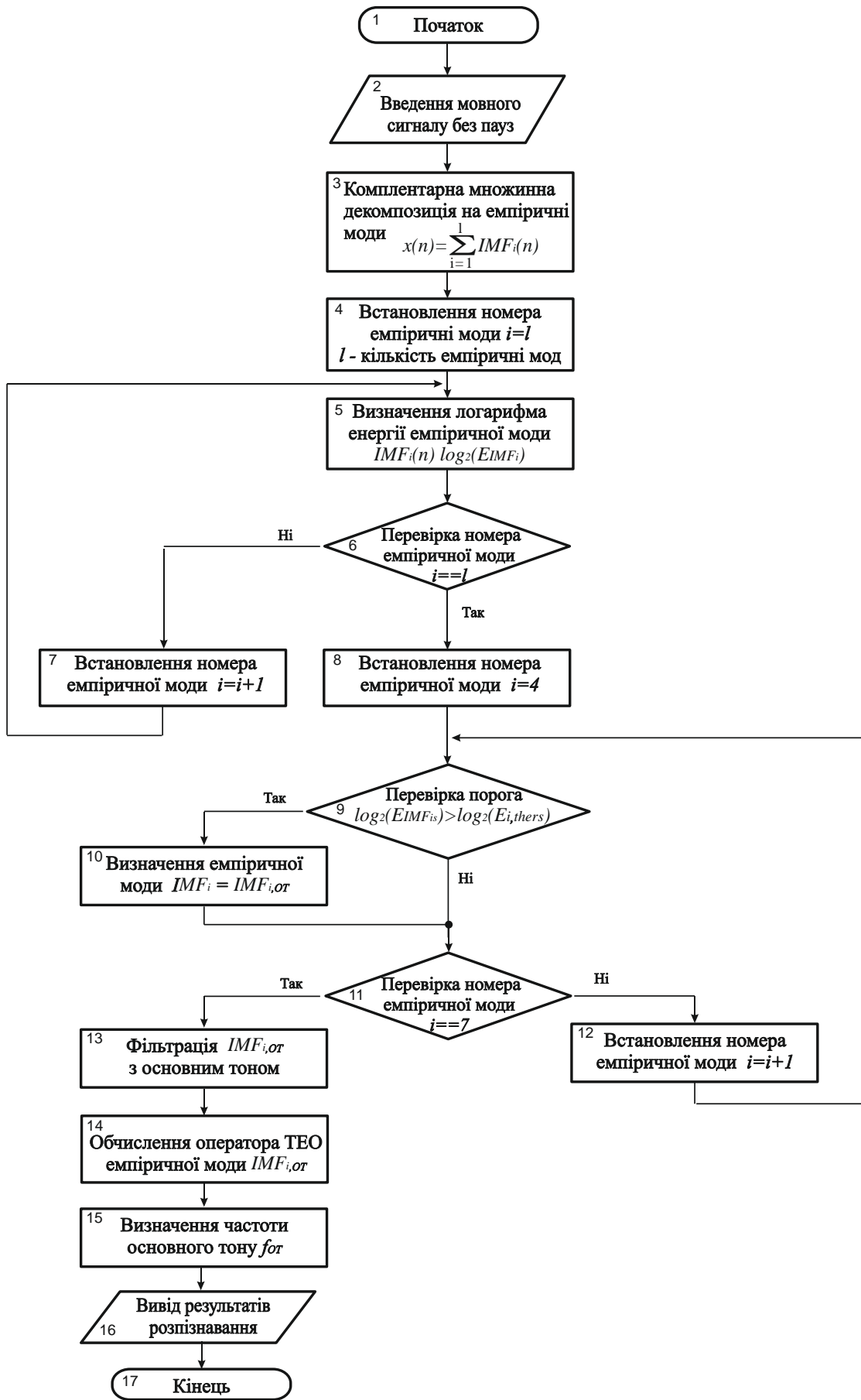


Рисунок 2.5 Алгоритм пристрою вимірювання частоти основного тону

Зм	Арк	№ докум.	Підпис	Дата

КвРКІ.180243.18.02.16 ПЗ

Арк.

36

Блоки 9, 10. Визначення емпіричної моди з основного тону полягає у обробці порогів - послідовному порівнянні порогових значень з логарифмами енергії інформативних емпіричних мод. Порогова обробка здійснюється в блоку 9. Якщо умова виконується, то емпірична мода визначається із основного тону (блок 10), якщо не виконується, перехід до блоку 11.

Блоки 11, 12. Після проведення аналізу першої інформативної емпіричної моди, порівнюються номер поточної моди з загальною кількістю (блок 11). Якщо умова не виконується, номер емпіричної моди збільшується на одиницю (блок 12), перехід до блоку 9 для проведення аналізу наступної інформативної емпіричної моди, якщо виконується, перехід до блоку 13.

Блок 13. Після визначення емпіричної моди з основного тону проводиться фільтрація поточної моди від частотних складових неінформативних.

Блоки 14. Для вимірювання миттєвого значення частоти основного тону використовується відповідна функція вимірювання миттєвої енергії сигналу.

Блок 15. Після обчислення функції оператора ТЕО здійснюється вимірювання частоти основного тону.

#### 2.4 Алгоритм та пристрій сегментації сигнал/пауза на основі дослідження логарифму енергетичних характеристик

Сегментація на інформативні ділянки та паузи мовних сигналів є одним із важливих задач у підсистемі голосового управління при обробці голосових команд. Виявлення точне меж мовної команди підвищує якість розпізнавання, зменшує кількість розрахункових і обчислювальних операцій.

Проведене дослідження показало, що на теперішній час існують різні підходи до сегментації сигнал/пауза, які вирішують проблему виявлення меж мовної команди. Серед найбільш відомих можна виділити наступні: підходи, що ґрунтуються на значеннях інформаційної ентропії; підходи, що ґрунтуються на використанні мел-частотних кепстральних коефіцієнтах; підходи, що базуються на статичних моделях; підходи, засновані на вейвлет перетвореннях.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						37
Зм	Арк	№ докум.	Підпис	Дата		

Найбільш практичними застосування є: алгоритм одновимірної відстані Махаланобіса та фоновому шуму з використанням статистичних властивостей; алгоритм на основі спільного використання ZCR (Short-time Zero-crossing Rate) та STE (Short-time Energy).

У першому випадку алгоритм ґрунтується на використанні статистичних властивостях фоновому шуму. Сегментація сигнал/пауза у мовній команді здійснюється з обчисленням відстані Махаланобісу для кожної ділянки сигналу. У даному випадку властивості функцій розподілу присутніх в мовному сигналі шумів є критеріями і використовуються, при цьому для сегментації сигнал/пауза.

У другому випадку в основі алгоритму лежить правило розмежування: якщо кількість переходів мовного сигналу короткочасної ділянки через нуль та значення енергії відрізняються від заданих порогових значень, то даний участок є фоновим шумом або тишею. Якщо значення меншою мірою відмінні від заданого порогу, то участок є корисним сигналом. Для коректної сегментації мовного сигналу в - невідомо, якими повинні бути призначенні граничні значення.

На рис. 2.6 наведено спрощений вид алгоритму сегментації на основі одновимірної відстані Махаланобіса. Проведені дослідження алгоритму при обмежених порогових значень та інтенсивних фонових шумах виявили низьку ефективність сегментації. Точність сегментації сигналу значно погіршується зі збільшенням відношення сигнал/шум, є наслідком використання, в даному випадку неадаптивних підходів аналізу та нестійких до шуму нелінійних та нестационарних складної форми мовних сигналів. Тому необхідно застосувати адаптивного аналіз даних для вирішення задач сегментації сигнал/пауза без порогових значень (будь-якої попередньої інформації). Алгоритм пристрою сегментації сигналів, що використовує адаптивний аналіз даних представлено на рис. 2.7. Суть даного підходу полягає у сегментації мовної команди на короткочасні фрагменти для проведення адаптивного аналізу методом комплементарної множинної декомпозиції на емпіричні моди з подальшим формуванням на основі оцінки енергії емпіричних мод та фізіологічного аспекту формування мови адаптивного порога. Розглянемо основні етапи роботи алгоритму пристрою сегментації сигнал/пауза.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						38
Зм	Арк	№ докум.	Підпис	Дата		

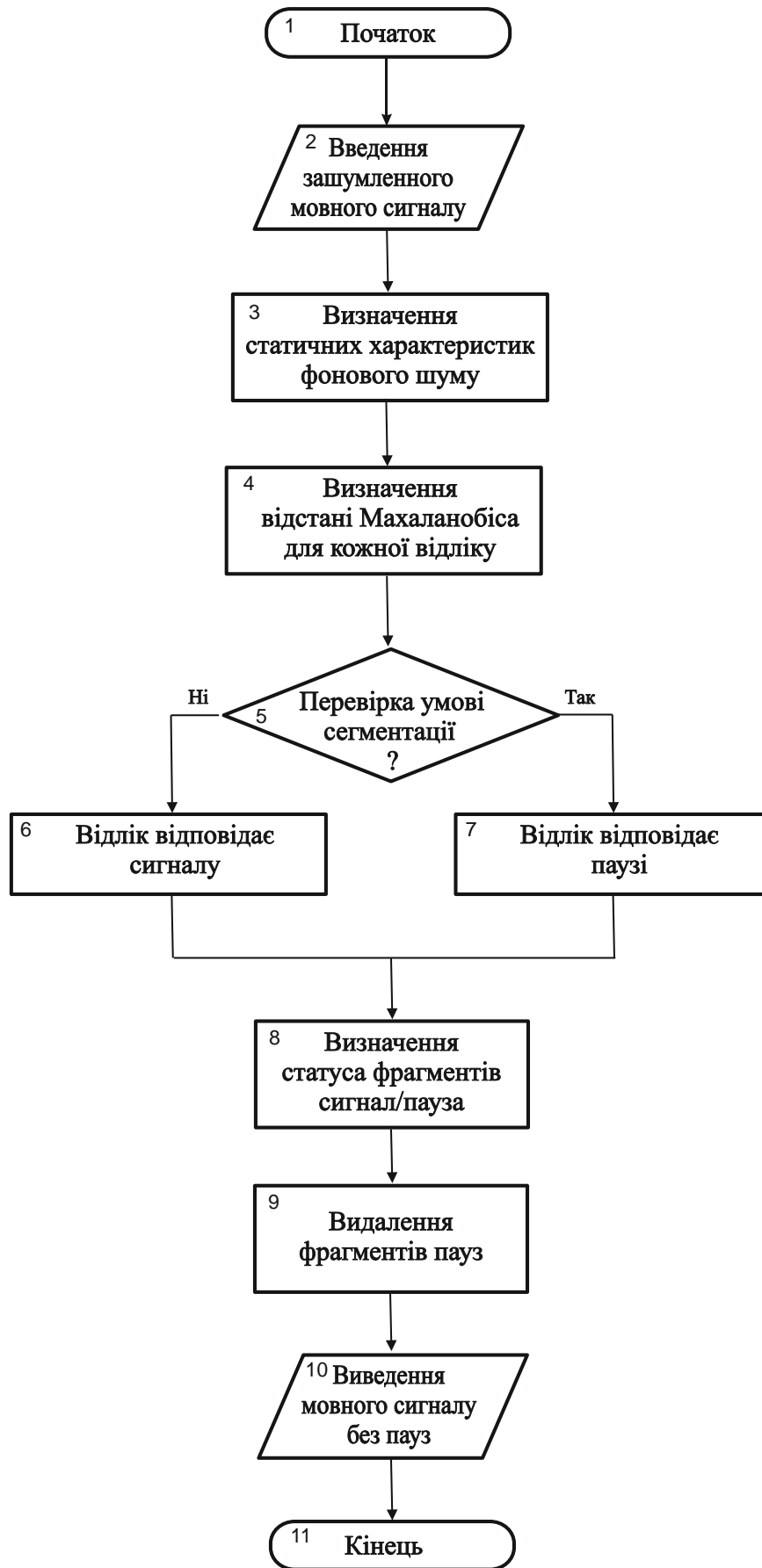


Рисунок 2.6 - Алгоритм сегментації на основі відстані Махаланобіса

Блок 2. Введення зашумленої мовної команди  $x(n)$  ( $n$  - дискретний відлік часу  $0 < n \leq N$ ,  $N$  - число дискретних відліків у мовному сигналі).

Блок 3. Мовна команда - процес, спектр залишається незмінними протягом короткого періоду часу. Це дозволяє розділити сигнал на фрагменти, у яких сигнал можна вважати стаціонарним.

Блок 4. Після завершення сегментації мовної команди отримаємо набір фрагментів, номер фрагмента -  $s$ , загальна кількість фрагментів -  $S$ . Подальша робота алгоритму пристрою здійснюється окремо з кожним фрагментом, встановлюється номер фрагмента.

Блоки 5, 6. Комплементарна множинна декомпозиція на емпіричні моди фрагмента. Встановлюється номер емпіричної моди фрагмента, подальша робота алгоритму пристрою здійснюється окремо з кожною емпіричною модою.

Блок 7. Зміни амплітуди мовної команди її характеру в часі - амплітудний розподіл, важливий інформативний параметр мовлення. Амплітуда мовної команди ділянок пауз чи невокалізованих участків менша за амплітуду вокалізованих ділянок. Зміна амплітуди мовного сигналу описується з використанням короткочасної функції сигналу.

Слуховий апарат людини фіксує мову нелінійно, різниця між енергіями паузи участків корисного сигналу є значною, щоб людина була в змозі фіксувати зміну амплітуди. Таким чином, для збільшення амплітуди в два рази енергію необхідно збільшити в вісім раз. Для проведення стиснення амплітуди мовного сигналу за великого динамічного діапазону використовують логарифмування енергії фрагментів, максимально наближаючи, при цьому роботу алгоритму пристрою до людини функціоналу слухового апарату.

Блоки 8, 9. Після завершення проведення аналізу першої емпіричної моди проводиться порівняння номера моди з загальною кількістю (блок 8). Якщо, при цьому умова не виконується, то встановлюється номер емпіричної моди на одиницю збільшений (блок 9) і в даному випадку здійснюється перехід до блоку 7 для проведення аналізу наступної емпіричної моди, при виконанні умови, здійснюється перехід до блоку 10.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						40
Зм	Арк	№ докум.	Підпис	Дата		

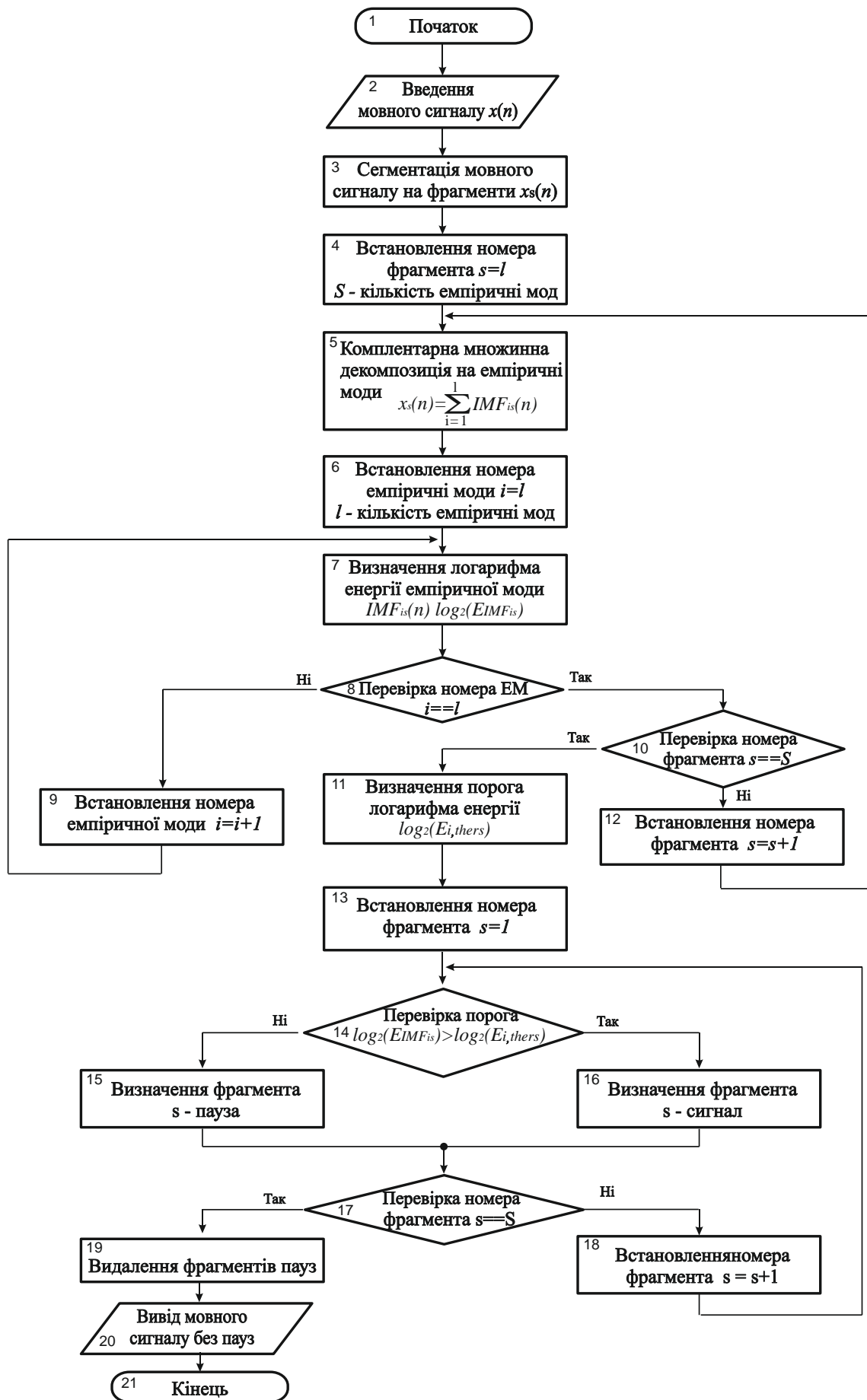


Рисунок 2.7 – Алгоритм пристрою сегментації сигнал/пауза

Зм	Арк	№ докум.	Підпис	Дата

Блоки 10, 12. Після завершення проведення аналізу всіх емпіричних мод проводиться порівняння номера фрагмента з загальною кількістю (блок 10). Якщо, при цьому умова не виконується, то встановлюється наступний номер фрагмента (блок 12) і здійснюється перехід до блоку 5 для проведення аналізу наступного фрагмента. Якщо умова виконується, перехід до блок 11.

Блок 11. Людина робить короточасну паузу перед вимовою голосової команди - 200 і більше мс. Дана ділянка відповідає тиші з фоновим шумом, не містить мовного сигналу, тому значення логарифмів енергії емпіричної моди фрагментів відповідатимуть значенням паузи перших 200 мс. Таким чином, використовуючи значення усереднених логарифмів енергії емпіричної моди, можна, відповідним чином сформуванати логарифмів енергії порогові значення для виявлення границь паузи і корисного сигналу.

Блоки 13. Після того як визначенні порогові значення логарифмів енергії для проведення порогової обробки фрагментів мовного сигналу рівним одиницю встановлюється номер фрагмента ( $s = 1$ ).

Блоки 14-16. Наступним етапом є виконання порівняння порогових значень значень зі значеннями логарифмів енергії емпіричних мод інших фрагментів. Якщо умова, в даному випадку виконується, то фрагмент відноситься до корисного сигналу, якщо не виконується, то фрагмент відноситься до пауз.

Блоки 17, 18. Після порогової обробки фрагмента здійснюється перевірка завершеності обробки, порівнянням номера фрагмента з загальною кількістю (блок 17). Якщо, при цьому, умова не виконується, то номер фрагмента збільшується на одиницю (блок 18) і виконується перехід до блоку 14 для проведення аналізу наступного фрагмента сигналу, якщо виконується, перехід до блоку 19.

Блоки 19, 20. Після визначення статусу всіх фрагментів та порогової обробки, фрагменти пауз в подальшому не використовуються. На виході алгоритму пристрою сегментації формується мовна команда без пауз.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						42
Зм	Арк	№ докум.	Підпис	Дата		

## 2.5 Висновки

1. Проведено аналіз частотно-вибіркових властивостей методів декомпозиції на емпіричні моди, показано недоліки та переваги, визначено відповідно до задач обробки мовних сигналів у підсистемі голосового управління керуючих систем, шляхи їх подолання.

2. Запропонований технологічний процес обробки мовних команд та показано основні етапи роботи у підсистемі голосового управління керуючих систем.

3. Запропонований алгоритм фільтрації з використанням підходу комплементарної множинної декомпозиції на емпіричні моди, нестационарних фонових шумів, використовується аналіз енергетичних та шумових характеристик емпіричних мод.

4. Запропонований алгоритм сегментації сигнал/пауза з використанням підходу комплементарної множинної декомпозиції на емпіричні моди, ґрунтується на проведенні аналізу логарифму енергетичних характеристик емпіричних мод з визначенням порогових значень паузи та сигналу.

5. Запропонований алгоритм вимірювання частоти основного тону з використанням підходу комплементарної множинної декомпозиції на емпіричні моди, базується на проведенні аналізу логарифму енергетичних характеристик емпіричних мод, визначенням моди, що містить основний тон.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						43
Зм	Арк	№ докум.	Підпис	Дата		

### 3 РЕАЛІЗАЦІЯ ПІДСИСТЕМИ ГОЛОСОВОГО УПРАВЛІННЯ КЕРУЮЧИХ СИСТЕМ

#### 3.1 Підсистема голосового управління у складі керуючих систем

У кожному житловому сучасному приміщенні функціонують інженерні системи, забезпечують безпеку, зв'язок та побут. Керуюча система забезпечує управління інженерними системами, можливість злагоджено працювати, їх інтеграція, підвищуючи, при цьому функціональність. Керуюча система – автоматизований комплекс, координує роботу інженерних систем: інженерна система безпеки; інженерна кліматична система; інженерна система управління; інженерна енергетична система; інженерна система механізації; інженерна система зв'язку та моніторингу.

На рис. 3.1 представлено структурну схему керуючої системи домашньої автоматизації.

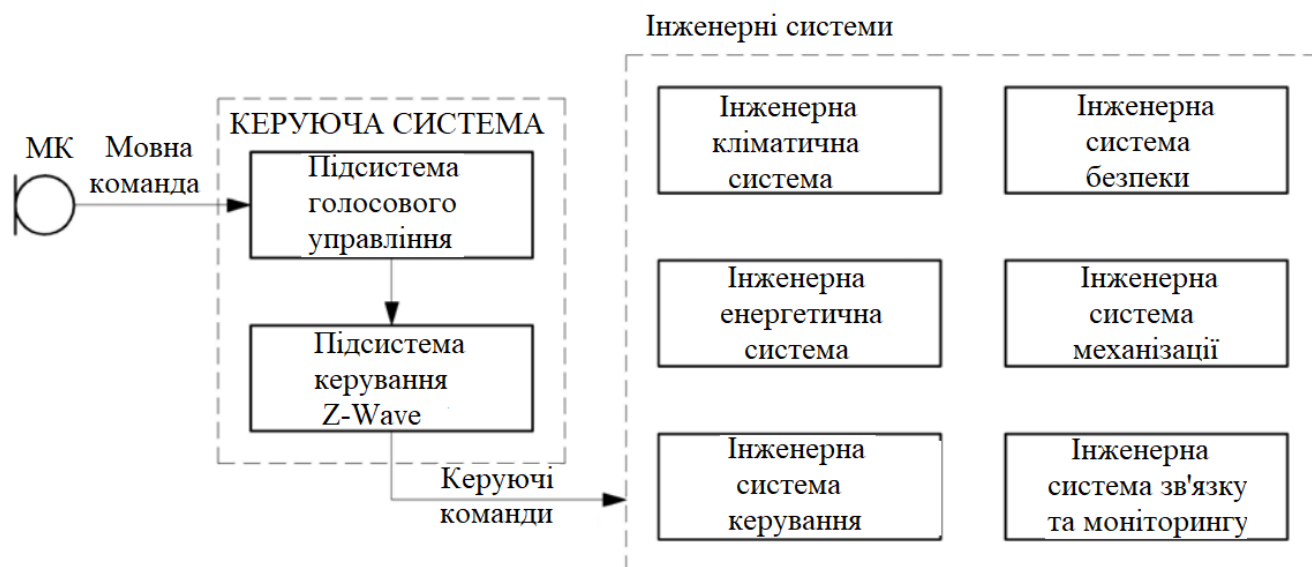


Рисунок 3.1 - Структурна схема керуючої системи домашньої автоматизації

Підсистема голосового управління керуючої системи домашньої автоматизації здійснює з допомогою мовних команд управління інженерними системами. Мовна команда реєструється з використанням мікрофона, обробляється в системі голосового управління та надходить в подальшому на підсистему управління Z-

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						44
Зм	Арк	№ докум.	Підпис	Дата		

Wave, відбувається перенапрямок та ідентифікація у вигляді керуючих команд для виконання в інженерній системі.

Взаємодія інженерних систем з керуючою системою здійснюється із застосуванням спеціалізованих протоколів та платформ двох видів: бездротові та провідні.

Найбільш практичну популярність серед провідних протоколів та платформ здобули:

1. EIB («Європейська інсталяційна шина» – European Installation Bus) – стандартний протокол через загальну мережу (шину) передачі даних. Швидкість передачі 9600 біт/с. Працездатність мережі забезпечують інтерфейсні модулі, блоки живлення, повторювачі та шинні з'єднувачі. Максимальна кількість, що підключаються до мережі, виконавчих пристроїв до 58 000.

2. Lan Drive – спеціалізована платформа побудови шинних систем управління. У платформі застосовується обладнання, підтримує швидкість передачі даних до 115200 біт/с. Платформа складається з виконавчих модулів, головного контролера, пов'язаних між собою шиною (загальною мережею). Максимальна кількість, що підключаються, виконавчих пристроїв, керованих одним головним контролером може досягати до 100. Можливе об'єднання головних контролерів декількох в одну мережу.

3. Smart-Bus – розподілена платформа побудови мереж керуючої системи домашньої автоматизації. Швидкість передачі потоку даних 9600 біт/с. Управління здійснюється виконавчими рівнозначними пристроями, які знаходяться у різних місцях платформи, не потребує головного контролера. Максимальна кількість виконавчих пристроїв може досягати 65280.

Практичну популярність серед бездротових протоколів та платформ здобули:

1. X10 - закритий бездротовий протокол передачі потоку даних, призначений для управління інженерної енергетичної системи, виконавчими пристроями. Швидкість передачі потоку даних 9600 біт/с на частоті 433 МГц у Європі та 310 МГц у США. Існує проблема з хибним спрацьовуванням у бездротового з'єднання.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						45
Зм	Арк	№ докум.	Підпис	Дата		

Можливі, також конфлікти пристроїв від різних виробників, що працюють за бездротовим протоколом X10.

2. ONE-NET – відкритий бездротовий протокол передачі потоку даних. Швидкість передачі потоку даних до 230 кбіт/с на частотах 868 МГц. Максимальна кількість з можливістю розширення мережі 212 виконавчих пристроїв. ONE-NET протокол може працювати від зовнішнього акумулятора та розрахований на низьке споживання енергії.

3. Z-Wave – запатентований бездротовий протокол передачі потоку даних. Швидкість передачі потоку даних 40 кбіт/с на частоті 869 МГц. Особливість протоколу полягає у використанні малопотужних та мініатюрних радіочастотних модулів, що вбудовуються у виконавчі пристрої. Максимальна кількість з можливістю розширення мережі виконавчих пристроїв 232.

Проведений аналіз бездротових і провідних протоколів та платформ слідує, що найбільш оптимальним варіантом передачі даних для управління є використання бездротового зв'язку. У табл. 3.1 представлені основні характеристики та параметри протоколів w Z-Wave, ONE-NET, X10.

Аналіз характеристиків та параметрів протоколів, показав, що найбільш для бездротового зв'язку оптимальним є використання протоколу Z-Wave.

Z-Wave є повністю бездротовою технологією, в основі якої лежить коміркова mesh-мережа. Кожен пристрій мережі Z-Wave є трансівером, збільшує надійність мережі. При виході зі строю одного з пристроїв мовний сигнал піде через сусідній пристрій. Зона покриття розширюється додаванням нових пристроїв, які можуть працювати, в даному випадку, як повторювачі сигналу керуючої команди. Для цього, достатньо, щоб пристрій мережі Z-Wave знаходився в радіусі дії сусіднього пристрою мережі.

Індивідуальний код має кожен пристрій мережі Z-Wave, відомий головному контролеру, забезпечує можливість програмування відповідних функцій управління пристроями. Таким чином, щоб створити мережу достатньо мати головний контролер та один виконавчий пристрій. Після побудови та запуску мережі до неї

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		46

можна підключити додаткові пристрої Z-Wave. Це надає змогу будувати систему управління зі зростанням потреб.

Таблиця 3.1 - Характеристики та параметри протоколів X10, ONE-NET, Z-Wave

Параметри	Протоколи		
	X10	ONE-NET	Z-Wave
Необхідність прокладання нових кабелів	ні	ні	ні
Робочий діапазон частот, МГц	433	865 - 868	869
Швидкість передачі інформації	1 команда за 0,7 – 1 с	38,4 – 230 кб/с	до 40 кб/с
Потужність, Вт	0,2	0,5	0,3
Максимальна кількість виконавчих пристроїв у мережі	64	212 з можливістю розширення мережі	232 з можливістю розширення мережі
Сумісність	Можливі конфлікти пристроїв різних виробників	Можливі конфлікти пристроїв різних виробників	Z-Wave Alliance відкритий консорціум
Стійкість до фізичних перешкод	ні	ні	Пошук обхідних шляхів, обумовлений комерційною структурою мережі

Технологія Z-Wave доступна у вигляді модуля з інтегрованим радіочастотним трактом та антеною а також у вигляді плати розширення Single Chip (рис. 3.2).

У платі розширення Z-Wave використовується універсальний асинхронний приймач-передавач, контролер Triac, флеш пам'ять, SDRAM, процесор 8051, введення/виведення, радіомодуль, АЦП. Система управління на основі технології Z-Wave є mesh-мережею з бездротовим дуплексним радіозв'язком.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						47
Зм	Арк	№ докум.	Підпис	Дата		

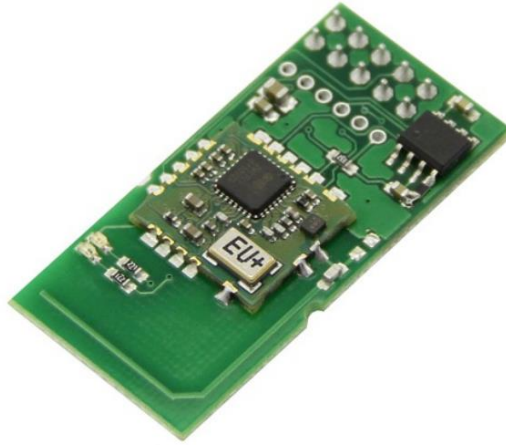


Рисунок 3.2 - Зовнішній вигляд плати розширення Single Chip Z-Wave

Компанія Sigma Designs, якій належить технологія Z-Wave, видає ліцензію на конструювання, інтерфейси прикладного програмування та стекове програмне забезпечення (Application Programming Interface, API) для виробників Z-Wave пристроїв, що гарантує прискорену та якісну розробку виробів. На теперішній час технологія Z-Wave є світовим лідером на ринку бездротового управління.

На основі розроблених алгоритмів та пристроїв обробки, розроблена загальна схема обробки мовних сигналів.

Загальна схема обробки мовних команд підсистеми голосового управління керуючих систем заснована на алгоритмах і пристроях обробки, складається з наступних дій: фільтрації нестационарних шумів фонових, сегментації сигнал/пауза, вимірювання частоти основного тону (рис. 3.3).

Особливість запропонованої схеми полягає в оптимізації повторюваних у кожному алгоритмі блоків обробки комплементарної множинної декомпозиції на емпіричні моди та сегментації на фрагменти (рис. 3.3 блоки з номерами 3/8/11, 4/7/14) та у можливості виконання одночасно декількох етапів (блоки 12-15, 16-19). Розглянемо етапи обробки мовних команд.

Блоки 1-5 - фільтрація, в рамках якої здійснюється: декомпозиція мовного сигналу на емпіричні моди; введення мовного сигналу з використанням пристроїв реєстрації; сегментація емпіричної моди на короткочасні фрагменти, для більш докладного аналізу характеристик емпіричної моди; фільтрація мовного сигналу з використанням моди з основним шумом; визначення моди з основним шумом на

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						48
Зм	Арк	№ докум.	Підпис	Дата		

основі обчислення шумових та енергетичних характеристик емпіричної моди та подальшої порогової обробки.

Блоки 6-9 - сегментація сигнал/пауза в яких здійснюється: декомпозиція фрагмента на емпіричні моди методом комплементарної множинної декомпозиції на емпіричні моди; сегментація мовного сигналу на короточасні фрагменти для визначення фрагментів сигналу та пауз та проведення більш докладного аналізу; пошук фрагментів, які містять паузу та сигнал з використанням обчислення логарифму енергетичних характеристик емпіричної моди та порогової обробки; видалення фрагментів, що містять паузу та визначення границь сигналу мовної команди.



Рисунок 3.3 – Загальна схема обробки мовних команд підсистеми голосового управління керуючих систем

Блок 10-14 є етапом вимірювання частоти основного тону, в яких здійснюється: визначення емпіричної моди з основного тону за використанням обчислення логарифму енергетичних характеристик емпіричної моди та порогової обробки; декомпозиція мовного сигналу без пауз на емпіричні моди методом комплементарної множинної декомпозиції на емпіричні моди; фільтрації емпіричної

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		49

моди з основного тону за використанням фільтра нижніх частот; вимір частоти основного тону кожного фрагмента; сегментація емпіричної моди з основного тону на короткочасні фрагменти з метою більш докладного аналізу.

Блоки 15-19 є етапом визначення та розпізнавання інформативних сигналів, в рамках яких здійснюється: формування набору інформативних сигналів - різниця мовної команди вихідного сигналу та емпіричною модою, з метою визначення корисної інформації мовної команди для розпізнавання; декомпозиція мовного сигналу без пауз на емпіричні моди; формування набору даних мовної команди і шаблонів бази даних; розпізнавання за використанням порівняння мовної команди з шаблоном з бази даних. Загальна схема є інструментом аналізу та обробки мовних команд.

### 3.2 Апаратна реалізація підсистеми голосового управління керуючої системи

Представлено апаратну реалізацію підсистеми голосового управління керуючої системи домашньої автоматизації. На рис. 3.4 наведена структурна схема підсистеми у складі керуючої системи домашньої автоматизації, яка складається з модулів: реєстрації, обробки, візуалізації, живлення, аудіо та Z-Wave.

Реєстрація мовних команд сигналів функціонально реалізована за використанням модуля реєстрації, що складається з аналогового підсилювача та мікрофонів. Мікрофон (МК 1) призначений для реєстрації команд, МК 2 - для шумоподавлення. Алгоритми та пристрої обробки мовних команд реалізовані апаратно та програмно, інтегровані в мікроконтролер Broadcom BCM2836, який виконує операції управління, контроль режимами роботи підсистеми голосового управління, обчислювальні операції. Мікроконтролер із додатковим периферійним устаткуванням функціонально реалізований як модуль обробки. Периферійне обладнання модуля забезпечує нормальну працездатність підсистем голосового управління відповідно до функціоналу.

Режими роботи передбачають, підсистеми голосового управління, попереднє навчання та налаштування. Попереднє налаштування здійснюється за рахунок

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						50
Зм	Арк	№ докум.	Підпис	Дата		

використанням сенсорного TFTдисплея для відображення та введення інформації про налаштування, реалізований у виді модуля візуалізації.

Аудіо модулем реалізується контроль режимів роботи системи, здійснює звукову індикацію режимів роботи.

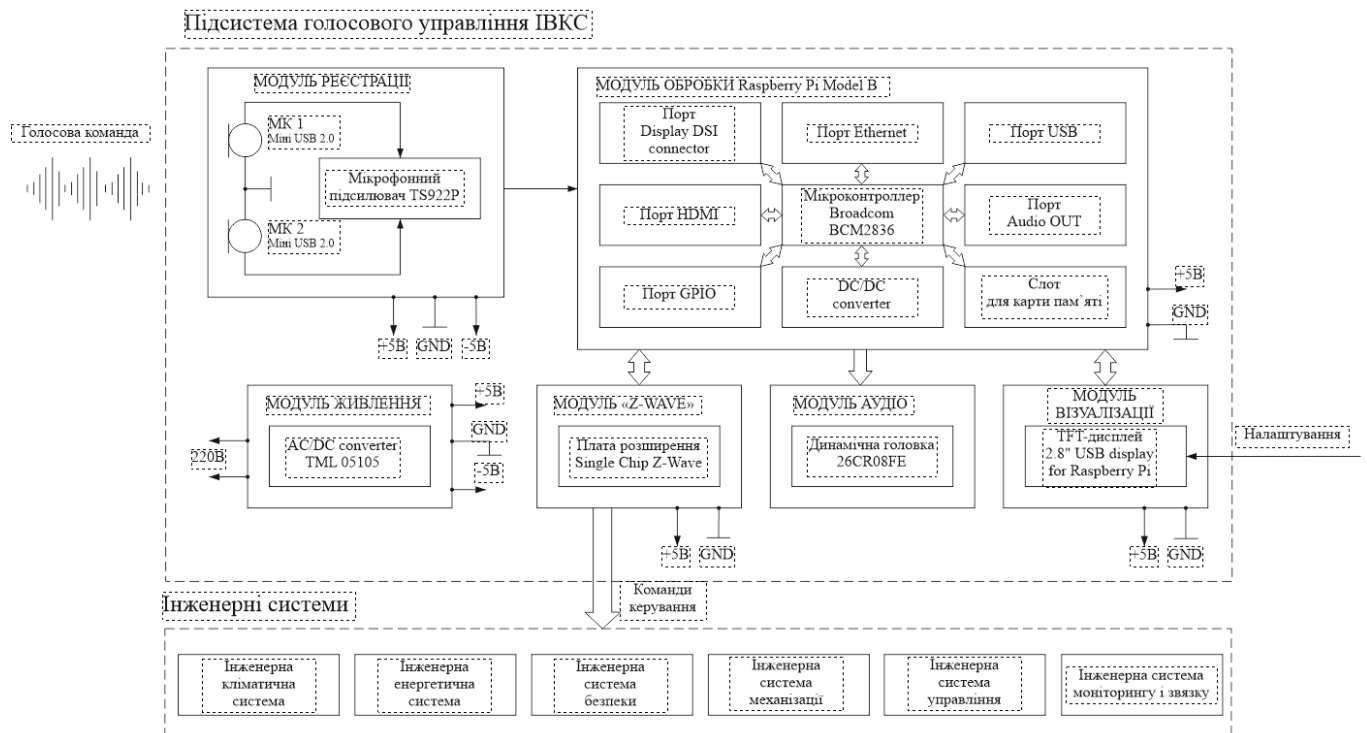


Рисунок 3.4 - Структурна схема підсистеми голосового управління у складі керуючої системи

Для забезпечення пристрою напругою живлення використовується модуль живлення TML05105. Для забезпечення безперебійної роботи підсистеми управління використовується модуль живлення, який виконує функції перетворення напруги.

Модуль обробки - основний компонент підсистеми голосового управління. При виборі модуля обробки враховувалися внутрішня архітектура та функціональні можливості. Вибір модуля обробки залежить від наявності (відсутності) системи забезпечення розробки (відповідного обладнання, програмного середовища), плат розширення для взаємодії з пристроями, інтерфейсів підключення периферійного обладнання які працюють за протоколом Z-Wave, ціни та розмірів. На теперішній час існують багатofункціональні налагоджувальні платформи, що використовуються

						<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата			51

для розробки на мікроконтролерах цифрових додатків, найбільш популярні: Arduino Uno R3, Beagle Bone-Black, Raspberry Pi.

Beagle Bone-Black - налагоджувальна платформа з архітектурою ARM Cortex-A8 від Texas Instruments на мікроконтролері AM3359 з тактовою частотою 1 ГГц, призначена для розробників, вивчення операційної системи Linux. До складу платформи входить оперативна пам'ять 512 МБ DDR3L, слот microSD карти, 2 ГБ eMMC пам'яті. Для підключення зовнішніх пристроїв призначенні інтерфейси USB, Ethernet, HDMI та роз'єми розширення. Платформа Beagle Bone-Black працює під управлінням ОС Linux Angstrom, встановлену у вбудовану eMMC Flash пам'ять. Плата підтримує ОС Ubuntu, Android і Fedora.

Arduino Uno R3 - налагоджувальна платформа на основі мікроконтролера ATmega 328. До складу платформи входить необхідне для зручної роботи з мікроконтролером: 6 аналогових входів, 14 цифрових входів/виходів, кварцовий роз'єм USB, резонатор на 16 МГц, роз'єм живлення, кнопка скидання, роз'єм для внутрішньосхемного програмування ICSP. Для початку роботи з пристроєм достатньо подати живлення від батарейки, АС/DC конвертора, або підключити його до комп'ютера за допомогою USB.

Raspberry Pi Model B - налагоджувальна платформа, одноплатним мініатюрним дешевим комп'ютером, на мікроконтролері Broadcom BCM2836, з тактовою частотою 0,7 ГГц, з архітектурою ARM 11 (можливість розгону процесора до 1 ГГц). До складу платформи входить Ethernet 10/100 Мбіт, система введення-виведення загального призначення - порт GPIO для керування зовнішніми пристроями, 512 Мб SDRAM оперативної пам'яті, два порти USB, слот для картки пам'яті SD/MMC/SDIO. Підтримуються наступні операційні системи: Raspbian (Debian, під архітектуру ARM), встановлений інтерпретатор Python, веб-браузер Midori.

Всі налагоджувальні платформи виконані в компактному форм-факторі, знаходяться в доступному ціновому діапазоні, можуть бути використанні для створення цифрових пристроїв. У табл. 3.2 наведено порівняльні характеристики платформ.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		52

Для апаратної реалізації підсистеми голосового управління, як видно з табл. 3.2 найбільш оптимальним варіантом є платформа налагодження Raspberry Pi Model B. Зовнішній вигляд платформи Raspberry Pi Model B представлений на рис. 3.5.

На даному етапі є одним із найпопулярніших одноплатний мініатюрний комп'ютер Raspberry Pi Model B, відкриває унікальні можливості перед програмістами, розробниками, інженерами для реалізації технічних проектів.

Таблиця 3.2 - Порівняння основних характеристик налагоджувальних платформ

Платформа	Arduino Uno R3	Beagle Bone-Black	Raspberry Pi Model B
Габаритні розміри	7,5 x 5,4	8,6 x 5,3	8,6 x 5,4
Мікроконтролер	ATmega 328	AM3359 з архітектурою ARM Cortex-A8	Broadcom BCM2836 з архітектурою ARM 11
Тактова частота ГГц	0,016	1	0,7 – 1,0
ОЗП, Мб	128	512	512
Flash-пам'ять	-	SD-карта	SD-карта
Напруга живлення, В	7 - 12	5	5
Мінімальне енергоспоживання, Вт	0,3	0,85	3,5
Інструменти розробки	Arduino IDE	Phyton, Scratch, Squeak, Cloud9/Linux	IDLE, Scratch, Squeak/Linux
Порт "Ethernet"	-	10/100 Мб RJ45	10/100 Мб RJ45
Інтерфейс «USB»	-	USB 2.0	2×USB 2.0
Відео вхід/вихід	-	HDMI	HDMI
Аудіо вхід/вихід	-	Через HDM	3,5 мм

Мінікомп'ютера Raspberry Pi Model B успішно може використовуватись у системах домашньої автоматизації і дозволяє вирішити широкий спектр задач. Реалізація технічних проектів на базі платформи Raspberry Pi Model B - досить проста при програмно-апаратній технічній підтримці. Мінікомп'ютер Raspberry Pi

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						53
Зм	Арк	№ докум.	Підпис	Дата		

Model B надає можливість реалізувати готові проєкти адаптувати та покращити їх під власні потреби.



Рисунок 3.5 - Платформи налагодження Raspberry Pi Model B

Модуль візуалізації. TFT-дисплеї – засновані на тонкоплівковому транзисторі активні екрани. У TFT-дисплеях є можливість керувати кожною точкою екрана використовується транзистор, у пасивних дисплеях використовується сітка горизонтальних і вертикальних зв'язків. Активні панелі мають більш яскраве зображення, більший кут огляду, чітке, чисте ніж пасивні.

При виборі TFT-дисплея необхідно виходити з раніше обраним модулем обробки на базі мінікомп'ютера Raspberry Pi з функціональних можливостей взаємодії. Вибір також залежить від роздільної здатності екрану, діагоналі, розмірів, яскравості та робочої температури, напруги живлення. На теперішній час існує достатньо TFT-дисплеїв, з можливостями взаємодії з мінікомп'ютером Raspberry Pi Model B, найбільш популярний: від компанії RoboPeak Mini USB Display 2.8" USB TFT display for Raspberry Pi. На рис. 3.6 наведено зовнішній вигляд TFT-дисплея, в табл. 3.3 основні характеристики і параметри.

Модуль реєстрації. При проектуванні модуля реєстрації необхідно виходити з принципу дії роботи мікрофонів. Дії мікрофонів за принципом: п'єзоелектричні, динамічні (електродинамічні), вугільні, конденсаторні.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		54

Таблиця 3.3 - Основні характеристики та параметри 2.8" USB TFT display for Raspberry Pi

Найменування параметра та характеристики	Дані параметра та характеристики
Діагональ екрану, дюйм	2,8
Роздільна здатність екрану, піксель	320x240
Розміри, мм	74x60x4
Тип підсвічування	LED
Вхідний інтерфейс	USB Full-Speed
Опції та склад комплекту	сенсорна панель, резистивна
Напруга живлення, В	3,3-5
Яскравість, кд/м <sup>2</sup>	350
Контраст	400 : 1
Глибина кольору, біт	16
Робоча температура, °С	від -10 до +60
Температура зберігання, °С	від -20 до +60

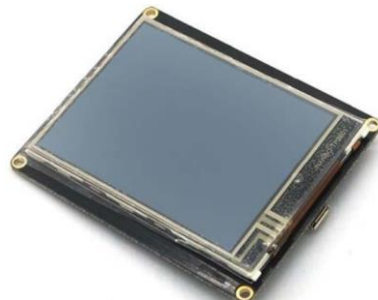


Рисунок 3.6 - Зовнішній вигляд TFT-дисплея

Вибір мікрофона залежить від динамічного діапазону, частотного, його спрямованості та нерівномірності. У таблиці 3.4 наведені порівняльні показники мікрофонів.

Для апаратної реалізації модуля реєстрації, як видно з табл. 3.4, оптимальним варіантом є конденсаторний мікрофон, має великі динамічні і частотні діапазони, більшою спрямованістю і найменшою нерівномірністю.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		55

Таблиця 3.4 – Порівняльні характеристики чотирьох типів мікрофонів

Характеристики	Види мікрофонів			
	Вугільний	Динамічний	Конденсаторний	П'єзоелектричний
Частотний діапазон, Гц	300 – 3400	30 – 15000	20 – 20000	100 - 7000
Нерівномірність частотної характеристики	30	10-12	5	15
Динамічний діапазон, дБ	20 – 30	11 – 160	11 – 160	до 80
Спрямованість, градус	65 – 360	45 – 180	45 – 360	90 – 360

При використанні конденсаторного мікрофона необхідно врахувати, що через особливості для мінікомп'ютера Raspberry Pi Model B немає можливості використовувати мікрофон зі стандартним роз'ємом аудіо TRS 3,5 мм. На сьогодні існують конденсаторні мікрофони, які використовують вихід порт USB, можна виділити найбільш популярний: Міні USB 2.0 мікрофон MIC. Особливістю мікрофона є наявність підсилювача вбудованого мікрофонного з високим рівнем чутливості. На рис. 3.7 наведений зовнішній вигляд Міні US B 2.0 мікрофон MIC. В табл. 3.5 основні характеристики і параметри.

Таблиця 3.5 - Характеристики та параметри Міні USB 2.0 мікрофон MIC

Найменування параметра та характеристики	Дані параметра та характеристики
Частотний діапазон, Гц	від 20 Гц до 20000
Якість запису, біт	16
Відношення сигнал/шум, дБ	96
Чутливість, дБмВ	- 67
Робоча температура, °С	від -10 до +60
Температура зберігання, °С	від -20 до +60

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		56



Рисунок 3.7 - зовнішній вигляд Міні US B 2.0 мікрофон MIC

Модуль живлення. Конвертори AC/DC призначені для перетворення змінної напруги 220В/50 Гц мережі на постійну напругу. Для забезпечення надійної роботи підсистеми голосового управління керуючої системи домашньої автоматизації необхідно задіяти стабілізований конвертор AC/DC з номінальним струмом не менше 1А та вихідною напругою 5В. Більшість конверторів знаходяться в одному ціновому сегменті, мають схожі характеристики, для підсистеми голосового управління важливо зберігати компактний форм-фактор.

Вибір конвертора AC/DC залежить від наступних параметрів: струму, вихідної та вхідної напруги, коефіцієнта корисної дії, габаритних розмірів, діапазону робочих температур, вихідної потужності. У табл. 3.6 наведені характеристики та параметри конверторів AC/DC п'яти провідних виробників.

Для апаратної реалізації модуля живлення, як показано в табл. 3.6 оптимальним варіантом є конвертор TRACO POWER TML 05105. Зовнішній вигляд конвертора AC/DC наведений на рис. 3.8.



Рисунок 3.8 – Зовнішній вигляд конвертора TRACO POWER TML 05105

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		57

Таблиця 3.6 - Параметри та характеристики AC/DC конверторів

Найменування параметру	TRACO POWER TML 05105	PEAK PPM03-S- ELF	POWER LINE RAC04- 05SA	MEAN WELL NFM-05-5	XP POWER VCP05
Вхідна змінна напруга, В	90 - 264	85 - 264	90 - 264	85 - 264	90 - 264
Частота вхідної напруги, Гц	47 - 440	47 - 440	47 - 440	47 - 440	47 - 63
Вихідна напруга, В	5±1%	5±2%	5±3%	5±2%	5±5%
Вихідний струм, А	2,0	0,3 – 1,2	0,8	1,0	1,0
Вихідна потужність, Вт	10	2,3 – 3,0	4	5	5
Напруга шумів та перешкод, мВ	> 10	> 100	> 200	> 80	> 150
Діапазон робочих температур, °С	від -30 до +71	від -27 до +70	від -40 до +70	від -20 до +70	від -20 до +60
ККД, %	84	79	72	71	77
Габаритні розміри, мм	52,4x27,2x23,0	37x23x15	36,7x21.0 x 27,0	58x45x19	58,5x17x35,5

Модуль аудіо. Динамічна головка - гучномовець, в якому відбувається перетворення електричного сигналу на звуковий завдяки переміщенню котушки зі струмом в постійному магніту магнітному полі, з перетворенням отриманих механічних коливань за допомогою дифузора в коливання навколишнього повітря.

Для модуля аудіо вибір динамічної головки залежить від параметрів: діапазону частот, номінального опору, габаритних розмірів, граничної потужності.. У табл. 3.7 наведені характеристики та параметри динамічних головок.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		58

Таблиця 3.7 - Характеристики та параметри динамічних головок

Найменування параметру	0,5 ГДШ 36-1-16	20CS08GI 4-V38ND	15CR08F	26CR08FE
Номінальний діаметр, мм	36	20	15	26
Номінальний опір, Ом	16	8	8	8
Діапазон відтворюваних частот, Гц	400 – 5000	600 – 20000	500 – 5000	450 - 20000
Гранична потужність, Вт	1	0,5	0,3	0,5

Для апаратної реалізації аудіомодуля, як наведено в табл. 3.7 оптимальним варіантом є головка 26CR08FE. Вигляд динамічної головки наведено на рис. 3.9.

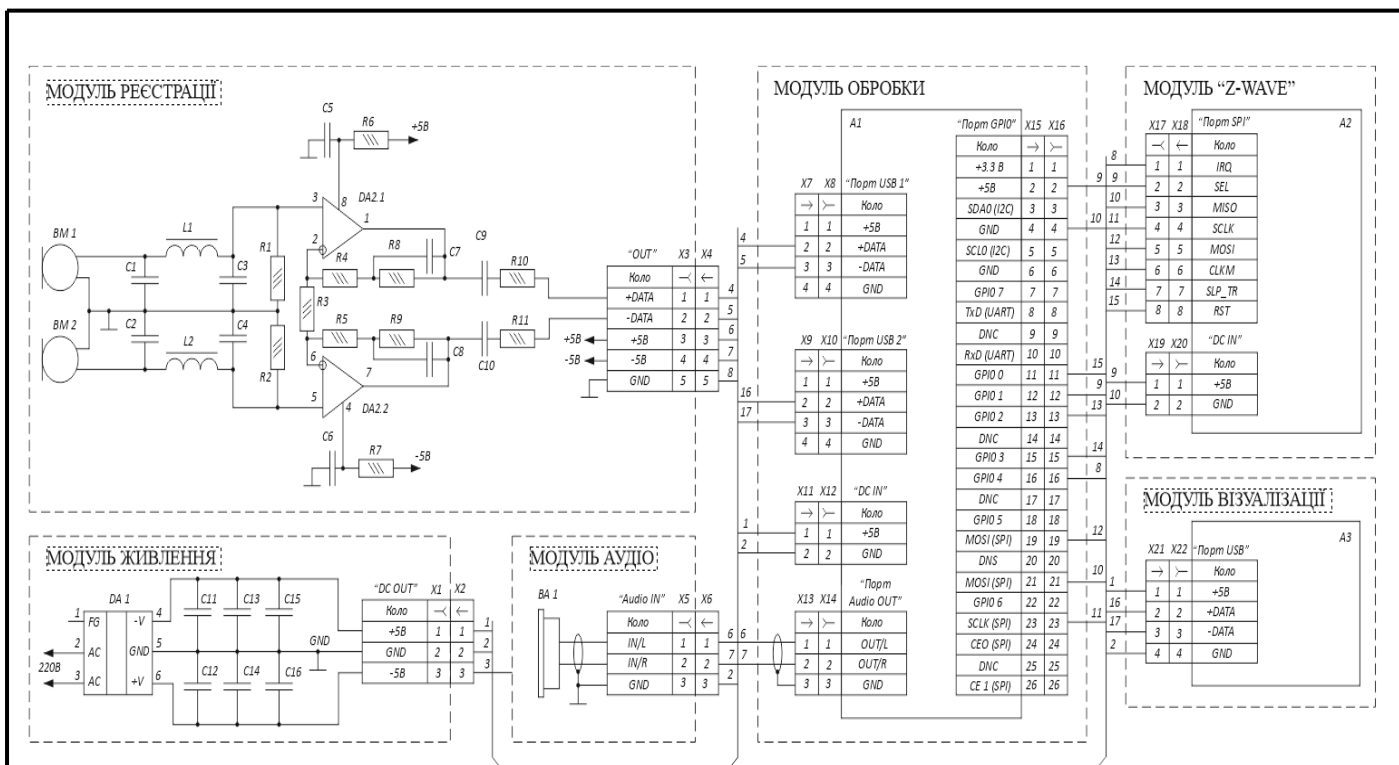


Рисунок 3.9 - Зовнішній вигляд динамічної головки 26CR08FE

На підставі обґрунтованих та вибраних елементів апаратної реалізації розроблено схему електричних з'єднань підсистеми голосового управління, яка використовується у складі керуючої системи домашньої автоматизації (рис.3.10).

Для забезпечення надійної роботи підсистеми голосового управління керуючої системи сумісності з персональним комп'ютером домашньої автоматизації, на основі використання мови об'єктно-орієнтованого програмування C# розроблено спеціальне програмне забезпечення, надбудова над програмним забезпеченням управління бездротовим протоколом передачі потоку даних Z-Wave.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		59



A1 - Мінікомп'ютер Raspberry Pi Model B  
A2 - Плата розширення Single Chip Z-Wave  
A3 - Дисплей 2.8" USB TFT display for Raspberry Pi Model B  
BA - Динамічна головка 26CR08FE  
BM 1, BM 2 - Мікрофон Міні USB 2.0 MIC  
DA 1 - AC/DC конвертер TRACO POWER TML 05105  
DA 2 - Операційний підсилювач TS922P

Рисунок 3.10 - Підсистема голосового управління керуючої системи  
схема електричних з'єднань

### 3.3 Висновки

1. Запропоновано підхід до обробки мовних сигнал підсистеми голосового управління керуючих систем, заснований на алгоритмах і пристроях обробки та оптимізовану послідовність дій: фільтрацію нестационарних фонових шумів, сегментацію сигнал/пауза, вимірювання чистот основного тону.

2. Розроблено структурну та електричну з'єднання схеми підсистеми голосового управління керуючої системи домашньої автоматизації за бездротовим протоколом передачі даних Z-Wave.

3. Апаратно реалізована підсистема голосового управління керуючої системи на базі модуля обробки – модуля візуалізації – TFT- дисплей 2.8" USB TFT display for Raspberry Pi; мінікомп'ютера Raspberry Pi Model B; модуля реєстрації -

мікрофона Міні USB 2.0 мікрофон MIC; модуля аудіо - динамічної головки 26CR08FE; модуля живлення - AC/DC конвертора TRACO POWER TML 05105.

4. Реалізована програмно підсистема голосового управління керуючої системи мовою об'єктно-орієнтованого програмування C# у вигляді програмного забезпечення голосового управління керуючої системи домашньої автоматизації, є надбудовою над програмним забезпеченням управління бездротовим протоколом передачі даних Z-Wave.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						61
Зм	Арк	№ докум.	Підпис	Дата		

## ВИСНОВКИ

У кваліфікаційній роботі проведено порівняльний аналіз існуючих систем голосового управління та застосовуваних підходів обробки мовних команд. Досліджено особливості обробки мовних команд у підсистемах голосового управління керуючих систем, проведено аналіз методів обробки мовних команд у підсистемі голосового управління керуючих систем.

У кваліфікаційній роботі вирішенні наступні задачі: запропоновано підхід до обробки мовних сигнал підсистеми голосового управління керуючих систем, заснований на алгоритмах і пристроях обробки та оптимізовану послідовність дій: фільтрацію нестационарних фонових шумів, сегментацію сигнал/пауза, вимірювання чистот основного тону.

Розроблено структурну та електричну схеми з'єднання підсистеми голосового управління керуючої системи домашньої автоматизації за бездротовим протоколом передачі даних Z-Wave.

Апаратно реалізована підсистема голосового управління керуючої системи на базі модуля обробки – модуля візуалізації – TFT – дисплей 2.8" USB TFT display for Raspberry Pi; мінікомп'ютера Raspberry Pi Model B; модуля реєстрації - мікрофона Міні USB 2.0 мікрофон MIC; модуля аудіо - динамічної головки 26CR08FE; модуля живлення - AC/DC конвертора TRACO POWER TML 05105.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
						62
Зм	Арк	№ докум.	Підпис	Дата		

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Голосовой інтерфейс // Вікіпедія. URL: <http://ru.wikipedia.org>.
2. Грод І.М. Аналіз ефективності деяких алгоритмів. Теорія і практика: навчальний посібник / І.М. Грод, С.В.Мартинюк, –Тернопіль:ТНПУ, 2017. – 64с.
3. Берко А. Ю. Системи баз даних та знань. Книга 1. Організація баз даних: навч. посібник [для студ. вищ. навч.закл.] / Берко А. Ю., Верес О. М., Пасічник В.В. – Львів : Магнолія, 2017. – 456 с.
4. Катренко А. В. Системний аналіз / А. В. Катренко. – Львів: Новий світ, 2019. – 396 с.
5. Комплекс мовних технологій Яндекс // Яндекс. URL: <https://tech.yandex.ru>.
6. Лавров Є. А. Математичні методи дослідження операцій : підручник / Є.А. Лавров, Л. П. Перхун, В. В. Шендрик. – Суми : Сумський державний університет, 2017. – 212 с.
7. Литвин В. В. Інтелектуальні системи / В. В. Литвин, В. В. Пасічник, Ю. В. Яцишин. – Львів: Новий Світ, 2015. – 406 с.
8. Надшвидке розпізнавання мови без серверів на реальному прикладі // Habrahabr. URL: <https://habrahabr.ru>.
9. Петрик М.Р. Моделювання програмного забезпечення : науково-методичний посібник / М.Р. Петрик, О.Ю. Петрик – Тернопіль: ТНТУ ім. І. Пулюя, 2015. – 200 с.
10. Розпізнавання мови для чайників//Habrahabr. URL:<https://habrahabr.ru>.
11. Ситнік Б. Т. Основи інформаційних систем і технологій: Навч. посібник. – Харків: УкрДУЗТ, 2019. – 175 с.
12. Табунщик Г. В. Проектування та моделювання програмного забезпечення сучасних інформаційних систем / Г. В. Табунщик, Т.І. Каплієнко, О.А. Петрова – Запоріжжя : Дике Поле, 2016. – 250 с.
13. Технології об'єктно-орієнтованого програмування: частина 1. Комп'ютерний практикум [Електронний ресурс]: навч. посіб. / В. І.Бендюг, Б. М. Комариста. – Київ: КПІ ім. І. Сікорського, 2018. – 225с.

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		63

14. Технології розпізнавання мови // ПостНаука. URL: <https://postnauka.ru>.

15. Що таке голосові технології // Теплиця соціальних технології. URL: <http://test.ru>.

16. Albahari Joseph. C# 8.0 in a Nutshell: The Definitive Reference/Joseph Albahari, Eric Johanssen – O'Reilly Media, 2020г. – 1104с.

17. Griffiths Ian, Programming C# 8.0: Build Cloud, Web, and Desktop Applications/ Ian Griffiths. – O'Reilly, 2020. – 800с.

18. Speak to Alexa on Products with Alexa Voice Service Built-In // Alexa Voice Service. URL: <https://developer.amazon.co>

					<i>КвРКІ.180243.18.02.16 ПЗ</i>	Арк.
Зм	Арк	№ докум.	Підпис	Дата		64

## ДОДАТОК А

(обов'язковий)

Лістинг програми «Визначення частоти основного тону мовного сигналу»

```
clear;
clc;
y = wavread ('Мовний сигнал для сегментації');
Y = fft (y, 512);
Pyy = y. * conj (Y) / 512;
plot(f,Pyy(1:65));
function allmode=eemd(Y,Nstd,NE) xsize=length(Y); dd=1:l:xsize; Ystd=std(Y);
Y=Y/Ystd;
TNM=fix(log2(xsize))-1; TNM2=TNM+2; для kk=1:1:TNM2, for ii=1:1:xsize,
llmode(ii,kk)=0.0; end end
for iii=1:1:NE, for i=1:xsize, temp=randn(1,1)*Nstd; X1 (i) = Y (i) + temp; end
for jj=1:1:xsize, mode(jj,1) = Y(jj); end
xorigin = XI; xend = xorigin; nmode = 1;
while nmode <= TNM, xstart = xend; iter = 1; while iter<=10,
[spmax, spmin, flag] = extrema (xstart); upper = spline (spmax (:, 1), spmax (:, 2),
dd); lower = spline (spmin (:, 1), spmin (:, 2), dd); meanul = (upper + lower) / 2; xstart =
xstart - meanul; iter = iter + 1; end
xend = xend - xstart; nmode = nmode + 1; for jj=1:1:xsize,
mode(jj,nmode) = xstart(jj);
end
end
for jj=1:1:xsize,
mode(jj,nmode+1)=xend(jj);
end
allmode=allmode+mode;
end
allmode=allmode/NE;
allmode=allmode*Ystd;
mode=eemd(y,2,50);
mode=mode;
subplot(5,2,[1 2]); plot(y)
Y = fft (mode (4, :), 512); Pyy = y. * conj (Y) / 512; T = i()()^(()*o4)/12*
subplot(2,2,1); plot(f,Pyy(1 :65))
Y = fft (mode (5, :), 512); Pyy = Y. * conj (Y) / 512;
subplot(2,2,2); plot(f,Pyy(1 :65))
Y = fft (mode (6, :), 512);
Pyy = Y. * conj (Y) / 512;
subplot(2,2,3); plot(f,Pyy(1 :65))
max(f,Pyy(1 :65));
```

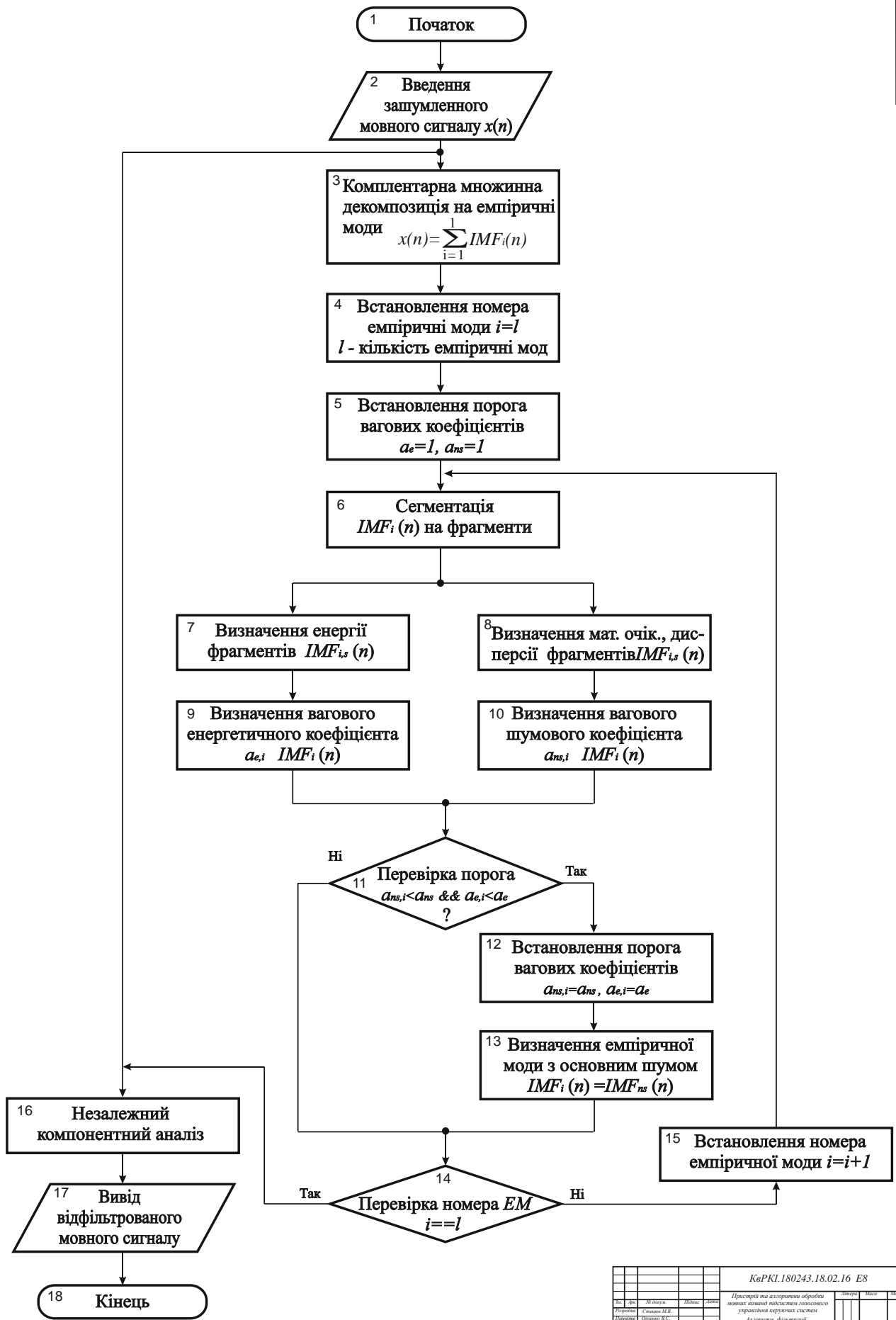
## ДОДАТОК Б

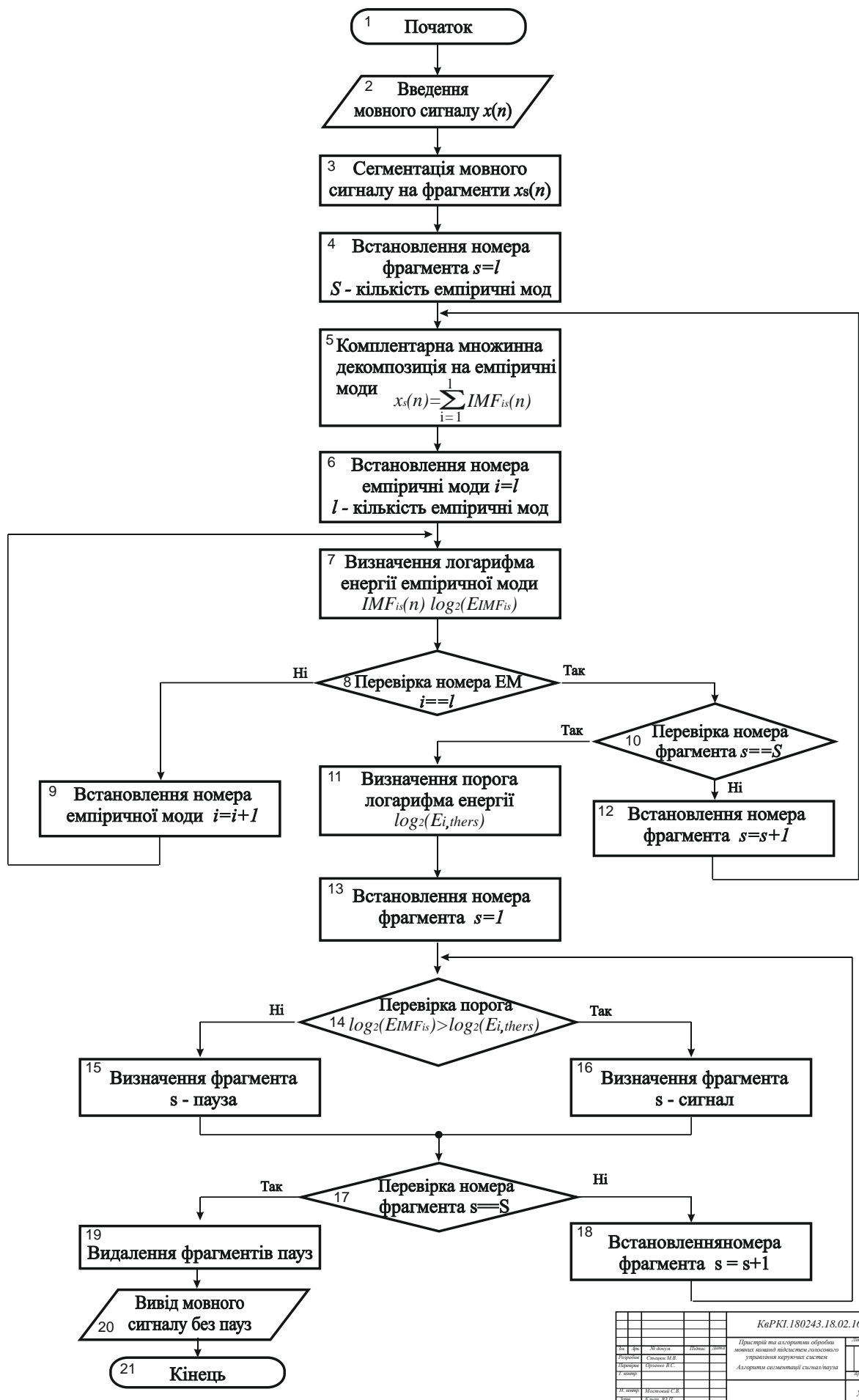
### Лістинг програми «Визначення перешкод у мовних сигналах»

```
clear;
clc;
y = wavread ('Мовний сигнал для фільтрації');
mode=eemd(y,2,50);
mode=mode';
t_inst=1 :length(y);
[A, f, tt] = hhspectrum (mode, t_inst);
freq_resol=300;
[im,tt,ff] = toimage(A,f,tt,length(tt),freq_resol);
mesh(im);
threch = 0.1;
iml=im>threch;
iml = double (iml);
mesh (iml);
t_inst=1 :length(rslt(:,1));
[A,f,tt] = hhspectrum(rslt(:,1),t_inst);
freq_resol=150;
[im,tt,ff] = toimage(A,f,tt,length(tt),freq_resol);
mesh(im);
threch=0.14;
iml=im>threch;
t_inst=1 :length(rslt(:,3));
[A,f,tt] = hhspectrum(rslt(:,3),t_inst);
freq_resol=50;
[im,tt,ff] = toimage(A,f,tt,length(tt),freq_resol);
mesh(im);
t_inst=1 :length(rslt(:,5));
[A,f,tt] = hhspectrum(rslt(:,5),t_inst);
freq_resol=25;
[im,tt,ff] = toimage(A,f,tt,length(tt),freq_resol);
mesh(im);
threch=0.4;
iml=im>threch;
t_inst=1 :length(rslt(:,7));
[A,f,tt] = hhspectrum(rslt(:,7),t_inst);
freq_resol=5;
[im,tt,ff] = toimage(A,f,tt,length(tt),freq_resol);
mesh(im);
threch = 0.8;
iml=im>threch;
```

```
t_inst=1:length(rslt(:,9));  
[A,f,tt] = hhspectrum(rslt(:,9),t_inst);  
[im,tt,ff] = toimage(A,f,tt,length(tt),freq_resol);  
mesh(im);  
t_inst=1:length(x);  
[A,f,tt] = hhspectrum(rslt(:,10),t_inst);  
[im,tt,ff] = toimage(A,f,tt,length(tt),freq_resol);  
mesh(im);  
save_result = save (C: \ Matlab \ im.m);  
clc;  
clear;
```

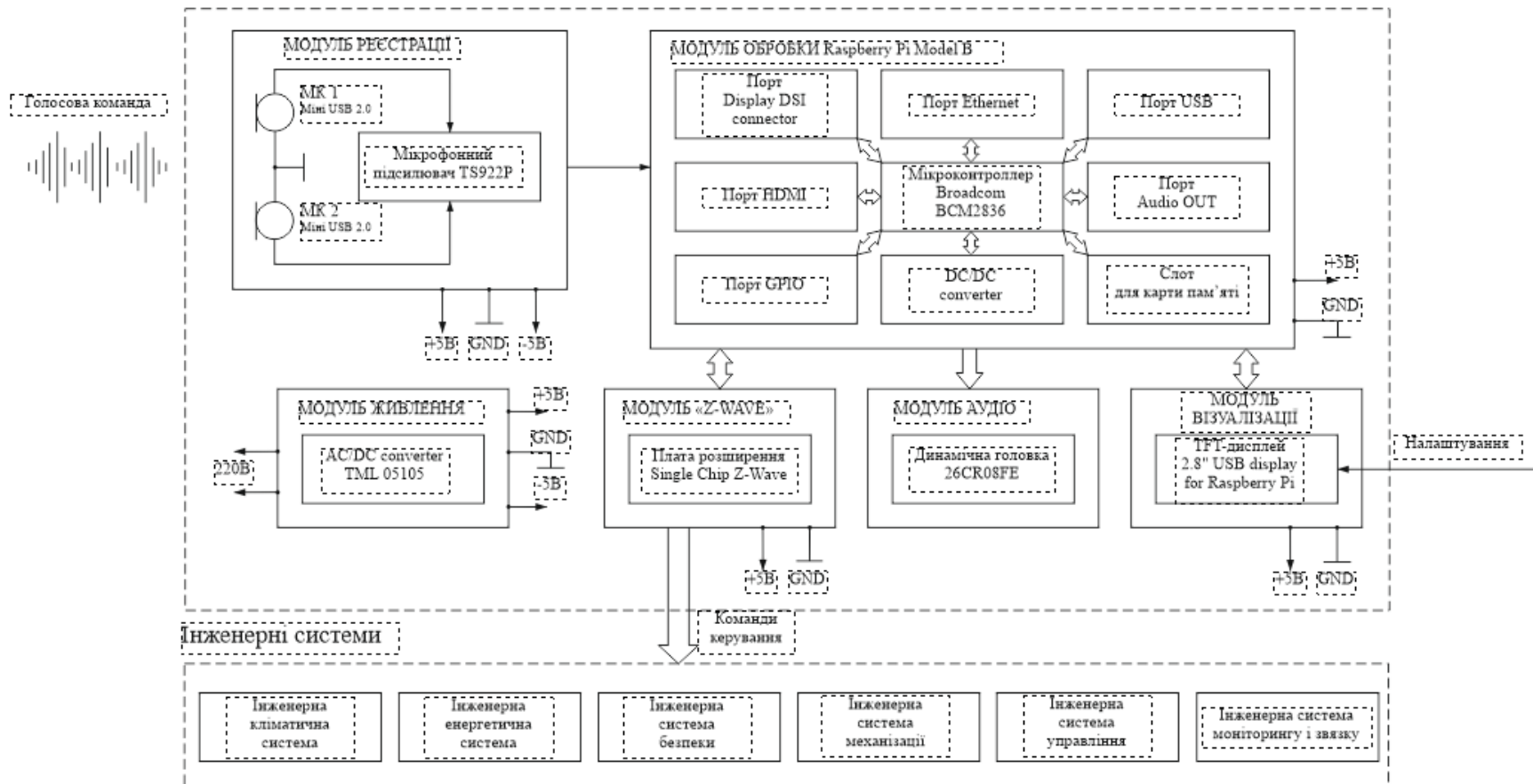




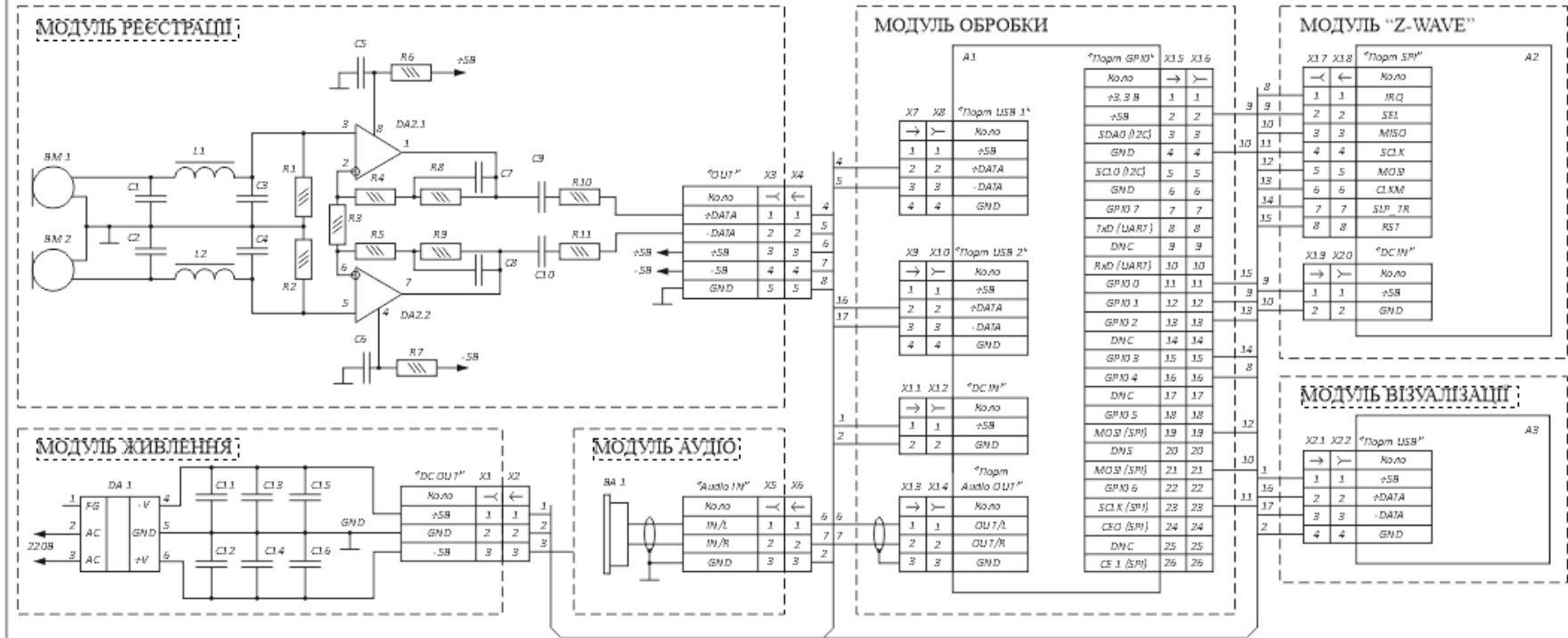


				KaPKI.180243.18.02.16 E8					
№	Фр.	Мова	Пауза	Сигнал	Пристрій та алгоритми обробки мовних сигналів відділення спеціального управління керування системою				
№	Фр.	Словесні М.Б.			Алгоритми сегментації сигнал/пауза				
№	Фр.	Фронтальні В.С.							
№	Фр.	Мовні М.Б.							
№	Фр.	Хоча В.С.							
					ХНУ -р. КІ-18-2				

Підсистема голосового управління ІВКС



КаPKI.180243.18.02.16 E2	
Проєкт на виконання роботи згідно технічної специфікації виконавця з метою реалізації програмної частини	
Сторінка інженерних функціоналів	
XIV/пр. 03-18-2	



- A1 - Мінікомп'ютер Raspberry Pi Model B
- A2 - Плата розширення Single Chip Z-Wave
- A3 - Дисплей 2.8" USB TFT display for Raspberry Pi Model B
- BA - Динамічна галівка 26CR08FE
- BM 1, BM 2 - Мікрофон Mini USB 2.0 MIC
- DA 1 - AC/DC конвертер TRACO POWER TML 05105
- DA 2 - Операційний підсилювач TS922P

## Anti-Plagiarism v-15.257

**Максимальное совпадение с одним документом 4.0%**

Словари проверки: en\_US, ru\_RU, ua\_UA. **Ошибок в документах: 6%**

ID: 104373 Название: Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем Добавлено в БД: 2022-06-02 Авторы: Стецюк Микола Валерійович Руководители: Орленко В.С. Консультанты: Опоненты:	Документ		Суммарное совпадение по Базе Данных	
	Символы	Лексемы	Символы	Лексемы
	77623	1146	3136 (4%)	47 (4%)

### Источник плагиата

ID	Описание	Наличие плагиата в документе	
		Символы	Лексемы

Ім'я користувача:  
Кафедра кібербезпеки

Дата перевірки:  
02.06.2022 11:53:48 EEST

Дата звіту:  
02.06.2022 12:08:36 EEST

ID перевірки:  
1011428845

Тип перевірки:  
Doc vs Internet + Library

ID користувача:  
100008300

Назва документа: Стецюк\_кваліфікаційна

Кількість сторінок: 69 Кількість слів: 12758 Кількість символів: 100267 Розмір файлу: 3.43 MB ID файлу: 1011309109

Виявлено модифікації тексту (можуть впливати на відсоток схожості)

**6.73%**  
**Схожість**

Найбільша схожість: 6.44% з джерелом з Бібліотеки (ID файлу: 1011309136)

2.75% Джерела з Інтернету

182

Сторінка 71

6.56% Джерела з Бібліотеки

121

Сторінка 73

**0% Цитат**

Вилучення цитат вимкнене

Вилучення списку бібліографічних посилань вимкнене

**0%**  
**Вилучень**

Немає вилучених джерел

**Модифікації**

Виявлено модифікації тексту. Детальна інформація доступна в онлайн-звіті.

Замінені символи

8

Підозріле форматування

13  
сторінок

**РЕЦЕНЗІЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ**  
освітнього ступеня «бакалавр»

Студент Стецюк Микола Валерійович

Тема Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем

Спеціальність: 123 - Комп'ютерна інженерія

**Обсяг кваліфікаційної роботи освітньо-кваліфікаційного рівня «бакалавр»:**

кількість листів креслень 5; кількість сторінок записки 64

1. Короткий зміст КР та прийнятих рішень В рамках кваліфікаційної роботи запропоновано підхід до обробки мовних сигналів підсистеми голосового управління керуючих систем, заснований на алгоритмах і пристроях обробки та оптимізовану послідовність дій: фільтрацію нестационарних фонових шумів, сегментацію сигнал/пауза, вимірювання чистот основного тону. Розроблено структурну та електричну схеми з'єднання підсистеми голосового управління керуючої системи домашньої автоматизації за бездротовим протоколом передачі даних Z-Wave.

2. Висновок про відповідність КР кваліфікаційному завданню Кваліфікаційна робота у повній мірі відповідає поставленому завданню як в теоретичній, так і в практичній частині представленої роботи

3. Характеристика виконання кожного розділу проекту, ступінь використання останніх досягнень науки і техніки і передових методів роботи: У першому, теоретичному, розділі кваліфікаційної роботи якісно та в повній мірі розглянуті методи вирішення поставленої задачі, проведено аналіз існуючих систем голосового управління та застосовуваних підходів обробки мовних команд. Досліджено особливості обробки мовних команд у підсистемах голосового управління керуючих систем, проведено аналіз методів обробки мовних команд у підсистемі голосового управління. Запропоновано підхід до обробки мовних сигналів підсистеми голосового управління керуючих систем, заснований на алгоритмах і пристроях обробки та оптимізовану послідовність дій: фільтрацію нестационарних фонових шумів, сегментацію сигнал/пауза, вимірювання чистот основного тону. В загальному усі розділи відповідають завданню та містять сучасні методи вирішення поставлених завдань.

4. Позитивні сторони проекту Кваліфікаційна робота відповідає сучасним вимогам до проектування інформаційних систем та містить ряд інноваційних рішень, зокрема, з обробки мовних сигналів підсистеми голосового управління керуючих систем, оптимізовано послідовність дій: фільтрацію нестационарних фонових шумів, сегментацію сигнал/пауза, вимірювання чистот основного тону.

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

5. Негативні сторони проекту В рамках кваліфікаційної роботи не наведені дані щодо конкретних затрат на проведення сегментації, фільтрації, розпізнавання мовних сигналів та виділення інформативних параметрів мовного сигналу системи голосового управління. Відсутні рекомендації щодо області використання запропонованих алгоритмів та підходів наведених в кваліфікаційній роботі.

6. Оцінка графічного оформлення та пояснювальної записки проекту Графічне оформлення виконане відповідно до теми кваліфікаційної роботи. На першому листку креслення наведено технологічний процес обробки мовних команд. В наступних листах креслення розглянуті питання - алгоритм фільтрації нестационарних фонових шумів, алгоритм сегментації сигнал/пауза. Схема електрична функціональна та схема електрична принципова. В загальному графічне оформлення виконане на достатньому рівні. Пояснювальна записка відповідає стандартам до її оформлення

7. Відгук про проект в цілому В загальному кваліфікаційна робота заслуговує позитивної оцінки. Весь матеріал кваліфікаційної роботи структурований, чіткий та послідовний. Усі розділи кваліфікаційної роботи послідовні та логічні, що дозволяє чітко розуміти матеріал викладений в рамках даної кваліфікаційної роботи. Графічний матеріал дозволяє наочно побачити доцільність та ефективність рішень, які були прийняті за основу при розробці пристрою та алгоритмів обробки мовних команд підсистем голосового управління керуючих систем

8. Інші зауваження

9. Оцінка дипломного проекту Розглянувши позитивні та негативні сторони представленого дипломного проекту, можна зробити висновок, що він заслуговує оцінку «добре».

РЕЦЕНЗЕНТ (прізвище, ім'я, по батькові, посада, місце роботи)

Гурман Іван Васильович, к.т.н., доцент  
кафедра інженерії програмного забезпечення, Хмельницький національний університет

« \_\_\_\_\_ » \_\_\_\_\_ 2022 р.

 (підпис)

Завідувачу кафедри кібербезпеки  
к.т.н., доц. Кльоцу Ю.П.

Стецюка Миколи Валерійовича  
ПІБ здобувача вищої освіти

студента ФІТ, 4 курсу, групи КІ-18-2

### ЗАЯВА

З правилами чинного Положення «Про дотримання академічної доброчесності в Хмельницькому національному університеті» від 26.09.2020 (зі змінами від 26.11.2020), згідно з яким виявлення плагіату є підставою для відмови в допуску кваліфікаційної роботи до захисту та застосування заходів дисциплінарної та академічної відповідальності, ознайомлений (а). Про використання програмно-технічних засобів для перевірки кваліфікаційних робіт здобувачів вищої освіти на плагіат оповіщений (а) та надаю свою згоду на обробку та збереження університетом моєї роботи в інституційному репозитарії університету.

Також надаю університету право на передачу моєї роботи для обробки та збереження в базах даних програмно-технічних засобів (Unicheck та Anti-Plagiarism) та використання роботи для виявлення плагіату в інших роботах, які перевіряються програмно-технічними засобами та користувачами, що мають доступ до цих програмно-технічних засобів, виключно в обмежених цілях для виявлення плагіату в текстах робіт.

Робота для перевірки університетом надається в друкованому та електронному варіанті. Електронна версія моєї роботи збігається (ідентична) з друкованою.

6.06.2022

дата



підпис

РІШЕННЯ ЕКСПЕРНОЇ КОМІСІЇ

КАФЕДРИ КІБЕРБЕЗПЕКИ

ПРО ДОПУСК КВАЛІФІКАЦІЙНОЇ РОБОТИ ДО ЗАХИСТУ

Підтверджуємо ознайомлення з результатом звіту подібності щодо роботи, генерованого системою виявлення текстових збігів/ідентичності/схожості:

Назва: Пристрій та алгоритми обробки мовних команд підсистем голосового управління керуючих систем

Автор: Стецюк Микола Валерійович

Спеціальність: 123 – Комп'ютерна інженерія

Освітня програма: Комп'ютерна інженерія

Науковий керівник: Орленко В.С.

Після аналізу звіту подібності зроблено такий висновок:

№	Висновок	Позначка про відповідність
1	Запозичення, виявлені в роботі, є законними і не є плагіатом (далі – зазначаються підстави віднесення запозичень до правомірних). Робота приймається до захисту.	відповідає
2	Виявлені запозичення не є плагіатом, розміщені в розділах, які не описують безпосередньо авторське дослідження, але кількість цитат перевищує обсяг, виправданий поставленою метою роботи (далі – зазначаються детальні та аргументовані підстави віднесення запозичень до правомірних). Робота приймається до захисту, але має бути відкоригована. Відкоригований варіант має бути поданий на кафедру за 2 дні до захисту, разом із заявою щодо самостійності виконання письмової роботи та ідентичності друкованої та електронної версії роботи.	
3	Виявлені запозичення не є плагіатом, але частково розміщені в розділах, які описують безпосередньо авторське дослідження, а кількість цитат перевищує обсяг, виправданий поставленою метою роботи. В зв'язку з цим мета роботи та поставлені завдання не були досягнені. Робота може бути допущена до захисту (наступного року) після того як буде відкоригована та допрацьована і успішно пройде повторну перевірку на академічний плагіат.	
4	Робота містить навмисні текстові спотворення, передбачувані спроби укриття запозичень або інші прояви академічного плагіату. Робота містить фабрикацію або фальсифікацію даних. Робота не допускається до захисту.	
5	Інше:	

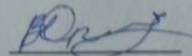
Підтвердження:

Запозичення, виявлені в роботі, є законними і не є плагіатом, оскільки:

- 1) запозичення розміщені в розділах аналізу існуючих методів та технологій, які не описують безпосередньо авторське дослідження і не стосуються результатів роботи;
- 2) усі запозичення фрагментарні, або мають належним чином оформленні посилання;
- 3) окремі виявлені збіги є загальноживаними фразами або виразами, про що свідчить посилання системи на збіг з 10-40 джерелами на один фрагмент речення;
- 4) всі зафіксовані системою ознаки модифікації тексту відносяться до комбінування латинських символів зі україномовними скороченнями індексів в формулах, що не є модифікацією тексту.

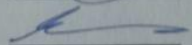
Сумарний обсяг всіх запозичень, визначений системою виявлення збігів/ідентичності/схожості, складає 11.0% і адресується до 136 першоджерел, що, з урахуванням наведених обґрунтувань, відповідає характеру наукового дослідження і свідчить на користь кваліфікаційної роботи.

Керівник роботи



В.С. Орленко

Завідувач кафедри КБ



Ю.П. Кльоц

Дата: 16.06.2022