

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

факультет прикладних інформаційних технологій та електроінженерії

(повна назва факультету)

кафедра радіотехнічних систем

(повна назва кафедри)

КВАЛІФІКАЦІЙНА РОБОТА

на здобуття освітнього ступеня

магістр

(назва освітнього ступеня)

на тему: Спосіб розпізнавання голосових команд керування промисловими
електроеханічними пристроями

Виконав(ла): студент(ка) 6 курсу, групи РАМ-61
спеціальності

172 Електронні комунікації та радіотехніка

(шифр і назва спеціальності)

Федунь А.М.
(підпис) (прізвище та ініціали)

Керівник
(підпис) Яворський Б.І.
(прізвище та ініціали)

Нормоконтроль
(підпис) Хвостівська Л.В.
(прізвище та ініціали)

Завідувач кафедри
(підпис) Дунець В.Л.
(прізвище та ініціали)

Рецензент
(підпис) Дедів Л.Є.
(прізвище та ініціали)

Тернопіль
2025

Міністерство освіти і науки України
Тернопільський національний технічний університет імені Івана Пулюя

Факультет прикладних інформаційних технологій та електроінженерії
(повна назва факультету)

Кафедра радіотехнічних систем
(повна назва кафедри)

ЗАТВЕРДЖУЮ

Завідувач кафедри

Дунець В.Л.
(підпис) (прізвище та ініціали)

«26» листопада 2025 р.

**З А В Д А Н Н Я
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

на здобуття освітнього ступеня магістр

(назва освітнього ступеня)

за спеціальністю 172 Електронні комунікації та радіотехніка

(шифр і назва спеціальності)

студенту Федуню Андрію Михайловичу

(прізвище, ім'я, по батькові)

1. Тема роботи Спосіб розпізнавання голосових команд керування промисловими електроеханічними пристроями

Керівник роботи Яворський Богдан Іванович, д.т.н., професор

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

Затверджені наказом ректора від «23» листопада 2025 року № 4/7-1038

2. Термін подання студентом завершеної роботи 11 грудня 2025 року

3. Вихідні дані до роботи Технічне завдання, розпізнавання голосових команд керування

4. Зміст роботи (перелік питань, які потрібно розробити)

1. Аналітична частина

2. Основна частина

3. Науково-дослідна частина

4. Охорона праці та безпека в надзвичайних ситуаціях

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень, слайдів)

6. Консультанти розділів роботи

[illegible]

7. Дата видачі завдання

КАЛЕНДАРНИЙ ПЛАН

[illegible]

Студент _____
(підпис)

Федунь Андрій Михайлович
(прізвище та ініціали)

Керівник роботи _____
(підпис)

Яворський Богдан Іванович

(прізвище та ініціали)

АНОТАЦІЯ

Спосіб розпізнавання голосових команд керування промисловими електромеханічними пристроями // Кваліфікаційна робота магістра // Федунь Андрій Михайлович // ТНТУ ім. І.Пулюя, ФПТ // Тернопіль, 2025 // с. - 83, рис. - 13, дод. - 2, бібл. - 47.

Ключові слова: КОМАНДА, КЕРУВАННЯ, ІНТЕРФЕЙС, ГОЛОС.

В роботі розроблено спосіб ідентифікації голосових команд керування промисловими електромеханічними пристроями. В способі розпізнавання використано метод опрацювання голосових сигналів, що являють собою записані з допомогою мікрофона команди, після чого виконується обробка сигналу та виділення інформативних ознак, за якими здійснюється розпізнавання. Основними етапами методу є: запис мовного сигналу та підготовка даних для аналізу; адаптивна фільтрація з метою зменшення впливу шумів; визначення часових меж (тривалості) команд методом ковзного вікна; фонемна сегментація команд на основі аналізу частоти основного тону. Частота основного тону оцінювалась методом формантного аналізу, а її наявність у кожному ковзному вікні фіксувалась через порогову функцію. Послідовність інтервалів із наявною/відсутю частоти основного тону, а також загальна тривалість команди запропоновано використати як критерії для ідентифікації команди та подальшого формування відповідного сигналу керування.

ANNOTATION

Method for voice commands recognizing for controlling industrial electromechanical devices // Fedun Andriy Mykhailovych // TNTU, FPT // Ternopil, 2025 // p. - 83, fig. - 13, appl. - 2, bibl. - 47.

Key words: COMMAND, CONTROL, INTERFACE, VOICE.

The paper developed a method for voice commands recognizing for controlling industrial electromechanical devices. The recognition method uses a method for processing voice signals, which are commands recorded using a microphone, after which the signal is processed and informative features are extracted, by which recognition is carried out. The main stages of the method are: recording a speech signal and preparing data for analysis; adaptive filtering to reduce the impact of noise; determining the time limits (duration) of commands using the sliding window method; phonemic segmentation of commands based on the analysis of the fundamental tone frequency. The fundamental tone frequency was estimated using the formant analysis method, and its presence in each sliding window was recorded using a threshold function. The sequence of intervals with the presence/absence of the fundamental tone frequency, as well as the total duration of the command are proposed to be used as criteria for identifying the command and further forming the corresponding control signal.

ЗМІСТ

ВСТУП.....	8
РОЗДІЛ 1. АНАЛІТИЧНА ЧАСТИНА.....	10
1.1 Аналіз проблеми голосового керування промисловим обладнанням.....	10
1.2 Задача керування промисловими верстатами за допомогою голосового інтерфейсу.....	13
1.3 Обробка природної мови (NLP).....	14
1.4 Інтерфейси людина–машина та голосовий інтерфейс.....	15
1.5 Голосові інтерфейси у виробництві.....	18
1.6 Архітектура голосового інтерфейсу.....	19
1.7 Концептуальна структура.....	20
1.8 Приклад реалізації голосового інтерфейсу.....	23
1.9 Переваги та недоліки голосових інтерфейсів для керування промисловим обладнанням.....	25
1.10 Висновки до розділу 1.....	28
РОЗДІЛ 2. ОСНОВНА ЧАСТИНА.....	29
2.1 Оцінка ефективності голосового керування промисловим обладнанням.....	29
2.2 Поняття голосової команди та задача розпізнавання.....	36
2.3 Загальна структура системи розпізнавання команд.....	36
2.4 Методи та ознаки для розпізнавання голосових команд у системах керування.....	37
2.5 Підходи до розпізнавання команд при голосовому керуванні.....	38
2.6 Висновки до розділу2.....	41
РОЗДІЛ 3. НАУКОВО-ДОСЛІДНА ЧАСТИНА.....	43
3.1 Принцип роботи пропонованого способу.....	43
3.2 Експериментальні результати.....	47

	7
3.3 Висновки до розділу 3.....	56
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ.....	57
4.1 Охорона праці.....	57
4.2 Безпека в надзвичайних ситуаціях.....	58
4.3 Висновки до розділу.....	65
ВИСНОВКИ.....	66
ПЕРЕЛІК ПОСИЛАНЬ.....	68
ДОДАТКИ	

ВСТУП

Актуальність. Актуальність розроблення методу ідентифікації голосових команд для керування промисловими електромеханічними пристроями зумовлена сукупністю технологічних, виробничих і ергономічних чинників, що посилюються в умовах цифровізації промисловості та переходу до концепцій Індустрії 4.0. Сучасні виробничі системи характеризуються зростанням складності технологічних процесів, підвищенням вимог до гнучкості переналагодження, скороченням виробничих циклів та необхідністю швидкого реагування на зміни умов експлуатації. У таких умовах ефективність функціонування приводів, двигунів, конвеєрів, маніпуляторів, верстатів/ЧПК та мехатронних вузлів суттєво залежить не лише від їх конструктивних і регуляторних характеристик, а й від якості людино-машинної взаємодії на рівні керування.

Традиційні засоби керування промисловим обладнанням (кнопкові пульти, сенсорні панелі, стаціонарні операторські термінали, клавіатурно-мишкові інтерфейси) залишаються домінантними, проте мають низку обмежень. Зокрема, вони часто потребують фізичного контакту та візуального контролю, що ускладнює роботу оператора в умовах багатозадачності, обмеженого простору, необхідності використання рукавичок чи масок або підвищених вимог до чистоти робочого процесу.

Голосове керування, як форма природного людино-машинного інтерфейсу, здатне частково усунути ці обмеження завдяки безконтактній взаємодії та можливості “hands-free” керування. Для промислових електромеханічних систем це потенційно означає скорочення часу введення команд, підвищення ергономіки, зниження навантаження на оператора та розширення доступності керування для працівників із обмеженими можливостями. Особливе значення

голосові команди можуть мати у сценаріях, де оператору необхідно одночасно виконувати механічні дії (налаштування вузлів, позиціонування заготовок, контроль інструмента), контролювати стан технологічного процесу та оперативно задавати керувальні впливи на приводи й виконавчі механізми.

Таким чином, розроблення методу ідентифікації команд за голосом для керування промисловими електромеханічними пристроями є актуальним науково-прикладним завданням, що спрямоване на підвищення ефективності й безпеки людино-машинної взаємодії в сучасних виробничих системах.

Мета. Розробка способу та методу ідентифікації команд за голосом для керування промисловими електромеханічними пристроями. **Задачі:**

- аналіз способів керування промисловим одладнанням;
- аналіз принципу роботи систем голосового керування;
- обґрунтування способу та методу ідентифікації голосових команд;
- експериментальне дослідження розробленого методу ідентифікації

голосових команд.

Об'єкт дослідження: процес ідентифікації голосових команд керування промисловими електромеханічними пристроями.

Предмет дослідження: метод ідентифікації голосових команд керування промисловими електромеханічними пристроями.

Наукова новизна. Запропоновано для ідентифікації голосових команд проводити опрацювання голосових сигналів в межах ковзного вікна із наступним виділенням ділянок з присутністю частоти основного тону. Це дасть можливість наступної сегментації команди на фонемні ділянки та наступної ідентифікації команд.

Практичне значення. Розроблений спосіб та метод можуть бути використані як окремий модуль виділення ознак команд голосового керування.

РОЗДІЛ 1

АНАЛІТИЧНА ЧАСТИНА

1.1 Аналіз проблеми голосового керування промисловим обладнанням

Традиційні засоби керування промисловим обладнанням (кнопкові пульти, сенсорні панелі, стаціонарні операторські термінали, клавіатурно-мишкові інтерфейси) залишаються домінантними, проте мають обмеження. Так, вони часто потребують фізичного контакту та візуального контролю, що ускладнює роботу оператора в умовах багатовимірності, обмеженого простору, необхідності використання рукавичок чи масок або підвищених вимог до чистоти робочого процесу. Також, стаціонарність терміналів зменшує мобільність оператора, що особливо критично під час обслуговування великогабаритного обладнання, налаштування технологічних режимів або керування розподіленими системами (лінії транспортування, багатопозиційні комплекси, роботизовані модулі). По-третє, навігація багаторівневими меню HMI/SCADA або CNC-інтерфейсів потребує часу, збільшує когнітивне навантаження та підвищує ризик помилок, особливо під час виконання складних або рідковживаних операцій.

Голосове керування, як форма природного людино-машинного інтерфейсу, здатне частково усунути ці обмеження завдяки безконтактній взаємодії та можливості “hands-free” керування. Для промислових електромеханічних систем це потенційно означає скорочення часу введення команд, підвищення ергономіки, зниження навантаження на оператора та розширення доступності керування для працівників із обмеженими можливостями. Особливе значення голосові команди можуть мати у сценаріях, де оператору необхідно одночасно виконувати механічні дії (налаштування вузлів, позиціонування заготовок,

контроль інструмента), контролювати стан технологічного процесу та оперативно задавати керувальні впливи на приводи й виконавчі механізми.



а)



б)



в)

Рис. 1.1. Приклади голосового керування: а – оператор станка з ЧПУ;
б – оператор складу; в – голосове керування автомобілем

Водночас промислове середовище створює специфічні виклики для розпізнавання мовлення. Основними з них є високий рівень акустичних перешкод (шум двигунів, редукторів, компресорів, потоків повітря), реверберації приміщень, наявність імпульсних шумів, а також різноманітність голосів операторів (темп, акцент, інтонаційні особливості). Додатковим фактором є вимога до мінімальної затримки та високої надійності: у виробничих процесах час реакції на команду, особливо в аварійних режимах (“стоп”, “пауза”, “відвід інструмента”), може бути критичним для безпеки персоналу та збереження обладнання. Тому метод розпізнавання голосових команд для промислових електромеханічних пристроїв має забезпечувати не лише високу точність, а й стабільність роботи за змінних акустичних умов, можливість виявлення невизначених або некоректних команд, а також механізми підтвердження та контролю безпечного виконання дій.

Окрему актуальність становить потреба адаптації методів розпізнавання до предметної області. Команди керування приводами, маніпуляторами чи ЧПК зазвичай мають формалізовану структуру та містять параметри (напрямок, координати, швидкість, подача, прискорення, номер програми, режим роботи), що потребує надійного вилучення числових значень і контекстних ознак. На практиці це означає необхідність розроблення методів, які поєднують стійку акустичну обробку сигналів (подавлення шумів, нормалізація, виділення інформативних ознак) із ефективними підходами до розпізнавання команд обмеженого словника та семантичного зіставлення команд із діями керування. Важливо також враховувати мовні особливості (зокрема українську мову), оскільки промислові рішення часто орієнтовані на найбільш поширені мовні моделі, що знижує якість розпізнавання для локальних виробництв і потребує спеціалізованих корпусів команд, словників та процедур адаптації.

Очікуваний практичний ефект полягає у скороченні часу керування, підвищенні зручності роботи оператора, зниженні ймовірності помилок під час виконання команд та створенні передумов для впровадження інтелектуальних, інклюзивних і мобільних НМІ у промисловості.

1.2 Задача керування промисловими верстатами за допомогою голосового інтерфейсу

Верстати є базовим елементом виробництва, адже їхня продуктивність безпосередньо впливає на ефективність і якість технологічних процесів. Концепція Індустрії 4.0 передбачає глибоку цифровізацію, зростання вимог клієнтів і підвищення рівня індивідуалізації продукції. У відповідь на це формується нове покоління обладнання – кіберфізичні верстати (CPMT), які можуть працювати більш автономно та забезпечувати вищу операційну ефективність, ніж традиційні ЧПК-системи [1].

Однією з ключових складових CPMT є розвиток інтелектуальних інтерфейсів людина–машина (НМІ). Сучасні ЧПК-верстати дають змогу виконувати точну й високорівневу автоматизацію обробки, однак їхні НМІ часто залишаються “слабким місцем”: вони вимагають постійного ручного керування та уваги добре підготовленого оператора [2]. Крім того, панель керування зазвичай фізично закріплена на верстаті, що обмежує мобільність оператора та може уповільнювати виконання типових дій.

Швидкий розвиток машинного навчання, зокрема технологій обробки природної мови (NLP), розпізнавання мовлення (STT) і синтезу мовлення (TTS), зробив голосові інтерфейси поширеним явищем (приклади — Amazon Alexa, Apple Siri). У промисловому контексті голосовий інтерфейс користувача (VUI) може бути перспективним доповненням до НМІ, оскільки здатен зменшити

типові обмеження традиційних інтерфейсів: недостатню інклюзивність, невисоку інтуїтивність та залежність оператора від стаціонарної панелі керування.



Рис. 1.2. Організація керування розумним будинком на основі сервісу Amazon Alexa

1.3 Обробка природної мови (NLP)

Обробка природної мови (NLP) — це напрям, що займається аналізом та інтерпретацією мовлення або тексту, тобто природної мови, якою люди спілкуються між собою [3]. Завдяки сучасним методам машинного навчання NLP-системи можуть визначати зміст висловлювання, намір користувача та, за потреби, виділяти параметри команди [4].

Важливим кроком у розвитку NLP стала поява архітектури Transformer [5], яка суттєво підвищила якість роботи з мовними даними. Саме на цій архітектурі базуються відомі моделі, зокрема BERT [6] і GPT [7]. Transformer використовує механізм уваги (attention), що дозволяє ефективно враховувати контекст і залежності між словами в реченні. На практиці трансформери застосовують не лише в NLP, а й у системах STT та TTS.

Для промислових застосувань важливо, що попередньо навчені мовні моделі можна донавчити на технічній лексиці та командному словнику конкретного верстата. Це дозволяє побудувати модуль NLU (розуміння мовлення), який буде коректно інтерпретувати команди в реальних виробничих умовах. У результаті VUI може забезпечити більш гнучке керування, скоротити час взаємодії та підвищити зручність роботи оператора — за умови правильного проєктування безпеки та контролю помилок.

1.4 Інтерфейси людина–машина та голосовий інтерфейс

У виробничих системах НМІ забезпечує двосторонню взаємодію: оператор отримує інформацію від машини та передає їй керувальні дії [8]. Інтерфейси НМІ включають технології введення та виведення.

Засоби введення можна умовно поділити на:

- матеріальні (контактні) — кнопки, сенсорні панелі, джойстики, клавіатури;
- нематеріальні (безконтактні) — розпізнавання жестів, відстеження руху, голосове керування тощо.

Безконтактні методи часто сприймаються як більш інтуїтивні і можуть працювати без “прив’язки” до конкретного місця у виробничому просторі. Водночас існує компроміс: такі інтерфейси інколи складніше зробити повністю

функціональними та надійними в усіх робочих сценаріях (особливо через шум, освітлення, відстань, безпеку команд).

Засоби виведення забезпечують зворотний зв'язок у візуальній або невізуальній формі: стаціонарні дисплеї, звукові повідомлення, а також носимі пристрої (наприклад, дисплеї на голові для доповненої реальності) [9].

Голосовий інтерфейс користувача (VUI) є частиною НМІ, де введення здійснюється голосом, а система відповідає мовленням, текстом або іншими каналами. На відміну від традиційних рішень, що здебільшого базуються на ручному введенні та візуальному відображенні, VUI може забезпечити швидшу взаємодію в ситуаціях, коли оператору потрібно діяти без використання рук або коли доступ до панелі керування обмежений.

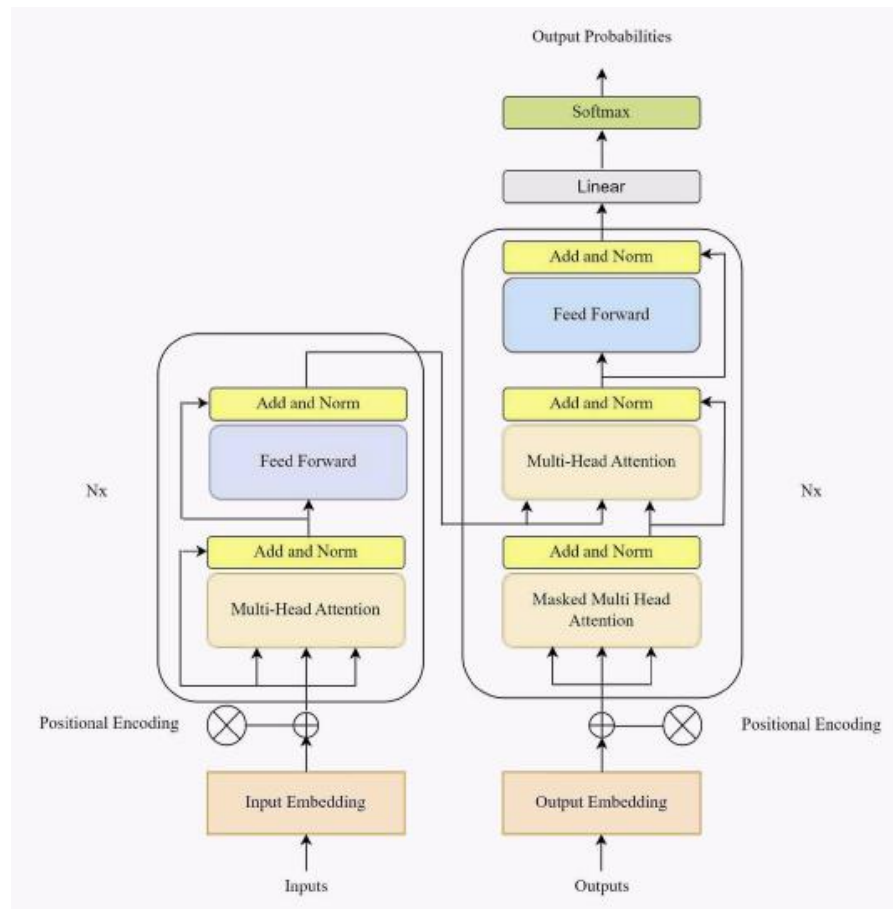


Рис. 1.3. Архітектура трансформера, запропонована в [5].

Голосові інтерфейси користувача (VUI) — це поєднання апаратних і програмних компонентів, які забезпечують взаємодію людини з машиною за допомогою мовлення.

З апаратного боку потрібні:

- засоби запису голосу (мікрофон або масив мікрофонів);
- засоби подання відповіді користувачу (динамік, дисплей або комбінований канал);
- обчислювальні ресурси (локальний контролер, edge-сервер або хмарні потужності), які виконують обробку аудіо та прийняття рішень.

З програмного боку базові компоненти VUI — це:

- ASR / STT (speech-to-text) — перетворює мовлення на текст для подальшого аналізу [10];
- NLP / NLU — інтерпретує зміст висловлювання, визначає намір (intent) і параметри команди;
- TTS (text-to-speech) — генерує голосову відповідь або підтвердження виконання команди.

Розпізнавання мовлення перетворює сказані слова в текст, після чого система може вилучати потрібну інформацію та формувати керувальну дію. Завдяки сучасним методам NLP тривають дослідження щодо підвищення точності ASR, особливо в складних умовах. Для формування аудіозворотного зв'язку використовується TTS. Хоча синтез мовлення відомий давно, саме нові підходи (зокрема на базі глибокого навчання та самонавчання) дозволили наблизити якість синтезу до природного звучання [11].

Для “розумної” взаємодії в промисловості, крім STT і TTS, потрібні ще керування діалогом та логіка вибору дій, щоб система не просто “чула”, а коректно розуміла контекст і виконувала допустимі операції. Загалом VUI роблять взаємодію інтуїтивнішою, бо мова є природним каналом введення.

Водночас наразі такі інтерфейси найчастіше застосовують у вигляді персональних помічників у побутовій електроніці та сфері охорони здоров'я.

1.5 Голосові інтерфейси у виробництві

Останніми роками з'явилися роботи, присвячені використанню VUI для керування промисловими ділянками, роботами та верстатами. Наприклад, у дослідженнях [12] описано проєктування ЧПК-верстатів із голосовим керуванням, де VUI реалізують через зовнішній модуль розпізнавання мовлення. У такому підході часто застосовують обмежений набір команд (наприклад, близько 80). Це підвищує безпеку, тому що система приймає лише наперед визначені фрази, але одночасно звужує функціональні можливості VUI.

Інші роботи (наприклад, [2]) акцентують увагу на проблемах промислового впровадження VUI. Ключовим фактором тут є точність і надійність: помилкова інтерпретація команди може призвести до некоректної роботи обладнання й створити ризик для оператора та машини. Щоб зменшити ці ризики, пропонують використовувати спеціалізовану “підмову” команд (умовно — опис мови голосових команд, VCD) — тобто стандартизовані формулювання, які легше однозначно розпізнати та перевірити.

Сучасні моделі глибокого навчання (зокрема трансформери) забезпечують високу точність у задачах розуміння мовлення (NLU), близьку до рівня людини. У роботі [13] запропоновано концепцію керування ЧПК-верстатом через VUI, але також підкреслено, що повна заміна традиційного НМІ голосом має обмеження. Важливим є етап навчання мовного модуля: його потрібно адаптувати під необхідну область застосування і реальні сценарії роботи. Для цього використовують підхід “Software-in-the-Loop” (SIL) і цифровий двійник верстата, який відтворює поведінку обладнання. Це дозволяє інтенсивно тренувати й

перевіряти мовні моделі (у т.ч. великі мовні моделі, LLM) без ризику для реального обладнання та без зупинки виробництва.

1.6 Архітектура голосового інтерфейсу

Головна мета впровадження VUI для керування верстатами — підвищити гнучкість взаємодії, зробивши керування менш залежним від фізичних елементів (кнопок, панелей, місця розташування оператора). Щоб це було можливим, потрібні відповідні апаратні й програмні умови.

Апаратні вимоги.

Для навчання та роботи модуля NLU потрібні ресурси для високопродуктивних обчислень.

Для розміщення компонентів, які повинні працювати з низькою затримкою та високою пропускнуою здатністю, може знадобитися продуктивний сервер або edge-обчислення.

Для взаємодії з цифровим двійником і VUI у деяких сценаріях можуть використовуватись якісні гарнітури та засоби віртуальної реальності (VR).

Іноді як інфраструктурну основу розглядають бездротові мережі (наприклад, 5G), щоб забезпечити високошвидкісний зв'язок між VUI, цифровим двійником і фізичною машиною. Однак у реальних умовах конкретні показники пропускнуої здатності та затримки залежать від впровадження мережі, конфігурації та завантаження.

Програмні вимоги.

Модулі STT/TTS можна використовувати як готові комерційні рішення або як власні компоненти.

Модель NLU зазвичай беруть попередньо навчену та донавчають під конкретні командні формулювання й термінологію.

Для промислової інтеграції потрібні протоколи обміну даними (наприклад, OPC UA, MQTT) та належна серверна інфраструктура.

Безпека та надійність.

Найважливіша умова для промислового застосування VUI — безпека.

Перед впровадженням потрібно:

- проаналізувати ризики для оператора та обладнання;
- оцінити роботу в аварійних сценаріях (наприклад, швидкість реакції на команду аварійної зупинки);
- спроектувати зрозумілий зворотний зв'язок (щоб оператор однозначно розумів, що саме система почула і що виконає);
- реалізувати логіку запобігання небезпечним послідовностям команд (наприклад, блокування дій, що можуть пошкодити інструмент або створити небезпечну ситуацію).

Захист даних і контроль доступу.

Ще одна вимога — дотримання принципів конфіденційності:

- система не повинна озвучувати або відображати персональні дані оператора без потреби;
- команди з підвищеним рівнем доступу (наприклад, сервісна діагностика) мають бути доступні лише обслуговуючому персоналу;
- VUI має підтримувати рольову модель доступу та автентифікацію користувачів.

1.7 Концептуальна структура

Узагальнено VUI можна представити як послідовність модулів: STT → NLU → логіка/діалог → TTS, які разом забезпечують прийом голосової команди, її розуміння, вибір дії та формування відповіді.

Для промислових сценаріїв корисно додати контур SIL (Software-in-the-Loop), який дозволяє навчати та тестувати VUI через цифровий двійник обладнання. Віртуальне введення в експлуатацію зменшує залежність від реальної машини та дозволяє довше й безпечніше проводити навчання, перевірку й оптимізацію NLU.

Побудову VUI доцільно починати з формування вимог і сценаріїв використання разом із кінцевими користувачами. На цій основі створюють специфікації, приклади команд природною мовою та правила їхнього зіставлення з діями. Дані для навчання бажано збирати під наглядом експертів і з урахуванням реальних виробничих шумів, щоб моделі працювали стабільно в типовому акустичному середовищі. Якщо використовуються готові модулі STT/TTS, обсяг необхідних даних для навчання системи під конкретний верстат зазвичай можна зменшити — тоді основна увага переходить на адаптацію NLU та логіки безпечного виконання команд.

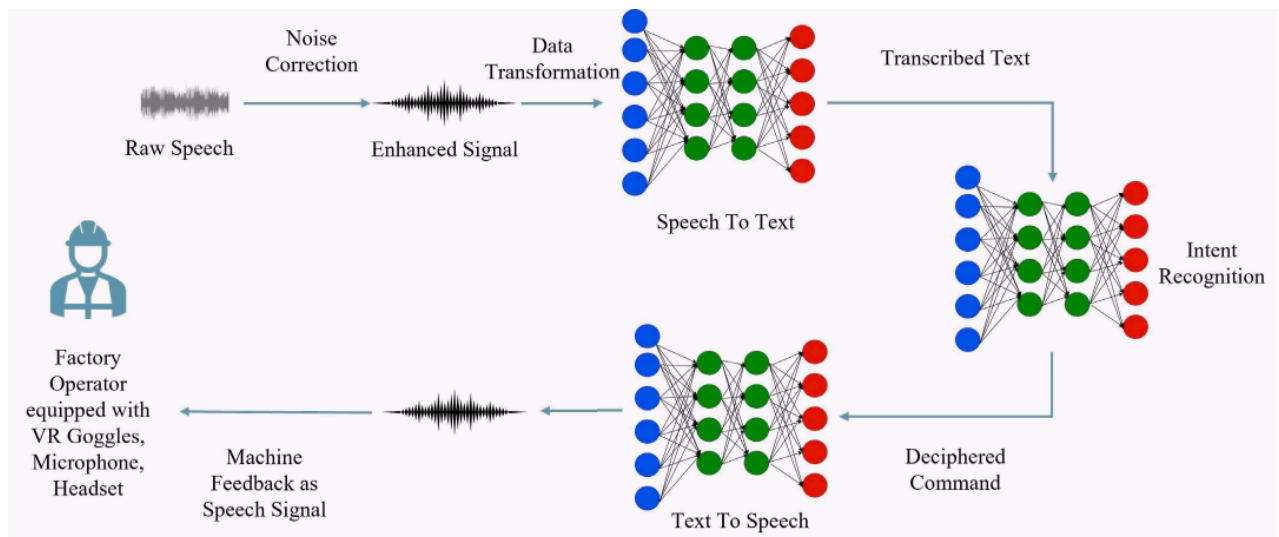


Рис. 1.4. Рівні реалізації голосового інтерфейсу користувача (VUI).

Модуль ASR (Automatic Speech Recognition) перетворює мовлення на текст (транскрипцію). На цьому етапі система працює лише з послідовністю символів

і слів, характерних для конкретної мови або діалекту, без розуміння змісту сказаного.

Щоб перейти від “тексту” до “сенсу”, застосовується модуль NLU (Natural Language Understanding). Він аналізує транскрибований текст, визначає намір користувача (що саме потрібно зробити) та виділяє ключові фрази/параметри команди. Сучасні підходи NLP досягли високої якості семантичної інтерпретації, наближеної до людського рівня. Для цього часто використовують архітектури Transformer, які попередньо навчають на великих корпусах текстів, а потім донавчають під конкретне прикладне завдання — наприклад, під набір команд для певного верстата.

Для такого донавчання збирають дані з різних джерел: описані сценарії використання, інструкції та документацію до обладнання, а також знання й типові формулювання команд від експертів. Після цього ці дані анотують (позначають мітками намірів та параметрів) і використовують для навчання моделі. У результаті трансформер визначає намір у транскрибованому тексті та витягує ключові елементи, пов’язані з керуванням ЧПК (наприклад, координати, режим роботи, вибір інструмента). Далі ці елементи зіставляються з командами, які підтримує модуль керування верстатом (ЧПК), і запускають відповідну дію.

Після виконання або під час виконання команди користувач отримує підтвердження та зворотний зв’язок. Для озвучування повідомлень застосовується модуль TTS (Text-to-Speech), який формує голосову відповідь відповідно до обраних користувачем налаштувань (гучність, мова, спосіб подання тощо). Використання готових, попередньо навчених сервісів ASR і TTS може спростити реалізацію системи, оскільки такі рішення часто вже підтримують адаптацію до темпу мовлення та акценту, а також (за потреби) роботу з кількома мовами.

1.8 Приклад реалізації голосового інтерфейсу

Розглянемо приклад застосування описаної концепції для 3-осьового фрезерного верстата з ЧПК. На рис. 1.5 подано схему потоку даних і процесів під час практичної роботи системи. Для подання структури VUI використано узагальнену діаграму нейромережевої моделі, яка ілюструє основні етапи обробки: від захоплення мовлення до інтерпретації команди та формування відповіді.

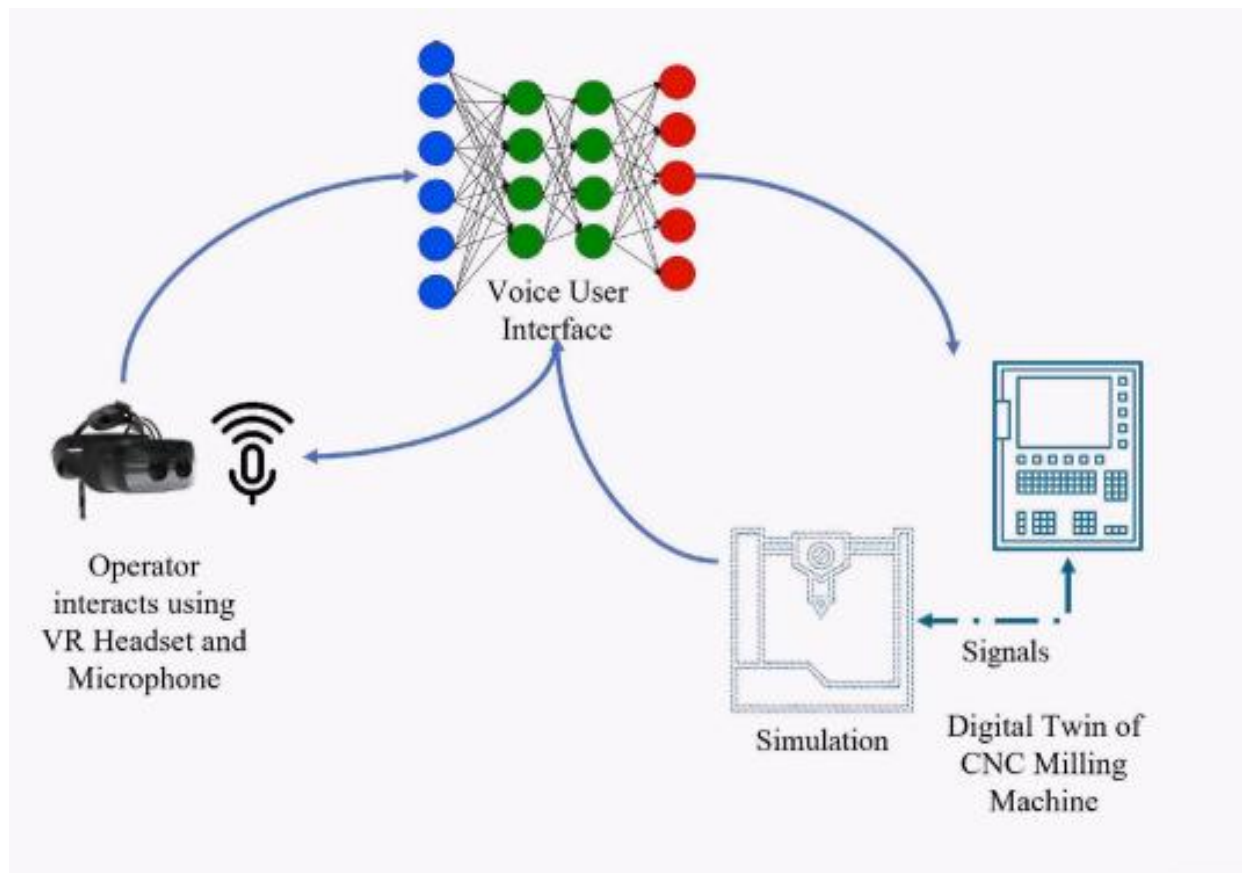


Рис. 1.5. Підхід Software-in-the-Loop (SIL) для навчання та тестування VUI на прикладі 3-осьового фрезерного верстата з ЧПК. Сині лінії показують потік даних у системі.

Команда «Hey CNC!» використовується для початку взаємодії з верстатом. Такі фрази називають словом пробудження (wake word) — вони активують систему й сигналізують, що далі буде подано команду. Аналогічний принцип застосовується в побутових помічниках, наприклад у фразі «Hey Google».

Після вимови слова пробудження модуль ASR перетворює мовлення на текст, а модуль NLP/NLU (на базі Transformer) розпізнає, що система активована. Далі модуль TTS озвучує коротке повідомлення-підтвердження (наприклад, «Готовий приймати команду»), щоб оператор розумів: інтерфейс працює та очікує наступну інструкцію.

Після активації оператор може подавати робочі команди, наприклад: «завантажити CAD-файл», «почати фрезерування», «перевірити наявність заготовки». Текст команди, отриманий від ASR, аналізується модулем розпізнавання намірів (intent classification) та модулем вилучення параметрів. На цьому етапі система визначає:

- що потрібно зробити (намір),
- які параметри потрібні для виконання (координати, напрям, значення, режими тощо).

Наприклад, фраза «перемістити шпиндель ліворуч на 10 см і вниз на 2 см» розкладається на окремі дії та параметри: переміщення по осі X на 10 см у заданому напрямку та переміщення по осі Z на 2 см. Далі ці параметри зіставляються з формальними командами, які підтримує ЧПК. Важливо, що система повинна обробляти не лише коректні формулювання, а й типові помилки: неповні команди, суперечливі параметри або небезпечні комбінації. Для таких випадків передбачаються альтернативні стратегії — уточнення, запит підтвердження або відмова у виконанні.

Після формування набору команд система створює зворотний зв'язок і, за потреби, запитує підтвердження перед виконанням. Повідомлення з підтвердженням озвучується через TTS. Оператор може підтвердити команду та,

за необхідності, запустити попередню перевірку на цифровому двійнику (DT). Це дозволяє виявити потенційні помилки або ризики для безпеки ще до того, як команда буде надіслана на реальну машину. Після успішної перевірки виконання може бути перенесене з DT на фізичний верстат через відповідні команди.

Оскільки технології ASR, TTS і NLP уже суттєво розвинені, зазвичай немає потреби навчати їх “з нуля”. Натомість ключовим стає адаптація під конкретний верстат і умови експлуатації. Для цього використовують дані від виробників і кінцевих користувачів: реальний набір ЧПК-команд, сценарії роботи, технологічні параметри. Ці дані застосовують як для побудови цифрового двійника, так і для навчання NLU-модуля. Окремо враховують акустичні умови: фонові шуми цеху моделюють або записують у реальних виробничих середовищах, а експерти предметної області допомагають сформулювати типові команди та правильні варіанти їх формулювання.

1.9 Переваги та недоліки голосових інтерфейсів для керування промисловим обладнанням

Головна перевага VUI — більш гнучка взаємодія людини з машиною. Голосове керування дозволяє працювати без використання рук, тому оператор може керувати обладнанням навіть тоді, коли виконує інші дії (наприклад, підготовку інструмента, складання, налаштування тощо).

Також голосовий інтерфейс корисний у ситуаціях, коли торкатися панелі керування незручно або небажано: у рукавичках, брудними руками тощо. Крім того, VUI підвищує доступність керування для людей з інвалідністю, які мають обмеження щодо роботи руками, що сприяє їх більшій участі у виробничих процесах.

Ще одна перевага — потенційне підвищення безпеки. Оператор може подавати команди на відстані, не перебуваючи постійно біля стаціонарної

консолі. Наприклад, керування з кількох метрів може зменшити потребу підходити до панелі заради коротких дій. Це особливо вигідно в багатоетапних операціях (зокрема під час переналаштування великого обладнання), де зростає продуктивність завдяки тому, що оператору не потрібно щоразу повертатися до терміналу.



Рис. 1.6. Віртуальне моделювання роботи 3-осьового верстата з ЧПК, що використовується як цифровий двійник.

Крім того, голосове керування майже не залежить від робочого простору, чистоти поверхонь чи умов освітлення, тому його можна використовувати навіть у складних виробничих умовах. Гнучкість і швидкість взаємодії також можна підвищити, якщо об'єднати кілька ручних дій в одну голосову команду. У багатьох випадках отримати потрібну інформацію або дати точну інструкцію

голосом швидше, ніж вручну переходити по меню інтерфейсу та виконувати серію натискань кнопок.

Разом із перевагами, впровадження VUI має і суттєві труднощі. Найбільшою проблемою в промисловому середовищі є точність роботи компонентів:

- STT (speech-to-text) — перетворення мовлення в текст,
- TTS (text-to-speech) — озвучування повідомлень,
- NLU (natural language understanding) — інтерпретація наміру та параметрів команди.

Важливо забезпечити, щоб ці модулі працювали надійно під час складних і “стресових” сценаріїв тестування, а можливі помилки не призводили до небезпечних дій. Невірне розпізнавання або неправильне трактування команди може спричинити некоректну роботу обладнання, що становить ризик як для техніки, так і для оператора.

Окрема складність — придатність голосового інтерфейсу до роботи в режимі реального часу. У критичних ситуаціях (наприклад, пошкодження інструмента, аварія, пожежа) може знадобитися миттєва зупинка обладнання, тому затримка реагування VUI стає критичною. Саме тому модулі, пов’язані з аварійними процедурами, потрібно проєктувати так, щоб мінімізувати час реакції, а також обов’язково передбачати альтернативні механізми керування (наприклад, фізичну кнопку аварійної зупинки або ручне керування).

Проблемою може бути і вимога до низької затримки в поєднанні з інфраструктурою: використання високошвидкісних мереж (наприклад, 5G) або продуктивних edge-серверів для розміщення ПЗ може бути дорогим. Тому при проєктуванні VUI важливо враховувати компроміс між можливостями моделей (зокрема, мовних моделей/LLM, якщо вони застосовуються) та обмеженнями апаратних ресурсів — пам’яті, обчислювальної потужності й енергоспоживання.

1.10 Висновки до розділу 1

Проведено аналіз проблеми голосового керування промисловим обладнанням. Показано, що традиційні методи керування мають недоліки та обмеження, які можуть бути усунуті завдяки використанню опосередкованих людино-машинних інтерфейсів, зокрема інтерфейсу голосового керування. Розглянуто етапи функціонування голосового інтерфейсу при керуванні промисловим обладнанням, зокрема етап обробки природної мови, архітектуру голосового інтерфейсу, концептуальну структуру інтерфейсу та приклади реалізації голосового інтерфейсу.

Показано, що головною перевагою таких інтерфейсів є більш гнучка взаємодія людини з машиною. Голосове керування дозволяє працювати без використання рук, тому оператор може керувати обладнанням навіть тоді, коли виконує інші дії. Також такий спосіб вирізняється підвищеною безпекою роботи для оператора. Крім того, голосове керування майже не залежить від робочого простору, чистоти поверхонь чи умов освітлення, тому його можна використовувати навіть у складних виробничих умовах.

Також виділено проблеми голосового керування, пов'язані з роботою промисловому середовищі, серед яких виділено точність роботи таких компонентів, як перетворення мовлення в текст та розпізнавання команд, озвучування повідомлень, інтерпретація наміру та параметрів команди.

РОЗДІЛ 2

ОСНОВНА ЧАСТИНА

2.1 Оцінка ефективності голосового керування промисловим обладнанням

Більшість досліджень, присвячених голосовому керуванню у виробництві, переважно описують загальні можливості та очікувані ефекти голосових інтерфейсів, часто орієнтуються на споживчі сценарії або розглядають потенціал застосування у майбутніх виробничих системах [14].

У практичних роботах зазвичай описують впровадження вже наявного (інколи відкритого) голосового застосунку, який підтримує обмежений набір команд. Такі рішення демонструвалися для:

- окремих дій промислового або навчального робота/маніпулятора [15];
- домашньої автоматизації [16];
- віртуального голосового помічника для токарного верстата [17];
- систем керування даними [18].

Мета цих підходів — показати, що голосові команди можна розпізнавати за допомогою ASR та перетворювати на дії або аудіозворотний зв'язок у середовищі, наближеному до промислового. У цілому ці дослідження підтверджують, що голосове керування машинами технічно можливе. Окремі роботи також порівнювали продуктивність та стійкість до шуму голосового керування зі звичайними HMI.

Для коректного порівняння двох інтерфейсів оператора — сенсорного та голосового — потрібно мати виробничу машину або стенд, на якому реалізовані обидва варіанти керування в однакових умовах. Хоча відомі виробники ПЛК та промислового обладнання вже працюють над впровадженням голосового

керування і відповідних стандартів для промисловості, такі рішення ще не є масово поширеними у реальних виробничих машинах [19].

Для дослідження голосового керування у виробничих машинах існує два основні підходи. Перший — використання парадигми WoZ, запропонованої Джоном Ф. Келлі. Суть методу полягає в тому, що учаснику створюють враження роботи з голосовим інтерфейсом, хоча насправді команди обробляє не система розпізнавання мовлення, а експериментатор. Він слухає зрозумілі голосові команди й вручну виконує відповідні дії через вже наявний НМІ [20]. Перевага такого підходу — не потрібно технічно реалізовувати ASR, що зменшує складність підготовки експерименту.

Однак WoZ додає в дослідження додатковий фактор — людину-експериментатора, і це створює низку проблем. По-перше, час реагування не є стабільним: швидкість виконання команд може змінюватися протягом експерименту. Це особливо критично, коли досліджується часова ефективність коротких команд.

По-друге, характер помилок у людини та в ASR різний. І експериментатор, і автоматичне розпізнавання можуть помилятися, але в реальній ASR-системі рівень точності та кількість помилок можна вимірювати і відтворювати (принаймні для кожного користувача та умов) — це робить результати більш порівнюваними. Сучасні ASR-системи в багатьох сценаріях демонструють точність близько 90% і більше, а в сприятливих умовах — ще вищу; істотний вплив має SNR. У складних акустичних умовах людина часто розпізнає мовлення краще, ніж алгоритм [21], але “якість” роботи експериментатора залежить від його втоми, уваги та когнітивного навантаження.

По-третє, в цьому дослідженні не використовувалися дуже гучні шуми, тому більш важливою є стабільність реакції (однаковий час і поведінка системи), ніж перевага людини в розпізнаванні мовлення в екстремальному шумі. Нарешті, сам факт взаємодії з “людським інтерфейсом” може викликати упередженість

учасників, навіть якщо їм не повідомляють, що замість ASR працює експериментатор [22].

Реалізований НМІ з голосовим і сенсорним керуванням

З огляду на зазначені обмеження WoZ було створено технологічну платформу Voice Control Production (VCP), яка дозволяє проводити експерименти з реальним голосовим і сенсорним керуванням. У середині технологічної камери встановлено:

- 3-осьовий фрезерний верстат (рис. 2.1, I);
- 4-осьовий робот (рис. 2.1, II).

Перед камерою розміщено:

- сенсорний екран для традиційного керування (рис. 2.1, III);
- всенаправлений мікрофон для голосових команд (рис. 2.1, IV).

Під камерою встановлено два динаміки (рис. 2.1, V). Вони використовуються для аудіозворотного зв'язку (підтвердження команд, повідомлення) та для імітації виробничого акустичного середовища.

Окрім обладнання, платформа містить магазин для заготовок і магазин для готових деталей. Оператор може:

- завантажувати заготовки роботом;
- вибирати програму обробки на фрезерному верстаті;
- виготовляти алюмінієві деталі;
- після завершення — робот переміщує готову деталь у відповідний магазин.

Також оператор має доступ до типових функцій керування: увімкнення освітлення в камері, запуск системи охолодження та змащення, ручне переміщення осей, переведення обладнання в заздалегідь визначені позиції тощо.

У системі реалізовано чотири рівні безпеки, які задають спосіб підтвердження команд:

- SL1 — короткий звуковий сигнал як підтвердження.

- SL2 — озвучення розпізнаної команди без затримки.
- SL3 — озвучення команди з 5-секундною паузою, щоб оператор міг скасувати дію.
- SL4 — максимальна безпека: команда озвучується і вимагає явного підтвердження оператором перед виконанням.

Завдяки такій структурі VCP є реалістичною моделлю автоматизованого виробництва, де оператор може працювати:

- або традиційно — виконуючи окремі команди для кожного пристрою,
- або як супервайзер — керуючи процесом на вищому рівні (координація дій системи загалом).

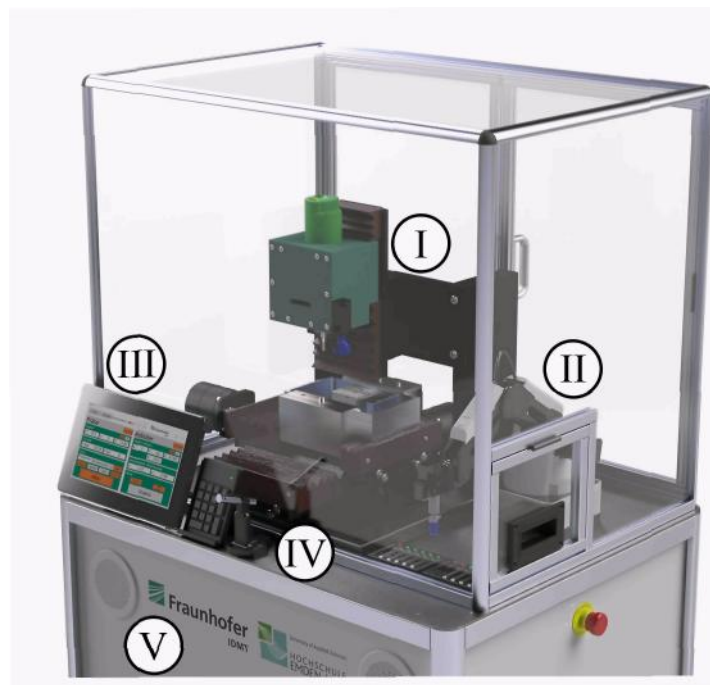


Рис. 2.1. Технологічна платформа VCP з голосовим і сенсорним НМІ
(I — 3-осьовий фрезерний верстат; II — 4-осьовий робот; III — сенсорний екран; IV — всепрямований мікрофон; V — динамік).

Технологічна платформа VCP, окрім власної системи автоматичного розпізнавання мовлення (ASR) для голосового керування, дозволяє виконувати ті

самі дії через графічний інтерфейс користувача (GUI). Сенсорні команди організовані у вигляді меню (до п'яти рівнів вкладеності), тоді як голосове керування працює за принципом “гарячого слова” (активації) та подальшої команди.

Автоматичне розпізнавання мовлення (ASR).

Обидва інтерфейси (голосовий і сенсорний) у рамках експерименту забезпечують коректне виконання команди за умови правильного введення. Тобто якщо оператор натискає потрібну кнопку або вимовляє правильну голосову команду, система має забезпечити її правильне розпізнавання і виконання. Для ASR це досягається насамперед тим, що систему навчають на невеликому спеціалізованому словнику команд, а також передбачають обробку синонімів. У результаті ASR “віддає перевагу” командним фразам платформи VCP, а не випадковим словам із загальної мовної моделі.

Щоб мінімізувати вплив акустичних факторів, рівень фонового шуму в лабораторії під час досліджень постійно контролюється і підтримується нижче 40 дБ(А). Це забезпечує прийнятне SNR навіть для учасників із тихим голосом. Такий рівень шуму, як правило, нижчий, ніж у реальних виробничих цехах, тому в умовах експерименту фоновий шум майже не впливає на якість розпізнавання. У більш шумному середовищі, особливо поруч із працюючими машинами, використання гарнітури могло б суттєво підвищити SNR.

Як ознаки використовували 40-вимірні log-Mel filterbank ознаки, сформовані з аудіо частотою 16 кГц. Архітектура акустичної моделі базувалася на поєднанні TDNN (нейромережа часової затримки) та LSTM, подібно до підходів, описаних у [23-24]. Модель навчали за критерієм максимальної взаємної інформації без решітки (lattice-free MMI) [25].

Для командної мовної моделі сформували текстовий набір на основі 33 команд та їх варіантів формулювань (загалом близько 2270 варіантів).

Графічний інтерфейс розроблено так, щоб він був схожим на інтерфейси поширених виробників ПЛК та промислового обладнання.

Ефективність з точки зору зручності використання HMI у стандарті DIN EN ISO 9241-11:2018 визначається як співвідношення між досягнутими результатами та витраченими ресурсами (наприклад, часом і зусиллями).



Рис. 2.2. Взаємодія учасників з технологічною платформою VCP

Ресурси, які витрачаються під час взаємодії з інтерфейсом, можуть бути різними: час, зусилля оператора, фінансові витрати або матеріальні ресурси. Для оцінювання ефективності HMI стандарт DIN EN ISO 9241-11:2018 рекомендує насамперед використовувати час, включно з часом, необхідним для взаємодії у разі помилки або відмови, як основний показник ефективності.

Вимірювання тривалості введення.

На практиці точно й однаково виміряти тривалість виконання кожної команди для всіх учасників складно (рис. 5). Основна проблема — коректно визначити момент початку взаємодії.

Після того як учасник читає наступну команду в інструкційному листі, він витрачає певний час на осмислення: як виконати команду через сенсорний екран або яку голосову фразу потрібно сказати. Лише після цього починається реальна

дія: натискання на екрані або вимовляння команди. Час від початку читання до першої дії складно зафіксувати без спеціальних систем, які могли б визначати момент початку читання/прийняття рішення. Тому було використано інший, більш надійний підхід: вимірювався інтервал між двома послідовними успішно виконаними командами — від моменту підтвердження попередньої команди до моменту підтвердження наступної (рис. 5, «вимірювана тривалість введення»).

Окрім об'єктивних вимірювань часу, учасники описали враження від двох способів керування.

Інтуїтивність (шкала 1–10):

- голосове керування — 9,2, сенсорне — 6,9.

Тобто голосовий інтерфейс учасники сприйняли як більш інтуїтивний (приблизно на 25%).

Когнітивне навантаження (шкала 1–10):

- сенсорне — 8,4, голосове — 4,1.

Сенсорний інтерфейс у середньому вимагав значно більше уваги й “ментальних зусиль”.

Комфортність використання (шкала 1–10):

- голосове — 8,8, сенсорне — 8,1.

Значення близькі, хоча учасники раніше не використовували голос як НМІ для виробничих машин.

Перевага у реальному сценарії:

- 8 із 15 (53,3%) обрали б лише голосове керування,

- 4 із 15 (26,7%) — голос як доповнення до сенсорного,

- 3 із 15 (20,0%) залишилися б лише з сенсорним керуванням.

2.2 Поняття голосової команди та задача розпізнавання

Розпізнавання голосових команд — це сукупність методів, які перетворюють акустичний сигнал мовлення на керувальне рішення: яку саме команду сказано і, за потреби, які параметри вона містить (числа, режими, об'єкти, напрямки тощо).

На практиці під “розпізнаванням команд” розуміють не одну задачу, а кілька взаємопов'язаних підзадач:

- 1) Виявлення мовлення (чи є в потоці мікрофона мовлення взагалі).
- 2) Сегментація фрази (де команда починається і закінчується).
- 3) Розпізнавання (класифікувати команду або розшифрувати у текст).
- 4) Інтерпретація (визначити намір/intent і параметри/slots).
- 5) Контроль помилок (відхилення невпевнених команд, підтвердження критичних дій, перевірка дозволеності команди у поточному стані системи).

Залежно від вимог (кількість команд, необхідність параметрів, робота в шумі, обмежені ресурси, затримка) застосовують різні методи, які умовно поділяють на класифікаційні та ASR+NLU підходи.

2.3 Загальна структура системи розпізнавання команд

Типовий конвеєр (pipeline) розпізнавання голосових команд містить такі блоки:

- захоплення сигналу. Мікрофон (гарнітура), оцифрування (частота дискретизації 8–48 кГц), нормалізація рівня.
- попередня обробка (front-end processing): фільтрація (наприклад, високочастотний фільтр для придушення низькочастотного гулу), автоматичне регулювання підсилення, придушення шуму (класичне або нейромережеве),
- VAD (Voice Activity Detection) визначає ділянки мовлення.

- Endpointing фіксує початок/кінець команди (важливо для коротких фраз).
- формування ознак (features) або робота з сирим сигналом.
- модель розпізнавання: або класифікація команд, або ASR (Automatic Speech Recognition) → текст, або спільні моделі intent+slots.
- постобробка та формування рішення: оцінка впевненості (confidence), відхилення “невідомих” фраз, правила безпечного виконання (підтвердження, обмеження станом системи), формування керувальної команди для виконавчого рівня.

2.4 Методи та ознаки для розпізнавання голосових команд у системах керування

Голосовий сигнал є нестационарним, тому часто аналізують його у коротких вікнах (20–30 мс) зі зсувом 10 мс. Базовий інструмент — короткочасне перетворення Фур’є (STFT):

$$X(t, \omega) = \sum_n x[n]w[n - t]e^{-j\omega n},$$

де $w[\cdot]$ — віконна функція (Hamming/Hann).

На основі STFT будують ряд ознак.

Спектральні ознаки мовлення: Log-Mel спектрограма (FBANK) та MFCC.

У більшості сучасних систем розпізнавання мовлення первинний аудіосигнал перетворюють у компактне представлення, яке зберігає основну акустичну інформацію та є зручним для подальшого машинного аналізу. Найпоширенішим варіантом є log-mel спектрограма (FBANK). Для її отримання спектр сигналу проєктують на mel-шкалу, що приблизно відповідає частотному розрізненню людського слуху. Далі для кожної mel-смуги обчислюють енергію

та застосовують логарифмування, що зменшує динамічний діапазон і наближує розподіл ознак до більш “лінійного” для моделей навчання.

Log-mel ознаки добре узгоджуються з архітектурами глибоких мереж, зокрема CNN, TCN та Transformer/Conformer, оскільки формують двовимірну часово-частотну структуру, подібну до зображення, де важливі локальні та глобальні патерни.

Альтернативою (або доповненням) виступають MFCC — кепстральні коефіцієнти, які отримують шляхом застосування DCT-перетворення до log-mel енергій. MFCC є більш компактними і історично широко використовувалися в класичних системах розпізнавання мовлення та класифікаторах, особливо в комбінації з HMM або DTW.

Нормалізація ознак.

Практичні системи суттєво залежать від акустичного тракту: типу мікрофона, відстані до нього, підсилення, а також від індивідуальних особливостей дикції. Для зменшення цього впливу застосовують CMVN (Cepstral Mean and Variance Normalization) — нормалізацію середнього значення та дисперсії ознак. CMVN підвищує стабільність розпізнавання при зміні обладнання й умов запису та зменшує систематичні спотворення, що вносяться мікрофоном і каналом передачі.

2.5 Підходи до розпізнавання команд при голосовому керуванні

2.5.1 Виявлення ключових слів і команд (Keyword Spotting, KWS).

Keyword Spotting (KWS) — це клас методів, які визначають, чи присутнє задане ключове слово або фраза в аудіопотоці (у реальному часі) або в короткому сегменті запису. KWS особливо ефективний у задачах:

- активації системи за “словом пробудження” (wake word);

- розпізнавання коротких критичних команд типу «стоп», «пуск», «додому».

Класичні методи KWS.

До традиційних підходів належать:

Шаблонне зіставлення + DTW (Dynamic Time Warping) за MFCC: команда порівнюється з еталонними зразками, а DTW виконує вирівнювання за часом, компенсуючи різну швидкість вимови.

HMM (Hidden Markov Models): слово моделюється як послідовність станів (фонем або їхніх підмножин), що дозволяє враховувати часову структуру мовлення та варіативність реалізації.

Нейромережеві KWS.

Сучасний домінуючий підхід у KWS — це класифікація спектральних ознак (найчастіше log-mel спектрограми) нейромережевою моделлю. Мережа оцінює ймовірність наявності ключового слова у вікні аудіо, що дозволяє реалізувати низьку затримку та високу точність при невеликому словнику команд.

2.5.2 Класифікація голосових команд (Speech Command Classification).

Класифікація голосових команд є узагальненням KWS. На вхід подається фіксований за тривалістю фрагмент (типово 1–2 секунди), а на виході система повертає клас команди з наперед визначеного набору. На практиці цей підхід доцільний для керування пристроями з обмеженою кількістю команд, де важлива швидкість прийняття рішення та контроль помилок.

Переваги: мала затримка (можливість швидкого реагування); менші вимоги до обсягу навчальних даних порівняно з повним ASR; спрощений контроль помилок (можна явно вводити клас “невідома команда” та пороги впевненості).

Обмеження: слабка масштабованість на “вільну” мову; складно підтримувати велику кількість формулювань та параметризовані команди без ускладнення моделі й датасету.

2.5.3 ASR - NLU (транскрипція мовлення з подальшим «розумінням»)

У підході ASR - NLU система спочатку виконує розпізнавання мовлення (ASR), отримуючи текстову транскрипцію, а потім компонент NLU перетворює цей текст у формалізовану керувальну структуру.

Цей підхід дає змогу працювати з природнішими фразами та комбінаціями параметрів, але підвищує вимоги до надійності та безпеки.

Класичний ASR (гібрид DNN–HMM). Історично ASR будувався як комбінація: акустичної моделі, мовної моделі, словника вимови, із процедурою декодування. Для командних систем важливим є використання обмеженої граматики, яка “дозволяє” лише валідні командні фрази. Це зменшує кількість хибних розпізнавань і підвищує керованість поведінки системи.

Переваги: висока гнучкість формулювань (природні фрази); простіше додавати нові команди та параметри без повної перебудови системи.

Недоліки: помилка ASR може призвести до небезпечної інтерпретації параметра (наприклад, неправильне число або вісь); вищі вимоги до даних і обчислювальних ресурсів; складніша логіка безпечної відмови та перевірки “сумнівних” команд.

Граматичні методи та обмеження простору команд. Для систем промислового керування часто доцільно обмежувати “мову команд”, щоб підвищити надійність і зменшити ризик хибного спрацювання.

Найнадійніший варіант — перелік стандартних команд із можливими синонімами. Такий підхід мінімізує неоднозначність і спрощує валідацію.

Формальні граматики. Формальні граматики обмежують простір можливих фраз і зменшують кількість помилкових розпізнавань, оскільки система не “розглядає” фрази, яких немає у правилах. Для промислового керування це часто є компромісом між гнучкістю та безпекою.

Робастність до шуму та варіативності мовлення.

Якість розпізнавання в реальних умовах знижують шум, реверберація, різні типи мікрофонів і відстань до них, а також індивідуальні відмінності дикторів (дикція, темп, акцент). Тому стійкість системи — одна з ключових вимог для промислового застосування.

Застосовують: спектральні методи подавлення шуму (зокрема спектральне віднімання), адаптивні фільтри, нейромережеві моделі шумопригнічення (denoise).

Навчання з аугментацією. Найефективніша стратегія підвищення робастності в ML — аугментація: додавання реальних шумів із різними SNR, моделювання реверберації через імпульсні характеристики приміщень (RIR), застосування SpecAugment до спектрограм.

Багатоступенева логіка безпеки.

Для керування обладнанням доцільно застосовувати багатоступеневі схеми: спочатку wake word - потім команда, відхилення команд при низькій впевненості, обов'язкове підтвердження критичних дій.

Оцінювання якості та метрики.

Для командних систем використовують такі метрики:

- Accurasy / F1 по класах команд;
- матрицю змішування (confusion matrix) для аналізу, які команди найчастіше плутаються;
- FAR/FRR (false accept / false reject), особливо для wake word;
- Latency (затримка від завершення фрази до рішення);
- оцінювання робастності до шуму (залежність якості від SNR).

2.6 Висновки до розділу2

На прикладі відомих досліджень показано переваги застосування голосового керування промисловим обладнанням, зокрема в плані зручності,

природності, зниження когнітивного навантаження на оператора, комфортності керування. Попри це, існує ряд обмежень та недоліків відомих систем організації голосового керування, які потребують покращення.

Розглянуто задачу розпізнавання голосових команд, що включає підзадачі виявлення мовлення, сегментація фрази (де команда починається і закінчується), розпізнавання (класифікувати команду або розшифрувати у текст), інтерпретація, контроль помилок.

Власне в дослідженні проводиться розробка методу розпізнавання голосових команд, в якому пропонуються шляхи вирішення перших трьох підзадач.

Розглянуто методи розпізнавання голосових команд, які розділено за трьома базовими підходами: KWS/класифікація команд (найпростіший і швидкий варіант для малого словника; забезпечує низьку затримку та відносно простий контроль помилок), ASR - NLU (більш гнучкий і масштабований підхід, придатний для команд із параметрами та різними формулюваннями, але він потребує ретельних механізмів безпеки, перевірки впевненості та більших ресурсів), граматично обмежені методи (компроміс між надійністю та підтримкою параметризованих команд; часто є найбільш практичним рішенням для промислового керування, де критичними є безпечність і прогнозованість поведінки системи).

Розроблюваний в дослідженні метод розпізнавання голосових команд будується на основі першого підходу, через можливість забезпечення низького значення затримки на опрацювання та прості способи контролю помилкових спрацювань.

РОЗДІЛ 3

НАУКОВО-ДОСЛІДНА ЧАСТИНА

3.1 Принцип роботи пропонованого способу

Запропонований спосіб призначений для обробки голосових сигналів з метою виділення слів-команд та формування на їх основі сигналів керування. У межах роботи передбачається розпізнавання чотирьох керувальних команд: «вліво», «вправо», «взад», «вперед».

Загальна послідовність роботи методу.

На першому етапі користувач вимовляє команду в мікрофон. Далі система виконує попередню обробку запису, виділяє інформативні ознаки та здійснює ідентифікацію команди. Для реалізації методу використано підходи, описані в попередніх розділах, із певними модифікаціями, що враховують особливості коротких команд та умови практичного застосування.

Послідовність обробки сигналу включає такі основні кроки:

1) Запис мовного сигналу.

Отримується цифровий сигнал із мікрофона для подальшої обробки.

2) Попередня фільтрація та подавлення завад.

Сигнал проходить обробку адаптивними методами фільтрації, що дозволяє зменшити вплив шуму та покращити співвідношення сигнал/шум.

3) Визначення часових меж команди.

Виконується пошук початку та кінця вимовленого слова, щоб відокремити команду від пауз і фонових звуків.

4) Фонемна сегментація слова.

Виділене слово поділяється на ділянки, які відповідають окремим фонемам або їх групам.

5) Виділення ділянок голосних і приголосних (вокалізованих) звуків.

На завершальному етапі визначаються фрагменти, що відповідають голосним та приголосним вокалізованим звукам, які мають характерні енергетичні та спектральні ознаки.

Ідея розпізнавання на основі фонемного складу.

Кожна з чотирьох команд має відмінний фонемний склад і різне чергування звуків, що можна використати як основу для розпізнавання. Зокрема:

- команда «вліво» містить 5 фонем: 2 голосні та 3 приголосні вокалізовані;
- команда «вправо» містить 6 фонем: 2 голосні, 3 приголосні вокалізовані та 1 глуха;
- команда «взад» містить 4 фонemi: 1 голосна та 3 приголосні вокалізовані;
- команда «вперед» містить 6 фонем: 2 голосні, 3 приголосні вокалізовані та 1 глуха.

Оскільки послідовність фонем у кожному слові різна, для виконання фонемної сегментації як інформативну ознаку доцільно використовувати частоту основного тону (F_0). Вона добре відображає наявність вокалізованих ділянок (особливо голосних) і дозволяє відокремлювати їх від глухих або слабо вокалізованих фрагментів.

Виявлення часових меж команди.

Першим практично важливим кроком є визначення часових меж вимовленого слова. Для цього застосовується метод із ковзним вікном, у межах якого обчислюється усереднене значення (або енергетична характеристика) голосового сигналу. Далі отримане значення порівнюється з заданим порогом:

- якщо усереднена величина дорівнює або перевищує поріг, то значення порогової функції h приймається рівним 1 (команда присутня);
- якщо величина менша за поріг, то $h=0$ (пауза або шум).

Цей механізм дозволяє відокремити активні ділянки мовлення від пауз та забезпечує стабільний початок подальших етапів обробки.

Щоб підвищити надійність у шумних умовах, на практиці замість “середнього значення сигналу” часто використовують короточасну енергію та/або частоту переходів через нуль (ZCR): енергія добре виділяє мовні фрагменти, а ZCR допомагає розрізняти вокалізовані та глухі ділянки. Однак реалізація такого підходу є складнішою та потребуватиме більших обчислювальних потужностей та часу на обробку.

На наступному етапі виконується оцінювання частоти основного тону F_0 — важливої інформативної (зокрема індивідуальної) характеристики мовлення. Як правило, F_0 чітко проявляється на ділянках, що відповідають голосним і вокалізованим приголосним фонемам, тобто звукам із вираженою періодичністю коливань голосових зв’язок. Саме тому аналіз F_0 доцільно використовувати для подальшої сегментації команди та розрізнення типів фонем.

Для визначення F_0 застосовано підхід, що ґрунтується на спектральному (формантному) аналізі: значення частоти основного тону ототожнюється з частотою першого домінуючого максимуму у спектрі потужності. Розрахунок виконується у режимі ковзного вікна, що дозволяє відстежувати зміну F_0 у часі та локалізувати ділянки з регулярною періодичною структурою.

Додатково варто зазначити, що на практиці для підвищення стійкості оцінки F_0 до шуму доцільно обмежувати пошук у типових межах голосового діапазону (наприклад, для дорослого диктора) та застосовувати згладжування траєкторії F_0 між сусідніми вікнами, щоб зменшити вплив випадкових спектральних піків.

На завершальному етапі виконується ідентифікація голосової команди. Для цього невідомий вхідний сигнал аналізується у режимі ковзного вікна. Спочатку визначається наявність частоти основного тону (F_0). Очікується, що F_0 спостерігається лише на тих ділянках сигналу в межах тривалості команди, які відповідають голосним та вокалізованим приголосним звукам.

Далі за результатами аналізу формуються часові інтервали, у яких F_0 присутня. Після цього команда розпізнається за послідовністю, кількістю та тривалістю таких інтервалів, оскільки для чотирьох заданих команд ці характеристики мають відмінні значення. Після визначення команди формування відповідних сигналів керування є технічно нескладним і реалізується шляхом зіставлення розпізнаної команди з потрібною дією виконавчого механізму.

Повна схема алгоритму ідентифікації команд показана на рис. 3.1.

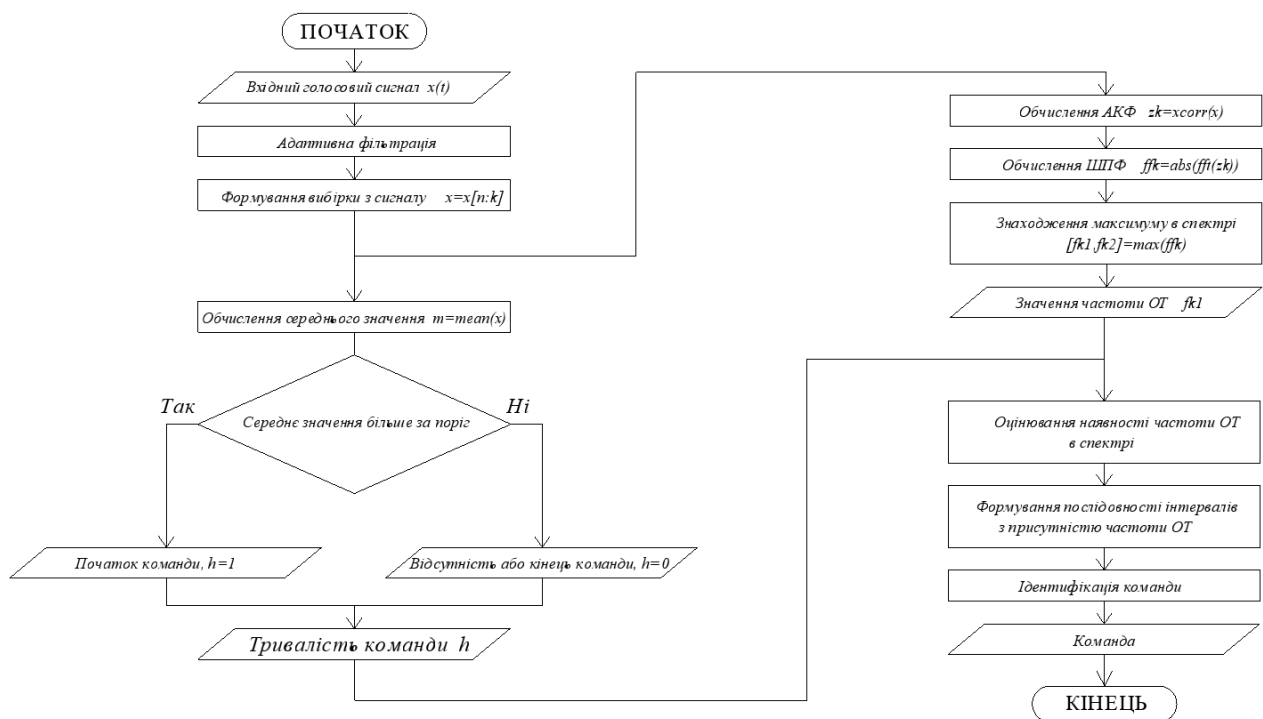


Рис. 3.1. Повна схема алгоритму ідентифікації команд для голосового керування

Також було розроблено програму дл Матлаб із графічним інтерфейсом, яка реалізує пропонуванний алгоритм обробки. Її вигляд наведено на рис. 3.2.

Власне у верхніх вікнах будуть висвічуватись вихідний сигнал, детектований сигнал та сигнал порогової функції, а в нижніх – сигнал однієї команди та відповідна порогова функція.

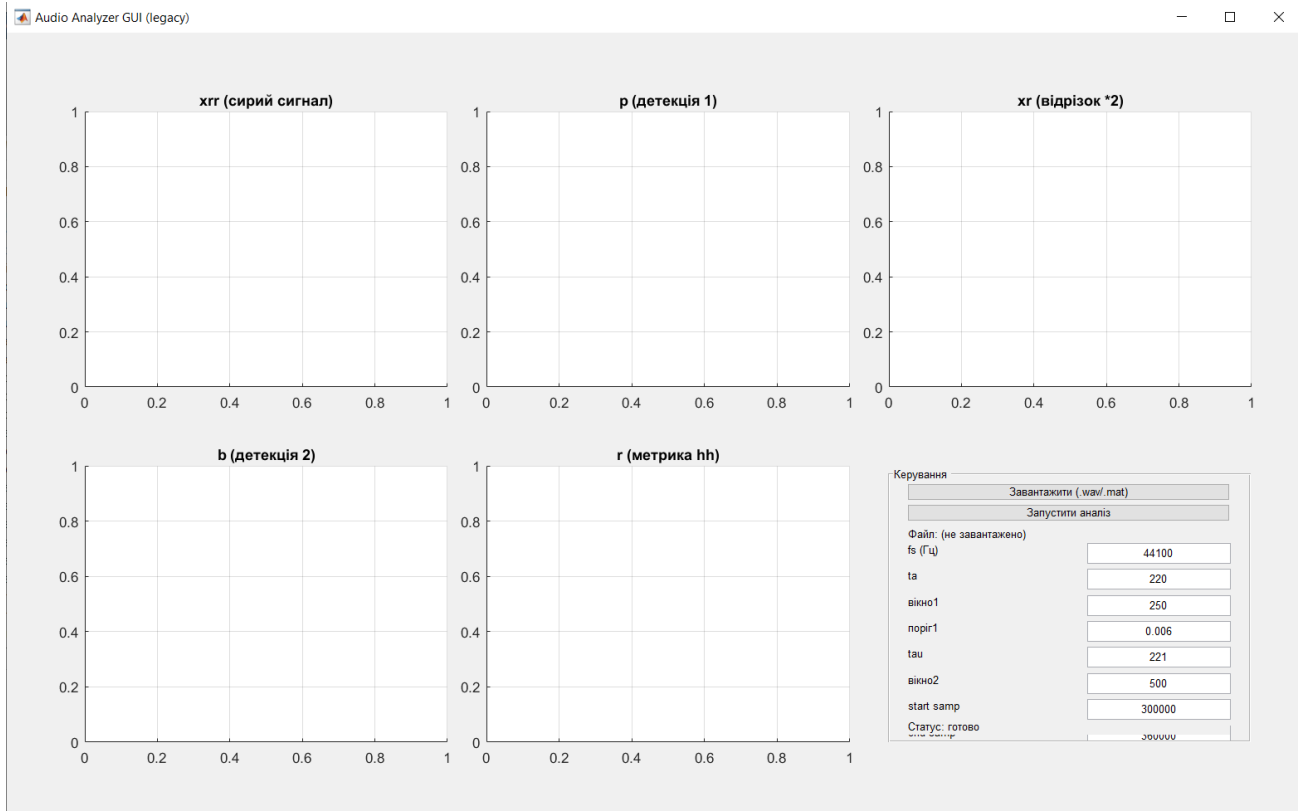


Рис. 3.2. Вигляд програми опрацювання голосових сигналів

3.2 Експериментальні результати

Для формування набору голосових сигналів із командами було використано стандартну комп'ютерну гарнітуру (мікрофон) та програмне середовище Adobe Audition, яке забезпечує зручний запис і попередній перегляд аудіосигналів. Записані фрагменти містили послідовне вимовляння чотирьох команд: «вліво», «вправо», «взад», «вперед».

Для подальшого аналізу аудіодані було збережено у цифровому форматі без додаткових ефектів обробки (окрім базового запису), щоб оцінювання роботи методу виконувалося на максимально “реальному” сигналі. Приклад часової реалізації записаного мовного сигналу, що містить усі чотири команди, наведено на рис. 3.3.

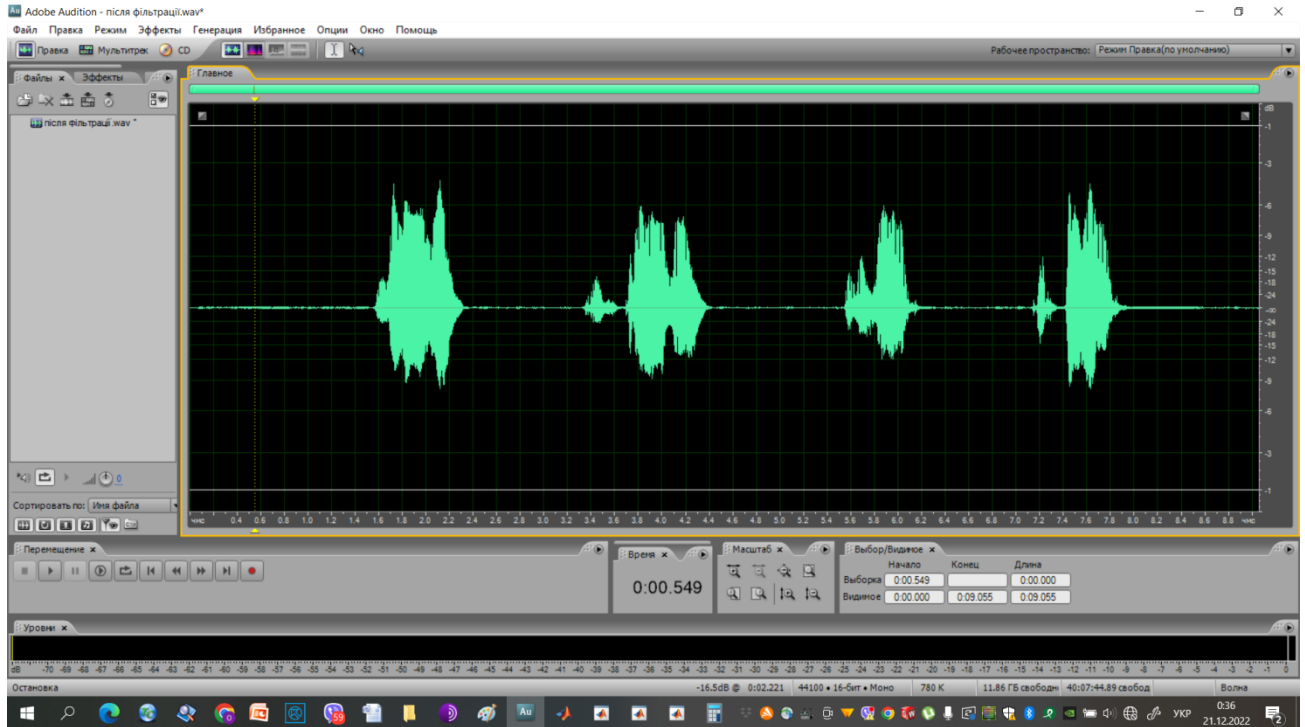


Рис. 3.3. Часова реалізація записаного мовного сигналу з чотирма командами.

Оскільки в записі наявні шумові складові, під час обробки на кожному етапі важливим є застосування адаптивного подавлення шуму. Для цього використано метод спектрального віднімання, який реалізований у програмі Adobe Audition. Суть підходу полягає в оцінюванні спектра шуму (як правило, за паузами або фоновими ділянками) та його подальшому відніманні зі спектра сигналу, що дозволяє підвищити співвідношення сигнал/шум і зробити мовні компоненти більш вираженими.

Часова форма сигналу після фільтрації та відповідна спектрограма наведені на рис. 3.4 і рис. 3.5 відповідно.

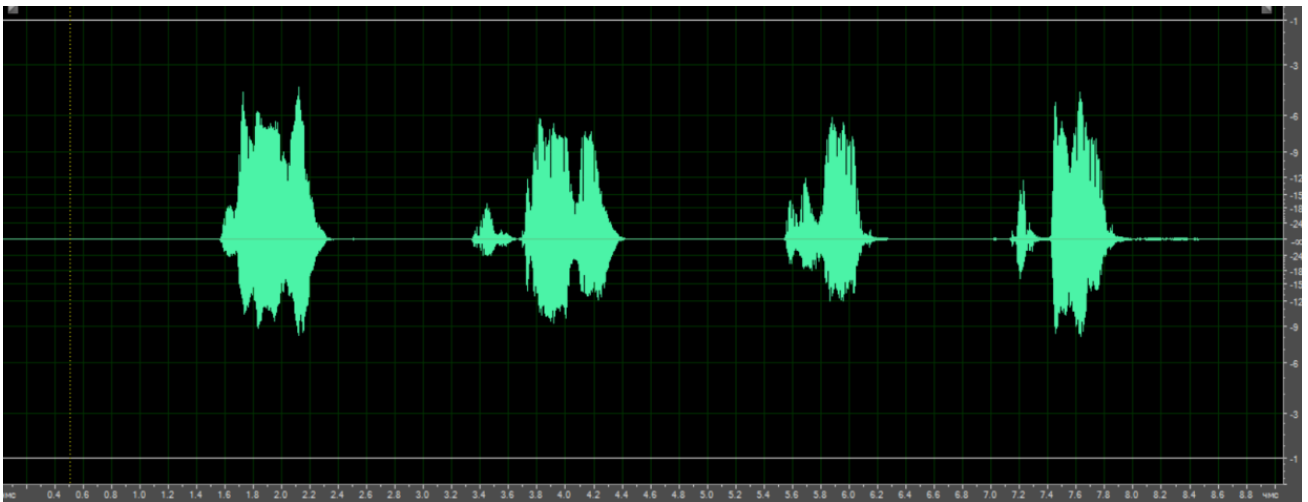


Рис. 3.4. Сигнал після адаптивної фільтрації

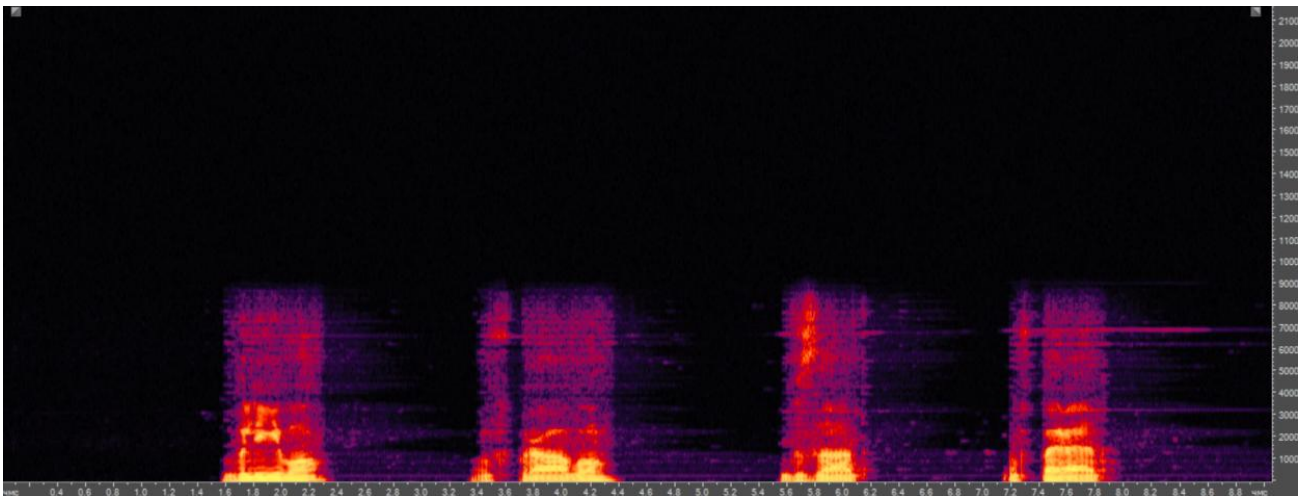


Рис. 3.5. Спектрограма після фільтрації

Далі очищений аудіосигнал було імпортовано в середовище MATLAB, де виконувалася його подальша обробка відповідно до запропонованого алгоритму. Часова реалізація цього сигналу наведена на рис. 3.6.

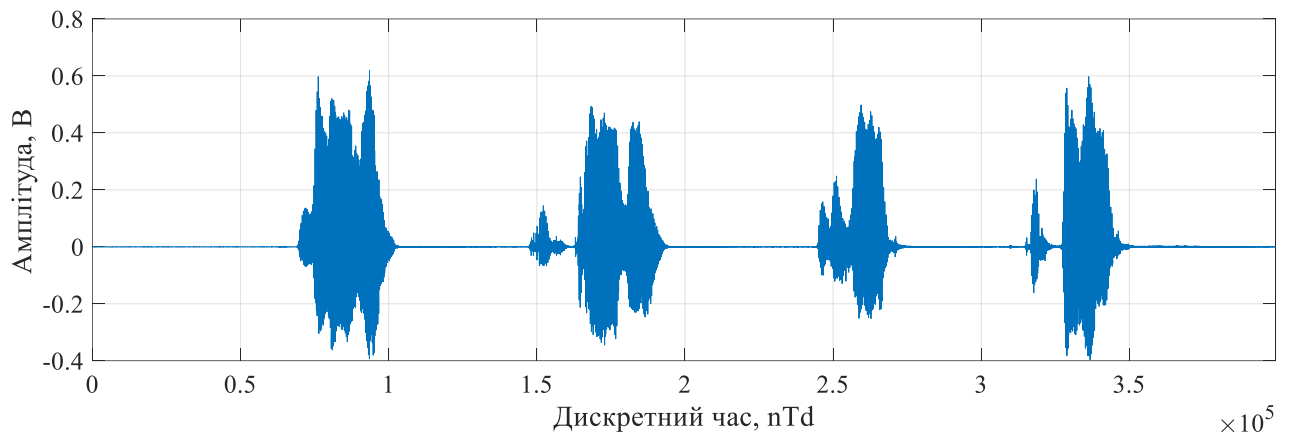


Рис. 3.6. Команди голосового сигналу

Далі, відповідно до алгоритму на рис. 3.1, виконано визначення часових меж (тривалості) кожної команди. Для цього використовувалося ковзне вікно тривалістю 5 мс. Таке значення було обрано апріорно, оскільки воно є суттєво меншим за тривалість вимовляння команди (понад у 100 разів), що забезпечує достатню часову роздільну здатність під час виявлення початку та завершення слова.

Порогове значення для прийняття рішення також задавали апріорно, з урахуванням рівня сигналу після шумоподавлення. Результат роботи порогового детектора — значення порогової функції $h(t)$, яка приймає 1 на ділянках мовної активності та 0 на паузах — наведено на рис. 3.7.

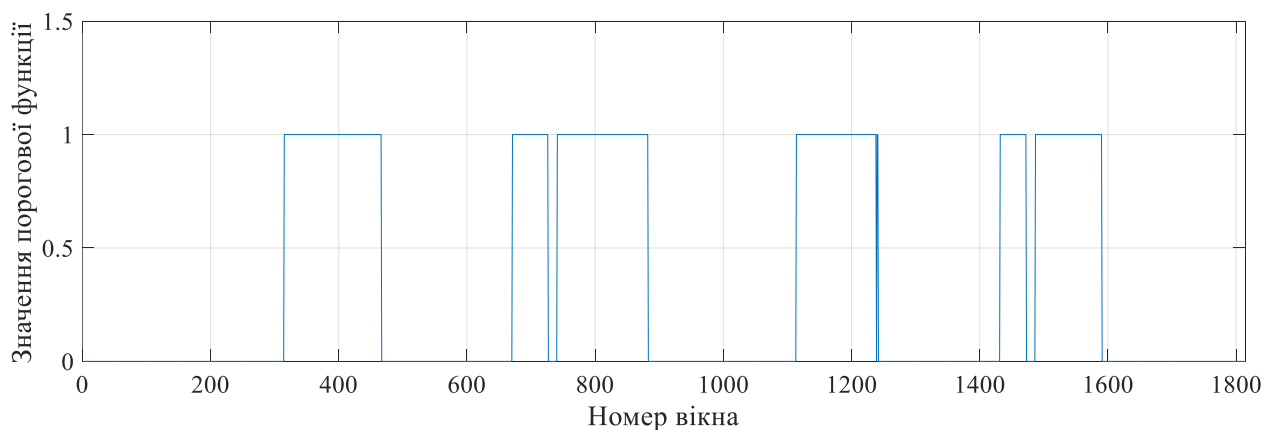


Рис. 3.7. Значення порогової функції $h(t)$.

Як видно з отриманого результату, порогова функція чітко виділяє ділянки мовної активності, тому за її переходами між значеннями 0 і 1 можна надійно визначити початок, кінець і тривалість кожної голосової команди.

На наступному етапі, відповідно до алгоритму на рис. 3.1, було виконано оцінювання частоти основного тону. Для наведеного прикладу значення F_0 становило приблизно 172 Гц.

Далі на завершальному етапі здійснено фонемну сегментацію кожної команди — тобто поділ слова на характерні звукові ділянки (зокрема голосні та вокалізовані приголосні), що використовуються як ознаки для ідентифікації. На рис. 3.8 показано часову реалізацію сигналу для першої команди «вліво».

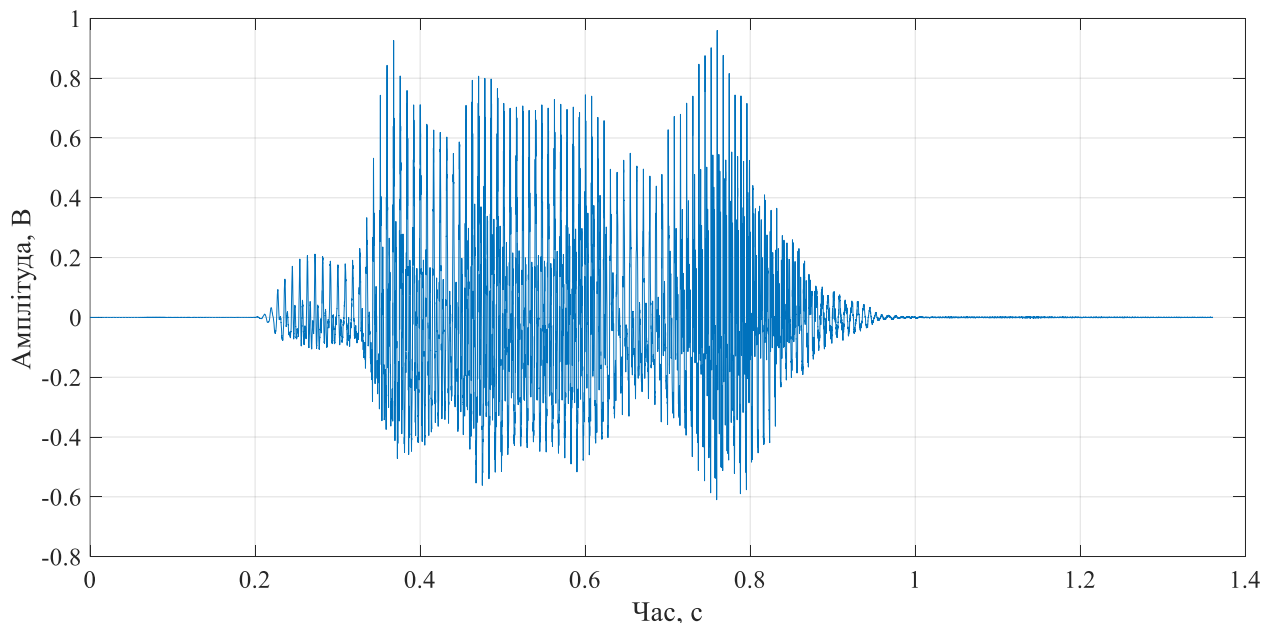


Рис. 3.8. Вигляд голосового сигналу команди «Вліво»

Далі, згідно з описаним вище алгоритмом, було сформовано оцінки спектра сигналу та виконано їх усереднення в околі F_0 . Обробка здійснювалася в режимі ковзного вікна, що дозволяє відстежувати спектральні зміни в часі.

Для кожного положення вікна обчислювався спектр, після чого визначалося середнє значення амплітуд (або потужності) в вузькій смузі частот навколо F_0 . Якщо в межах даного фрагмента сигналу присутня F_0 , то це проявляється як виражений пік (або підвищена енергія) поблизу F_0 , і усереднене значення в цій смузі зростає.

Отримані усереднені значення для кожного вікна було нанесено на спільну часову вісь відповідно до поточного положення ковзного вікна. Таким чином сформовано часовий графік, який відображає наявність/відсутність F_0 у різних ділянках слова та може використовуватися для подальшої сегментації. Результат наведено на рис. 3.9.

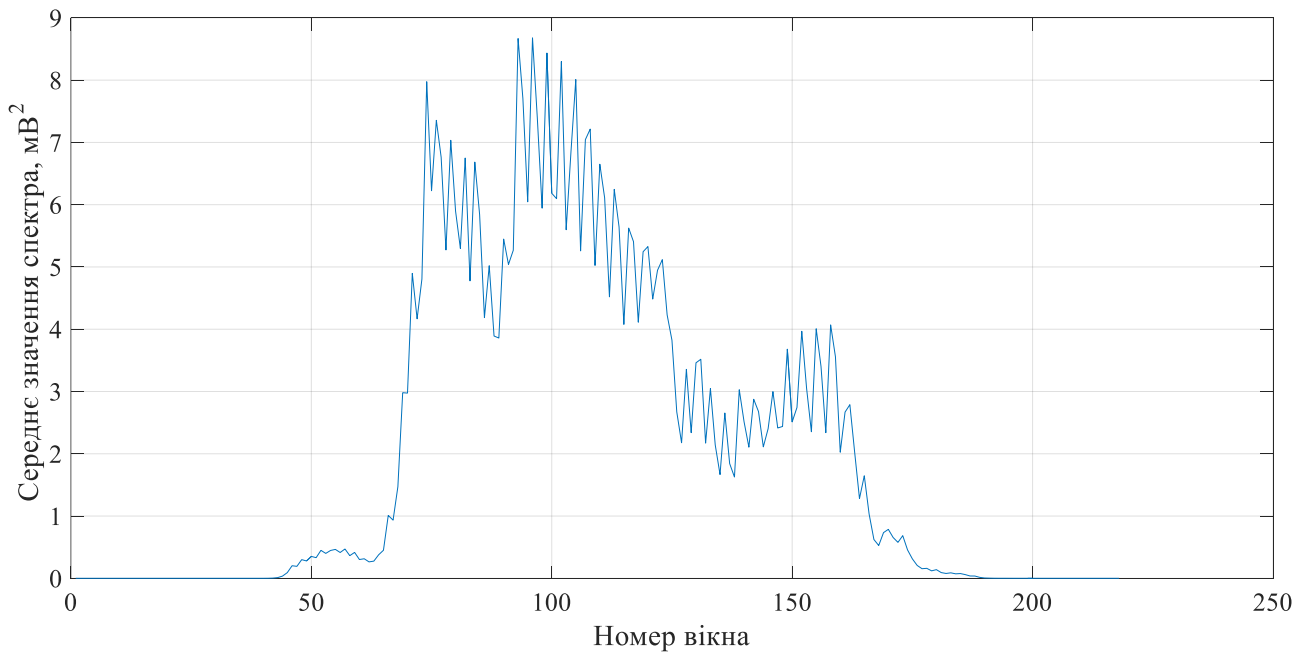


Рис. 3.9. Часова залежність показника, що відображає наявність F_0 у мовному сигналі.

На основі отриманого графіка далі сформовано порогову функцію, яка дозволяє автоматично виділяти інтервали з наявною F_0 . Для цього було задано поріг 0,1 (обраний апріорно), і відповідно:

- якщо значення показника перевищує поріг — фіксується наявність F_0 ;

- якщо не перевищує — вважається, що F0 відсутня або недостатньо виражена.

Графік порогової функції наведено на рис. 3.10.

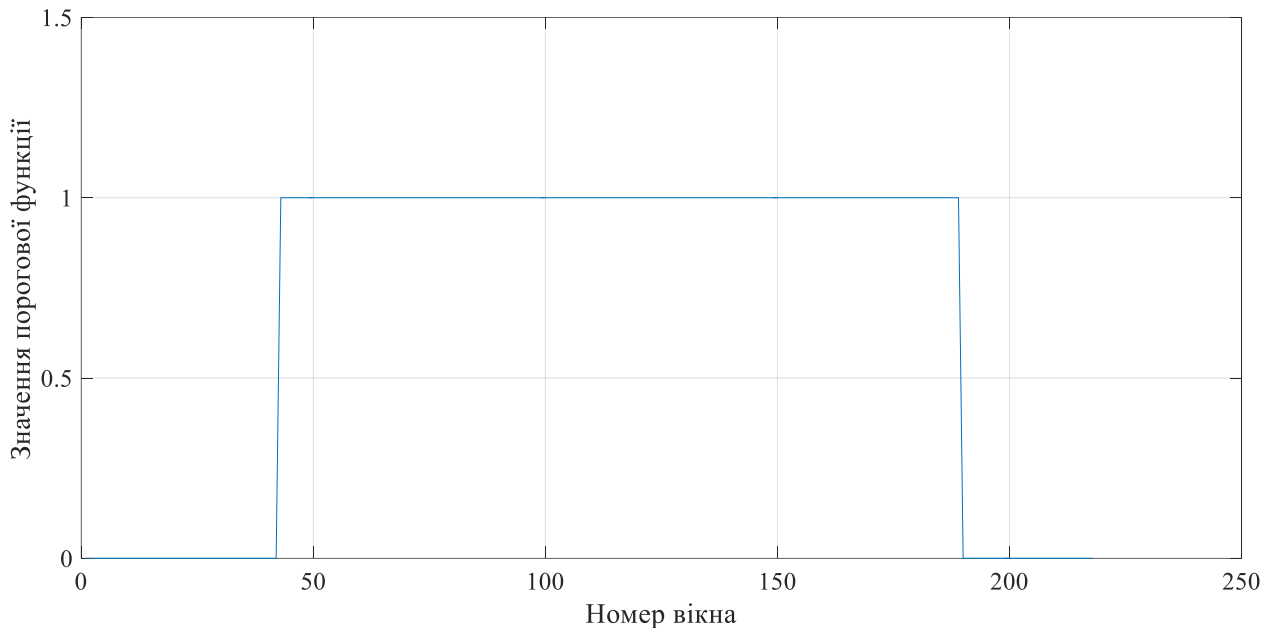
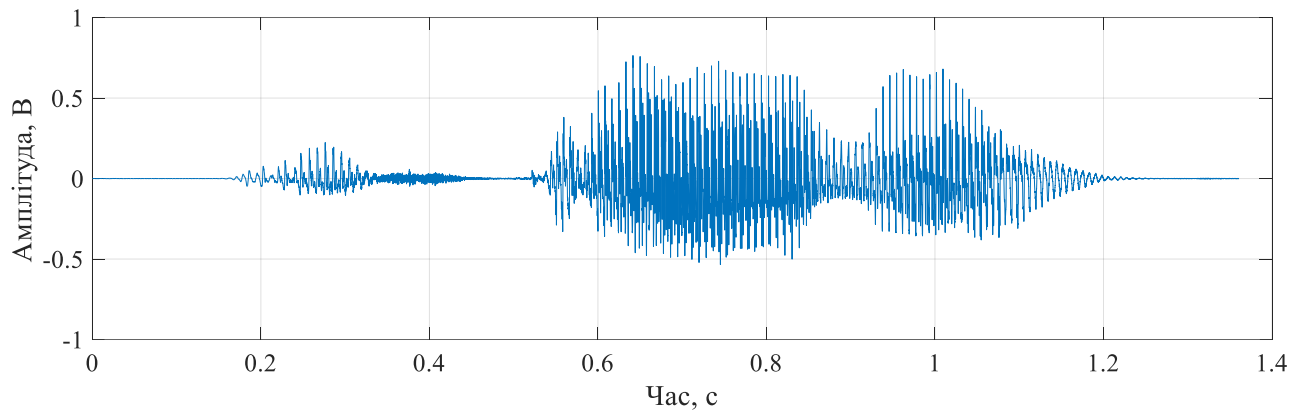


Рис. 3.10. Порогова функція для виявлення ділянок із наявною F0

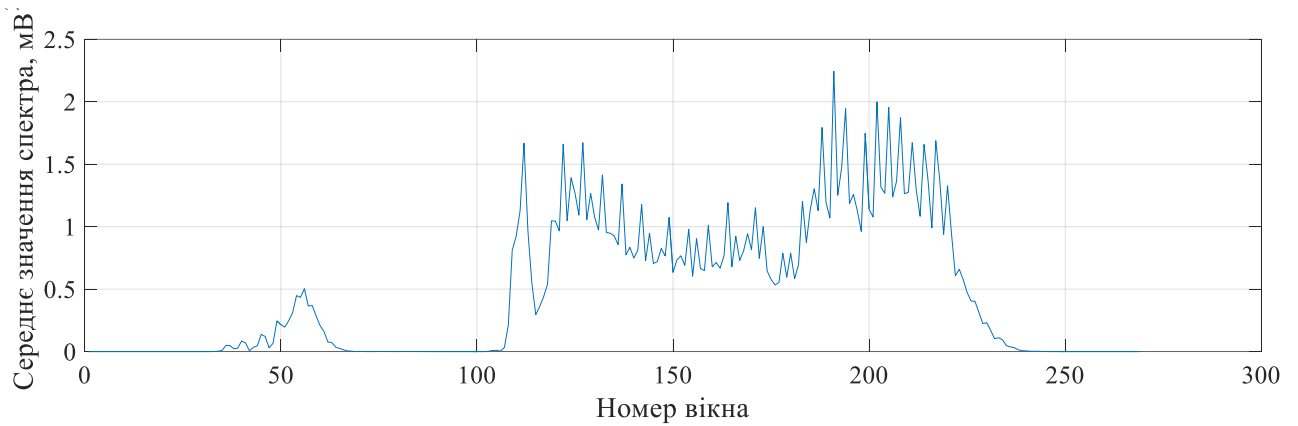
.

З отриманого графіка видно, що F0 присутня практично протягом усієї тривалості команди. Це узгоджується з фонетичною структурою слова, оскільки більшість його звуків належать до вокалізованих (тобто таких, що мають виражену періодичну складову), зокрема голосних і дзвінких/вокалізованих приголосних. У результаті на більшій частині сигналу спостерігається стабільний прояв F0.

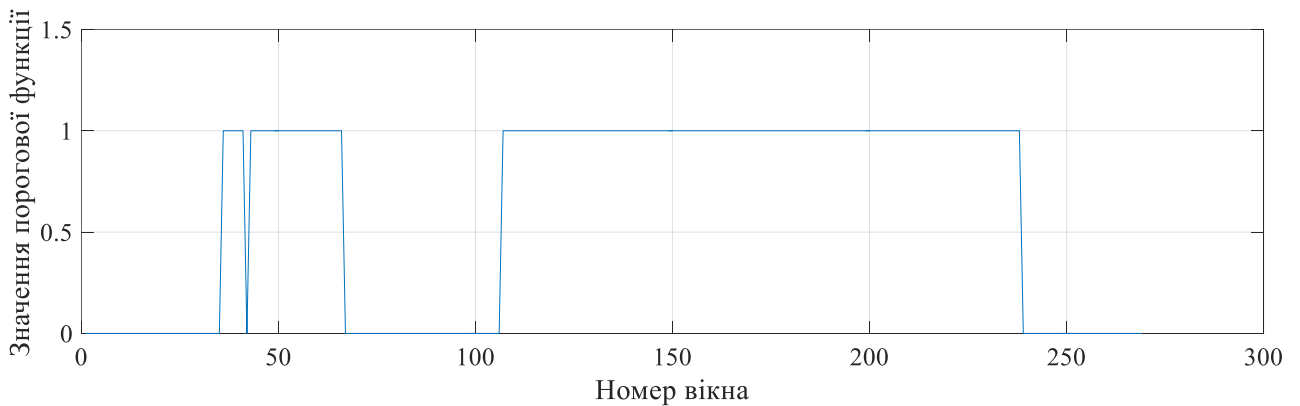
На рис. 3.11 наведено результати для другої команди «вправо»: часову реалізацію сигналу, графік показника наявності F0 та відповідну порогову функцію, сформовану за тим самим принципом.



а)



б)



в)

Рис. 3.11. Сигнал команди «вправо»: (а) часова реалізація, (б) показник наявності F0, (в) порогова функція.

У цьому випадку на графіку помітний короткий інтервал “провалу”, тобто відсутності вираженої F0. Така ділянка відповідає глухій приголосній фонемі [п],

для якої характерна відсутність регулярної періодичності голосових зв'язок і, відповідно, відсутність стабільної F0. Це підтверджує, що обрана ознака дозволяє відрізняти вокалізовані та глухі звуки в межах команди.

На рис. 3.12 наведено повне вікно розробленої програми із основними результатами опрацювання.

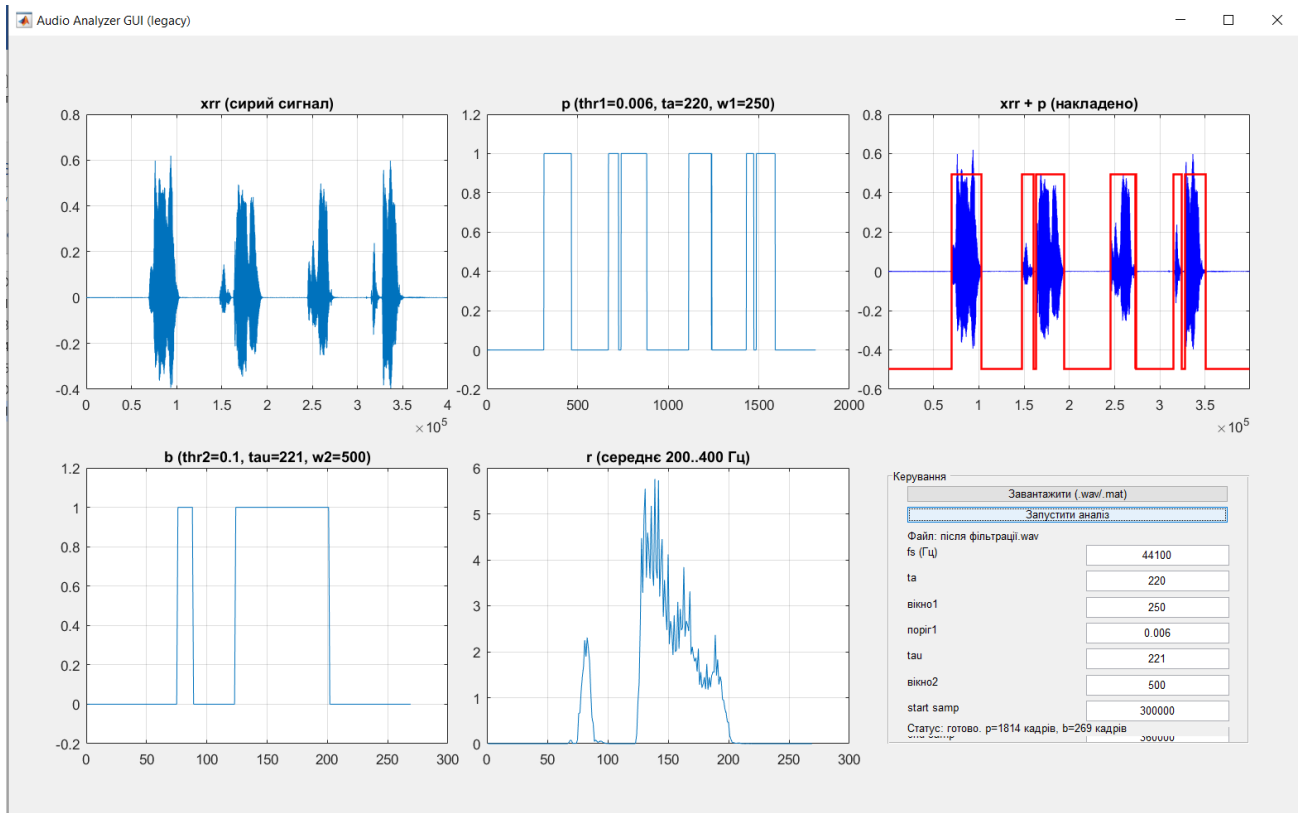


Рис. 3.12. Вікно розробленої програми із основними результатами опрацювання.

Загалом порівняння результатів для всіх чотирьох команд показує, що форма порогової функції, кількість «активних» ділянок та їх тривалість відрізняються. Це дає підстави стверджувати, що за сукупністю двох ознак — тривалості команди та послідовності інтервалів із наявною/відсутньою F0 — можна виконувати ідентифікацію команди та формувати відповідні керувальні сигнали.

3.3 Висновки до розділу 3

У розділі розроблено та апробовано метод обробки голосових сигналів, призначений для розпізнавання окремих слів-команд і подальшого формування керувальних впливів. Метод орієнтований на чотири команди: «вліво», «вправо», «взад», «вперед».

Суть методу така: користувач вимовляє команду в мікрофон, після чого система виконує обробку сигналу та виділяє інформативні ознаки, за якими здійснюється розпізнавання. Основними етапами методу є: запис мовного сигналу та підготовка даних для аналізу; адаптивна фільтрація з метою зменшення впливу шумів; визначення часових меж (тривалості) команд методом ковзного вікна; фонемна сегментація команд на основі аналізу частоти основного тону. Частоту основного тону оцінювали методом формантного (спектрального) аналізу, а її наявність у кожному ковзному вікні фіксували через другу порогову функцію: вона дорівнює 1, якщо F_0 присутня (вокалізовані ділянки), і 0 — якщо відсутня (наприклад, глухі приголосні або міжфонемні переходи).

Експериментальні результати для всіх чотирьох команд підтвердили, що послідовність інтервалів із наявною/відсутньою частотою основного тону, а також загальна тривалість команди відрізняються для різних слів. Отже, ці характеристики можуть бути використані як критерії для ідентифікації команди та подальшого формування відповідного сигналу керування.

Для підвищення стійкості методу до зміни гучності, темпу мовлення та шуму доцільно виконувати нормалізацію рівня сигналу, а порогові значення визначати не лише апіорно, а й адаптивно (за статистикою фонових ділянок запису).

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1 Охорона праці

Акустичне середовище є важливим компонентом середовища існування: людина живе у світі звуків. Параметри акустичного середовища можуть суттєво впливати на загальний стан людини та її працездатність і успіхи діяльності (у системах зв'язку робота оператора пов'язана з прийманням сигналів). Екстремальні умови виникають, якщо людина через сторонній звуковий тиск не може розпізнати корисні сигнали і якщо рівень звукового тиску наближається до больової межі.

Величина звукового тиску звичайно оцінюється в децибелах (дБ). Шепіт людини, який сприймається на відстані 1,5...2 метри від того, хто говорить, становить 1...18дБ; шум у приміщенні, де працюють люди, але нема машин (студентська аудиторія без чутної лекторської мови), - 40дБ; мова людини середньої нормальної гучності, що приймається людиною, котра стоїть поруч - 60 дБ; шум двигуна легкового автомобіля - 75дБ; шум електропоїзда метро, який сприймається на віддалі 3 м від нього - 95дБ; шум реактивного літака - 115дБ.

Вже при рівні звуку 100 дБ виникає загальна втома, знижується працездатність та якість праці. При рівні звуку 100...120 дБ шум та звук викликають гнітючу дію. При рівні звуку (шуму) 110 дБ неможливе мовне спілкування. Больова межа рівня звуку становить 120-130 дБ.

При проектуванні робочих місць рівень звуку понад 80 дБ вважається недопустимим. Рекомендований рівень звуку в приміщеннях для конструкторських та теоретичних робіт і опрацювання експериментальних даних - 50 дБ, в приміщеннях керування, робочих кімнатах - 60 дБ, на робочих місцях у виробничих приміщеннях - не більше 80 дБ.

Якщо рівень звукового тиску перевищує допустимий, використовують індивідуальні та колективні засоби захисту (ізоляція джерела звуку чи робочих приміщень) - вушні заглушки, навушники, шоломи.

Механічні коливання, які виникають при роботі працюючих машин, можуть викликати не лише звуки та шуми, але й вібрації.

Вібрація - це механічні коливання пружних тіл, які характеризуються амплітудою, швидкістю та прискоренням.

За характером дії вібрацію поділяють на загальну та місцеву. Загальні вібрації діють на все тіло людини. Внутрішні органи людини утворюють коливні системи з власною частотою коливань (в межах десятків та сотень Гц). Резонансна частота серця, живота та грудної клітки - 5 Гц, голови - 20 Гц, очних яблук - 60 Гц, центральної нервової системи - 250 Гц. Дія зовнішніх коливань з кратними частотами може викликати резонансні явища та призвести до зміщення та механічних ушкоджень внутрішніх органів. Частота власних коливань людей, що сидять, складає 4...8 Гц.

Вібрація сприймається людиною як природне навантаження, що подібне на важку працю. Вібрації частотою понад 200 Гц перевантажують нервову систему людини, потребують підвищеної психічної напруги.

4.2 Безпека в надзвичайних ситуаціях

Порядок дій працівників підприємства з виготовлення телекомунікаційних систем у разі виникнення надзвичайних ситуацій [2]

Хімічна небезпека.

Аварії (катастрофи) на підприємствах, транспорті та продуктопроводах можуть супроводжуватися викидом (виливом) в атмосферу і на прилеглу територію сильнодіючих отруйних речовин (СДОР), таких як хлор, аміак, синильна кислота, фосген, сірчаний ангідрид та інші. Це являє серйозну

небезпеку для населення, заражене повітря уражає органи дихання, а також очі, шкіру та інші органи.

Фактори небезпеки викиду (розливу) хімічно небезпечних речовин: забруднення навколишнього середовища, небезпека для всього живого, що опинилося на забрудненій місцевості (загибель людей, тварин, знищення посівів та ін.), крім того, внаслідок можливого хімічного вибуху виникнення сильних руйнувань на значній території. [1]

Дії у випадку загрози виникнення хімічної небезпеки:

1. Сирени і переривчасті гудки підприємств - це сигнал "Увага всім". негайно ввімкніть приймач радіотрансляційної мережі або телевізор. Слухайте інформацію про надзвичайну ситуацію та порядок дій.

2. При оголошенні небезпечного стану уникайте паніки.

3. Попередьте сусідів, надайте допомогу інвалідам, дітям та людям похилого віку.

4. Виконайте заходи щодо зменшення проникнення отруйних речовин в квартиру (будинки): щільно закрийте вікна та двері, щілини заклейте.

5. Підготуйте запас питної води: наберіть воду у герметичні ємності, підготуйте найпростіші засоби санітарної обробки (мильний розчин для обробки рук).

6. Дізнайтеся у місцевих органів влади про місце збору мешканців для евакуації та уточніть час її початку. Підготуйтеся: упакуйте у герметичні пакети та складіть у валізу документи, цінності та гроші, предмети першої необхідності, ліки, мінімум білизни та одягу, запас консервованих продуктів на 2-3 доби.

7. Перед виходом з будинку вимкніть джерела електро-, водо- і газопостачання, візьміть підготовлені речі, одягніть засоби захисту.

Дії у випадку раптового виникнення хімічної небезпеки:

1. Уникайте паніки. З одержанням повідомлення (по радіо або іншим засобам оповіщення) про викид (розлив) в атмосферу СДОР та про небезпеку хімічного зараження, виконайте передбачені заходи.

2. Надягніть засоби індивідуального захисту органів дихання та найпростіші засоби захисту шкіри.

3. По можливості негайно залишіть зону хімічного забруднення.

4. Якщо засобів індивідуального захисту немає і вийти із району аварії неможливо, залишайтеся у приміщенні і негайно та надійно герметизуйте приміщення! Зменшить можливість проникнення СДОР (парів, аерозолів) у приміщення: щільно закрийте вікна та двері, димоходи, вентиляційні люки, щілини в рамах вікон та дверей заклейте, вимкніть джерела газо-, електропостачання та загасіть вогонь у печах. Чекайте повідомлень органів влади з питань надзвичайних ситуацій через засоби зв'язку.

5. Знайте, що вражаюча дія конкретної СДОР на людину залежить від її концентрації у повітрі та тривалості, тому якщо немає можливості покинути небезпечну зону, не панікуйте і продовжуйте вживати заходи безпеки.

6. Швидко зберіть необхідні документи, цінності, ліки, продукти, запас питної води та інші необхідні речі у герметичну валізу та підготуйтеся до евакуації.

7. Попередьте сусідів про початок евакуації. Надайте допомогу дітям, інвалідам та людям похилого віку. Вони підлягають евакуації в першу чергу.

8. Залишаючи приміщення (квартиру, будинок) вимкніть джерела електро-, водо- і газопостачання, візьміть підготовлені речі, одягніть засоби захисту.

9. Виходьте із зони хімічного зараження в бік, перпендикулярний напрямку вітру та обходьте тунелі, яри, лощини - в низинах може бути висока концентрація СДОР.

10. При підозрі на ураження СДОР уникайте будь-яких фізичних навантажень, необхідно пити велику кількість рідини (чай, молоко, сік, вода) та звернутися до медичного закладу.

11. Вийшовши із зони зараження, зніміть верхній одяг, ретельно вимийте очі, ніс та рот, по можливості прийміть душ.

12. З прибуттям на нове місця перебування, дізнайтеся у місцевих органів державної влади та місцевого самоврядування адреси організацій, що відповідають за надання допомоги потерпілому населенню.

Надання першої допомоги при ураженні СДОР

В першу чергу негайно захистіть органи дихання від подальшої дії СДОР. Надягніть на потерпілого протигаз або ватно-марлеву пов'язку, попередньо змочивши її водою або 2% розчином питної соди у випадку отруєння хлором, а у разі отруєння аміаком - водою або 5% розчином лимонної кислоти. Винесіть потерпілого із зони зараження та забезпечте йому спокій і тепло.

Радіаційна небезпека [1]

Факторами небезпеки радіації є: забруднення навколишнього середовища, небезпека для всього живого, що опинилося на забрудненій місцевості (загибель людей, тварин, знищення посівів та ін.), крім того, внаслідок можливого атомного вибуху виникнення сильних руйнувань на значній території.

Дії у випадку загрози виникнення радіаційної небезпеки:

1. При оголошенні небезпечного стану не панікуйте, слухайте повідомлення.

2. Попередьте сусідів, надайте допомогу інвалідам, дітям та людям похилого віку.

3. Дізнайтеся про час та місце збору мешканців для евакуації.

4. Уменшіть проникнення радіаційних речовин в квартиру (будинок): щільно закрийте вікна та двері, щілини заклейте.

5. Підготуйтеся до можливої евакуації: упакуйте у герметичні пакети та складіть у валізу документи, цінності та гроші, предмети першої необхідності, ліки, мінімум білизни та одягу, запас консервованих продуктів на 2-3 доби, питну воду. Підготуйте найпростіші засоби санітарної обробки (мильний розчин).

6. Перед виходом з приміщення від'єднайте всі споживачі електричного струму від електромережі, вимкніть газ та воду.

Дії у випадку раптового виникнення радіаційної небезпеки:

1. З одержанням повідомлення про радіаційну небезпеку негайно укрийтеся в будинку. Стіни дерев'яного будинку послаблюють іонізуюче випромінювання в 2 рази, цегляного - у 10 разів; заглиблені укриття (підвали): з покриттям із дерева у 7 разів, з покриттям із цегли або бетону - у 40-100 разів.

2. Уникайте паніки. Слухайте повідомлення органів влади з питань надзвичайних ситуацій.

3. Уменшіть можливість проникнення радіаційних речовин в приміщення.

4. Проведіть йодну профілактику. Йодистий калій вживати після їжі разом з чаєм, соком або водою 1 раз на день протягом 7 діб: дітям до двох років - по 0,040 г на один прийом; дітям від двох років та дорослим - по 0,125 г на один прийом. Водно-спиртовий розчин йоду приймати після їжі 3 рази на день протягом 7 діб: дітям до двох років - по 1-2 краплі 5% настоянки на 100мл молока або годувальної суміші; дітям від двох років та дорослим - по 3-5 крапель на стакан молока або води. Наносити на поверхню кінцівок рук настоянку йоду у вигляді сітки 1 раз на день протягом 7 діб.

5. Уточніть місце початку евакуації. Попередьте сусідів, допоможіть дітям, інвалідам та людям похилого віку. Вони підлягають евакуації в першу чергу.

6. Швидко зберіть необхідні документи, цінності, ліки, продукти, запас питної води, найпростіші засоби санітарної обробки та інші необхідні вам речі у герметичну валізу.

7. По можливості негайно залишіть зону радіоактивного забруднення.

8. Перед виходом з будинку вимкніть джерела електро-, водо- і газопостачання, візьміть підготовлені речі, одягніть протигаз (респіратор, ватно-марлеву пов'язку), верхній одяг (плащ, пальто, накидка), гумові чоботи.

9. З прибуттям на нове місце перебування, проведіть дезактивацію засобів захисту, одягу, взуття та санітарну обробку шкіри на спеціально обладнаному пункті або ж самостійно (зняти верхній одяг, ставши спиною проти вітру, витрясти його; повісити одяг на перекладину, віником або щіткою змести з нього радіоактивний пил та вимити водою; обробити відкриті ділянки шкіри водою. Для обробки шкіри можна використовувати марлю чи просто рушники.

10. Дізнайтеся у місцевих органів державної влади адреси організацій, що відповідають за надання допомоги потерпілому населенню.

Повінь, паводок

Фактори небезпеки повеней та паводків: руйнування будинків та будівель, мостів; розмив залізничних та автомобільних шляхів; аварії на інженерних мережах; знищення посівів; жертви серед населення та загибель тварин.

Внаслідок повені, паводку починається просідання будинків та землі, виникають зсуви та обвали.

Дії у випадку загрози виникнення повені, паводка:

1. Уважно слухайте інформацію про надзвичайну ситуацію та інструкції про порядок дій, не користуйтеся без потреби телефоном, щоб він був вільним для зв'язку з вами.

2. Зберігайте спокій, попередьте сусідів, надайте допомогу інвалідам, дітям та людям похилого віку.

3. Дізнайтеся у місцевих органах державної влади та місцевого самоврядування місце збору мешканців для евакуації та готуйтеся до неї.

4. Підготуйте документи, одяг, найбільш необхідні речі, запас продуктів харчування на декілька днів, медикаменти. Складіть все у валізу. Документи зберігайте у водонепроникному пакеті.

5. Від'єднайте всі споживачі електричного струму від електромережі, вимкніть газ.

6. Перенесіть більш цінні речі та продовольство на верхні поверхи або підніміть на верхні полиці.

7. Переженіть худобу на підвищену місцевість.

Пожежа

Щорічно внаслідок пожеж гине та страждає значна кількість людей. В дим та попіл перетворюються цінності на мільярди гривень.

Якщо виникла пожежа - рахунок часу йде на секунди. Не панікуйте та остерігайтеся: високої температури, задимленості та загазованості, обвалу конструкцій будинків і споруд, вибухів технологічного обладнання і приладів, падіння обгорілих дерев і провалів. Знайте, де знаходяться засоби пожежогасіння, та вмійте ними користуватися.

Заходи щодо рятування потерпілих з будинків, які горять, та під час гасіння пожежі:

1. Перед тим, як увійти в приміщення, що горить, накрийтеся мокрою ковдрою, будь-яким одягом чи щільною тканиною.

2. Відкривайте обережно двері в задимлене приміщення, щоб уникнути посилення пожежі від великого притоку свіжого повітря.

3. В сильно задимленому приміщенні рухайтесь поповзом або пригинаючись.

4. Для захисту від чадного газу необхідно дихати через зволожену тканину.

5. У першу чергу рятуйте дітей, інвалідів та старих людей.

6. Пам'ятайте, що маленькі діти від страху часто ховаються під ліжку, в шафу та забиваються у куток.

7. Виходити із осередку пожежі необхідно в той бік, звідки віє вітер.

8. Побачивши людину, на якій горить одяг, зваліть її на землю та швидко накиньте будь-яку ковдру чи покривало (бажано зволожену) і щільно притисніть до тіла, при необхідності, викличте медичну допомогу.

9. Якщо загорівся ваш одяг, падайте на землю і перевертайтеся, щоб збити полум'я, ні в якому разі не біжіть - це ще більше роздуває вогонь.

10. Під час гасіння пожежі використовуйте вогнегасники, пожежні гідранти, воду, пісок, землю, кошму та інші засоби гасіння вогню.

11. Бензин, гас, органічні масла та розчинники, що загорілися, гасіть тільки за допомогою пристосованих видів вогнегасників, засипайте піском або ґрунтом, а якщо осередок пожежі невеликий, накрийте його азбестовим чи брезентовим покривалом, зволоженою тканиною чи одягом.

12. Якщо горить електричне обладнання або проводка, вимкніть рубильник або електричні пробки, а потім починайте гасити вогонь. [1]

4.3 Висновки до розділу

В розділі «Охорона праці та безпека в надзвичайних ситуаціях» описано екстремальні умови, пов'язані з впливом шуму та порядок дій у разі виникнення надзвичайних ситуацій

ВИСНОВКИ

Проведено аналіз проблеми голосового керування промисловим обладнанням. Показано, що традиційні методи керування мають недоліки та обмеження, які можуть бути усунуті завдяки використанню опосередкованих людино-машинних інтерфейсів, зокрема інтерфейсу голосового керування. Розглянуто етапи функціонування голосового інтерфейсу та встановлено, що основними проблемами голосового керування є перетворення мовлення в текст та розпізнавання команд, озвучування повідомлень, інтерпретація наміру та параметрів команди.

Розглянуто задачу розпізнавання голосових команд, що включає підзадачі виявлення мовлення, сегментації фрази (де команда починається і закінчується), розпізнавання (класифікація команд), інтерпретація, контроль помилок.

В дослідженні проведено розробку методу розпізнавання голосових команд, в якому запропоновано шляхи вирішення перших трьох підзадач. Розроблено та апробовано метод обробки голосових сигналів, призначений для розпізнавання окремих слів-команд і подальшого формування керувальних впливів.

Суть методу така: користувач вимовляє команду в мікрофон, після чого система виконує обробку сигналу та виділяє інформативні ознаки, за якими здійснюється розпізнавання. Основними етапами методу є: запис мовного сигналу та підготовка даних для аналізу; адаптивна фільтрація з метою зменшення впливу шумів; визначення часових меж (тривалості) команд методом ковзного вікна; фонемна сегментація команд на основі аналізу частоти основного тону. Частоту основного тону оцінювали методом формантного (спектрального) аналізу, а її наявність у кожному ковзному вікні фіксували через порогову функцію.

Експериментальні результати для розглянутих голосових команд підтвердили, що послідовність інтервалів із наявною/відсутю частоти основного тону, а також загальна тривалість команди відрізняються для різних слів. Отже, ці характеристики запропоновано використати як критерії для ідентифікації команди та подальшого формування відповідного сигналу керування.

ПЕРЕЛІК ПОСИЛАНЬ

1. Liu, C., Zheng, P., Xun, X., 2023. 'Digitalisation and servitisation of machine tools in the era of Industry 4.0: a review'. In: International Journal of Production Research, 61, 4069-4101.
2. Chao L., Hrishikesh V., Yuqian L., Xun X., 2019. 'Cyber-Physical Machine Tools Platform using OPC UA and MTConnect'. In: Journal of Manufacturing Systems 51, 61–74.
3. Reshamwala, A., Mishra, D., Pawar, P., 2013. 'REVIEW ON NATURAL LANGUAGE PROCESSING'. In: IRACST – Engineering Science and Technology: An International Journal (ESTIJ) 3, 113–116
4. Nadkarni, P. M., Ohno-Machado, L., Chapman, W., 2011. Natural language processing: an introduction. In: Journal of the American Medical Informatics Association, 18, 544–551.
5. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, Aidan N., Kaiser, Ł., Polosukhin, I., 2017. 'Attention is All you Need'. In: Advances in Neural Information Processing Systems 3.
6. Devlin, J., Chang, M., Lee, K., Toutanova, K., 2018. 'BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding'. In: ArXiv
7. Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Amodei, D., 2020. 'Language Models are Few-Shot Learners'. In: ArXiv
8. Krupitzer, C., Muëller, S., Lesch, V., M., Edinger, J., Lemken, A., Schaëfer, D., Kounev, S., Becker, C., 2020. 'A Survey on Human Machine Interaction in Industry 4.0.' In: ArXiv

9. Mertes J., Lindenschmitt D., Amirrezai M., Tashakor N.,...,2022. 'Evaluation of 5G-capable framework for highly mobile, scalable human-machine interfaces in cyber-physical production systems'. In: Journal of Manufacturing Systems. 64
10. Benzeghiba M., De Mori R., Deroo O., Dupont S., Erbes T., Jouvet D., Fissore L., Laface P., Mertins A., Ris C., Rose R., Tyagi V., Wellekens C. 'Automatic speech recognition and speech variability: A review'. In: Speech Communication. 49, 763–786
11. Valle, R., Shih, K., Prenger, R., Catanzaro, B., 2020. Flowtron: An Autoregressive Flow-based Generative Network for Text-to-Speech Synthesis. In: ArXiv.
12. Gayathri, N., Sundar M., Sargurunathan R., Sudharsan R., Sajith A., 2022. Design of Voice Controlled Multifunctional Computer Numerical Control (CNC) Machine, In: International Conference on Inventive Computation Technologies (ICICT), 657–663.
13. Fedosov, Y., Katridi, A., 2018. Concept of Implementing Computer Voice Control for CNC machines Using Natural Language Processing. In: Manufacturing Letters, 15, 89–92.
14. M. Ya. Afanasev et al., "A concept for integration of voice assistant and modular cyber-physical production system," in Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN), Helsinki, Finland, vol. 1, Jul. 2019, pp. 27-32.
15. V Priyadarshana et al., "Voice controlled robot manipulator for industrial applications," in Proc. IEEE 13th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON), Vancouver, BC, Canada Oct. 2022, pp. 0160-0165.
16. S. S. Yadav, P. Kumar, S. Kumar, and S. Singh, "Google assistant controlled home automation with voice recognition," in Proc. 3rd Int. Conf. Adv. Comput., Commun. Control Netw. (ICACN), Noida, India, Dec. 2021, pp. 1039-1044.

17. F. Longo and A. Padovano, "Voice-enabled assistants of the operator 4.0 in the social smart factory: Prospective role and challenges for an advanced human-machine interaction," *Manuf. Lett.*, vol. 26, pp. 12-16, Oct. 2020.
18. M. Ya. Afanasev et al., "A concept for integration of voice assistant and modular cyber-physical production system," in *Proc. IEEE 17th Int. Conf. Ind. Informat. (INDIN)*, Helsinki, Finland, vol. 1, Jul. 2019, pp. 27-32.
19. D. Hirschfeld and K. Rottmayr, *Voice Control—Ready for Automation: Warum Sprachinteraktion in Zukunft Grafische Bedienkonzepte Erweitern Oder Sogar ablösen Wird*, voice INTER connect, Dresden, Germany, Spectra GmbH & Co. KG, Reutlingen, Germany, Aug. 2020. [Online].
20. J. F. Kelley, "An empirical methodology for writing user-friendly natural language computer applications," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, Boston, MA, USA, 1983, pp. 193-196.
21. C. Spille, B. Kollmeier, and B. T. Meyer, "Comparing human and automatic speech recognition in simple and complex acoustic scenes," *Comput. Speech Lang.*, vol. 52, pp. 123-140, Nov. 2018, doi: 10.1016/j.csl.2018.04.003.
22. J. Lazar, J. H. Feng, and H. Hochheiser, *Research Methods in Human-Computer Interaction*. Cambridge, MA, USA: Elsevier Morgan Kaufmann Publishers, 2017, pp. 33-67.
23. V. Peddinti, D. Povey, and S. Khudanpur, "A time delay neural network architecture for efficient modeling of long temporal contexts," in *Proc. Interspeech*, Sep. 2015, pp. 3214-3218.
24. S. Ganapathy and V. Peddinti, "3-D CNN models for far-field multi-channel speech recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 5499-5503.
25. D. Povey et al., "Purely sequence-trained neural networks for ASR based on lattice-free MMI," in *Proc. Interspeech*, Sep. 2016, pp. 2751-2755.

26. Dozorskyi V., Dediv I., Sverstiuk S., Nykytyuk V., Karnaukhov A. The Method of Commands Identification to Voice Control of the Electric Wheelchair. Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2023. P.233-240. ISSN 1613-0073.

27. Khvostivska L., Khvostivskyi M., Dediv I., Yatskiv V., Palaniza Y. Method, Algorithm and Computer Tool for Synphase Detection of Radio Signals in Telecommunication Networks with Noises. Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2023. P.173-180. ISSN 1613-0073.

28. Khvostivska L., Khvostivskyi M., Dunets V., Dediv I. Mathematical, algorithmic and software support of synphase detection of radio signals in electronic communication networks with noises. Scientific Journal of TNTU (Tern.), vol 111, no 3, 2023. pp. 48–57.

29. Основи технології радіоелектронних апаратів : навчальний посібник / Р. А. Ткачук, В. Г. Дозорський, Л. Є. Дедів, І. Ю. Дедів. - Тернопіль : Тернопільський національний технічний університет імені Івана Пулюя, 2017. - 336 с.

30. Dozorskyi V., Dediv L., Kovalyk S., Dozorska O., Dediv I. (2024) Design of the endoskeleton of a biocontrolled hand prosthesis. Scientific Journal of TNTU (Tern.), vol. 115, no 3, pp. 100-111.

31. Дедів І. Ю. Метод автентифікації користувачів за параметрами голосових сигналів / І. Ю. Дедів, М. М. Кузик // Збірник тез доповідей VI Міжнародної науково-технічної конференції молодих учених та студентів „Актуальні задачі сучасних технологій“, 16-17 листопада 2017 року. Т.: ТНТУ, 2017. Том 2. С. 47.

32. Методичні рекомендації до виконання, оформлення та захисту кваліфікаційних робіт для здобувачів другого (магістерського) рівня вищої освіти за спеціальністю 172 «Електронні комунікації та радіотехніка» галузі знань 17 «Електроніка, автоматизація та електронні комунікації»/уклад.: Дунець В.Л., Хвостівська Л.В., Дедів І.Ю. Тернопіль: ТНТУ, 2024. 56 с.

33. Математичне та комп'ютерне моделювання електрокардіосигналів у системах голтерівського моніторингу / Л.Є. Дедів, А.С. Сверстюк, І.Ю. Дедів, М.О. Хвостівський, В.Г. Дозорський, Є.Б. Яворська. – Львів: Видавництво «Магнолія - 2006», 2021. – 120 с. ISBN 978-617-574-218-1.

34. Математичне моделювання, методи та програмне забезпечення опрацювання дихальних шумів у комп'ютерних аускультативних діагностичних системах / І.Ю. Дедів, А.С. Сверстюк, Л.Є. Дедів, В.Г. Дозорський, М.О. Хвостівський. – Львів: Видавництво «Магнолія - 2006», 2021. – 126 с. ISBN 978-617-574-219-8.

35. Методичні рекомендації з оформлення кваліфікаційних робіт бакалавра за спеціальністю 172 «Телекомунікації та радіотехніка» уклад.: Дунець В.Л., Хвостівський М.О. Дедів І.Ю. Тернопіль: ТНТУ імені Івана Пулюя, 2021 р. – 72с.

36. Дозорський В.Г., Дозорська О.Ф., Дедів Л.Є., Дедів І.Ю., Паньків І.М., Яворська Є.Б. Структура системи відбору біосигналів для задачі відновлення комунікативної функції людини. Вісник Хмельницького національного університету: технічні науки. – Хмельницький: редакція журналу "Вісник Хмельницького національного університету". – 2019. - №2(271) – с. 183-186.

37. Dozorska O., Yavorska E., Dozorskyi V., Pankiv I., Dediv L. Dediv I. The Method of Indirect Restoration of Human Communicative Function. Proc. of the 15th International Conference on the Experience of Designing and Application of CAD Systems (CADSM), CADSM'2019, (pp. 19–22). Polyana-Svalyava (Zakarpattya), UKRAINE 978-1-7281-0053-1/19.

38. Дозорська О.Ф., Яворська Є.Б., Дозорський В.Г., Дедів Л.Є., Дедів І.Ю. Метод виявлення ознак основного тону в структурі електроміографічних сигналів для задачі компенсації порушеної комунікативної функції людини», Вісник НТУУ "КПІ". Серія Радіотехніка, Радіоапаратобудування, (81), с. 56-64. doi: 10.20535/RADAP.2020.81.56-64.

39. Khvostivska L., Khvostivsky M., Dunetc V., Dediv I. Mathematical and Algorithmic Support of Detection Useful Radiosignals in Telecommunication Networks. 2nd International Workshop on Information Technologies: Theoretical and Applied Problems, ITTAP 2022. CEUR Workshop Proceedings. Ternopil 22- 24 November 2022. Vol 3309, P. 314-318. ISSN 1613-0073.

40. Гевко О.В., Дозорський В.Г., Дедів Л.Є., Дедів І.Ю., Дозорська О.Ф. Структурний синтез вібротерапевтичної апаратури. Перспективні технології та прилади, № 20, Луцьк, 2022. – с. 23-31.

41. Dozorskyi V., Dediv I., Sverstiuk S., Nykytyuk V., Karnaukhov A. The Method of Commands Identification to Voice Control of the Electric Wheelchair. Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2023. P.233-240. ISSN 1613-0073.

42. Khvostivska L., Khvostivskyi M., Dediv I., Yatskiv V., Palaniza Y. Method, Algorithm and Computer Tool for Synphase Detection of Radio Signals in Telecommunication Networks with Noises. Proceedings of the 1st International Workshop on Computer Information Technologies in Industry 4.0 (CITI 2023). CEUR Workshop Proceedings. Ternopil, Ukraine, June 14-16, 2023. P.173-180. ISSN 1613-0073.

43. Khvostivska L., Khvostivskyi M., Dunets V., Dediv I. Mathematical, algorithmic and software support of synphase detection of radio signals in electronic communication networks with noises. Scientific Journal of TNTU (Tern.), vol 111, no 3, 2023. pp. 48–57.

44. Основи технології радіоелектронних апаратів : навчальний посібник / Р. А. Ткачук, В. Г. Дозорський, Л. Є. Дедів, І. Ю. Дедів. - Тернопіль : Тернопільський національний технічний університет імені Івана Пулюя, 2017. - 336 с.
45. Dozorskyi V., Dediv L., Kovalyk S., Dozorska O., Dediv I. (2024) Design of the endoskeleton of a biocontrolled hand prosthesis. Scientific Journal of TNTU (Tern.), vol. 115, no 3, pp. 100-111.
46. Техноекологія та цивільна безпека. Частина «Цивільна безпека». Навчальний посібник / В.С. Стручок, – Тернопіль: ТНТУ ім. І.Пулюя, 2022. – 150 с.
47. Стручок В.С. Безпека в надзвичайних ситуаціях. Методичний посібник для здобувачів освітнього ступеня «магістр» всіх спеціальностей денної бо та заочної (дистанційної) форм навчання / В.С.Стручок. — Тернопіль: ФОП Паляниця В. А., 2022. — 156 с.

ДОДАТКИ

Додатор А

Текст програми обробки голосових сигналів

```

function AudioAnalyzerGUI
    if exist('OCTAVE_VERSION','builtin')
        try, pkg load signal; catch, end
    end
    S = struct();
    S.data = [];
    S.fs = 44100;
    % ----- Вікно -----
    S.f = figure('Name','Audio Analyzer GUI (legacy)', ...
        'NumberTitle','off','MenuBar','none','ToolBar','none', ...
        'Position',[50 50 1250 760]);
    % ----- Оси (5 графіків) -----
    S.ax1 = axes('Parent',S.f,'Units','normalized','Position',[0.06 0.55 0.28 0.35]); grid(S.ax1,'on');
    title(S.ax1,'xrr (сирій сигнал)');
    S.ax2 = axes('Parent',S.f,'Units','normalized','Position',[0.37 0.55 0.28 0.35]); grid(S.ax2,'on');
    title(S.ax2,'p (детекція 1)');
    S.ax3 = axes('Parent',S.f,'Units','normalized','Position',[0.68 0.55 0.28 0.35]); grid(S.ax3,'on');
    title(S.ax3,'xr (відрізок *2)');
    S.ax4 = axes('Parent',S.f,'Units','normalized','Position',[0.06 0.10 0.28 0.35]); grid(S.ax4,'on');
    title(S.ax4,'b (детекція 2)');
    S.ax5 = axes('Parent',S.f,'Units','normalized','Position',[0.37 0.10 0.28 0.35]); grid(S.ax5,'on');
    title(S.ax5,'r (метрика hh)');
    % ----- Панель керування -----
    S.panel = uipanel('Parent',S.f,'Title','Керування','Units','normalized','Position',[0.68 0.10 0.28
0.35]);
    y = 0.88; dy = 0.10; xL = 0.05; xE = 0.55; wL = 0.48; wE = 0.40; h = 0.08;
    S.btnLoad = uicontrol('Parent',S.panel,'Style','pushbutton','String','Завантажити (.wav/.mat)', ...
        'Units','normalized','Position',[0.05 0.92 0.90 0.07], ...
        'Callback',@onLoad);
    S.btnRun = uicontrol('Parent',S.panel,'Style','pushbutton','String','Запустити аналіз', ...
        'Units','normalized','Position',[0.05 0.84 0.90 0.07], ...
        'Callback',@onRun);
    S.txtFile = uicontrol('Parent',S.panel,'Style','text','String','Файл: (не завантажено)', ...
        'Units','normalized','HorizontalAlignment','left', ...
        'Position',[0.05 0.76 0.90 0.06]);
    % Поля параметрів
    y = 0.68;
    [S.edFs] = addEdit('fs (Гц)', '44100', y); y=y-dy;
    [S.edTa] = addEdit('ta', '220', y); y=y-dy;
    [S.edW1] = addEdit('вікно1', '250', y); y=y-dy;
    [S.edThr1] = addEdit('попір1', '0.006', y); y=y-dy;
    [S.edTau] = addEdit('tau', '221', y); y=y-dy;

```

```

[S.edW2] = addEdit('вікно2', '500', y); y=y-dy;
[S.edStart]= addEdit('start samp', '300000',y); y=y-dy;
[S.edEnd] = addEdit('end samp', '360000',y); y=y-dy;
[S.edFLow] = addEdit('f low (Гц)', '200', y); y=y-dy;
[S.edFHigh]= addEdit('f high (Гц)', '400', y); y=y-dy;
[S.edThr2] = addEdit('попір2', '0.1', y); y=y-dy;
S.txtStatus = uicontrol('Parent',S.panel,'Style','text','String','Статус: готово', ...
    'Units','normalized','HorizontalAlignment','left', ...
    'Position',[0.05 0.02 0.90 0.06]);
guidata(S.f, S);
% ----- вкладки функції -----
function [hEdit] = addEdit(lbl, def, yy)
    uicontrol('Parent',S.panel,'Style','text','String',lbl, ...
        'Units','normalized','HorizontalAlignment','left', ...
        'Position',[xL yy wL h]);
    hEdit = uicontrol('Parent',S.panel,'Style','edit','String',def, ...
        'Units','normalized','Position',[xE yy wE h]);
end
function setStatus(msg)
    S = guidata(S.f);
    set(S.txtStatus,'String',['Статус: ' msg]);
    drawnow;
end
function onLoad(~,~)
    S = guidata(S.f);
    [file, path] = uigetfile({'*.wav;*.mat','WAV або MAT (*.wav, *.mat)'} , 'Оберіть файл');
    if isequal(file,0), return; end
    full = fullfile(path,file);
    try
        [~,~,ext] = fileparts(full);
        if strcmpi(ext,'.wav')
            if exist('audioread','file')
                [y, fs] = audioread(full);
            else
                % дуже старий MATLAB/Octave
                [y, fs] = wavread(full); %#ok<DWVRD>
            end
            if size(y,2) > 1, y = mean(y,2); end
            S.data = y(:);
            S.fs = fs;
            set(S.edFs,'String',num2str(fs));
        else
            T = load(full);
            if isfield(T,'data') && isnumeric(T.data)
                S.data = T.data(:);
            else
                fn = fieldnames(T);
                found = false;
                for i=1:numel(fn)

```

```

        v = T.(fn{i});
        if isnumeric(v) && ~isempty(v)
            S.data = v(:);
            found = true;
            break;
        end
    end
    if ~found
        errordlg('У .mat не знайдено numeric даних (ні "data", ні іншого numeric
поля).','Помилка');
        return;
    end
end
end
set(S.txtFile,'String',['Файл: ' file]);
guidata(S.f, S);
axes(S.ax1); cla(S.ax1); plot(S.ax1, S.data); grid(S.ax1,'on');
title(S.ax1,'xrr (сирий сигнал)');
setStatus(sprintf('завантажено %d семплів', numel(S.data)));
catch ME
    errordlg(ME.message,'Помилка завантаження');
end
end
function onRun(~,~)
    S = guidata(S.f);
    if isempty(S.data)
        errordlg('(.wav або .mat).','Немає даних');
        return;
    end
    try
        setStatus('аналіз...');
        fs = str2double(get(S.edFs,'String'));
        ta = round(str2double(get(S.edTa,'String')));
        w1 = round(str2double(get(S.edW1,'String')));
        thr1 = str2double(get(S.edThr1,'String'));
        tau = round(str2double(get(S.edTau,'String')));
        w2 = round(str2double(get(S.edW2,'String')));
        thr2 = str2double(get(S.edThr2,'String'));
        sStart = round(str2double(get(S.edStart,'String')));
        sEnd = round(str2double(get(S.edEnd,'String')));
        fLow = str2double(get(S.edFLow,'String'));
        fHigh = str2double(get(S.edFHigh,'String'));
        xrr = S.data(:);
        % ----- p -----
        nFrames1 = floor((numel(xrr) - w1) / ta);
        if nFrames1 < 1
            errordlg('Сигнал занадто короткий для частини 1','Помилка');
            return;
        end
    end
end

```

```

p = zeros(1,nFrames1);
for kk=0:nFrames1-1
    idx1 = kk*ta + 1;
    idx2 = idx1 + w1 - 1;
    x112 = abs(xrr(idx1:idx2));
    p(kk+1) = (max(x112) >= thr1);
end
% ----- відрізок -----
sStart = max(1, min(sStart, numel(xrr)));
sEnd = max(1, min(sEnd, numel(xrr)));
if sEnd < sStart, tmp=sStart; sStart=sEnd; sEnd=tmp; end
x1 = xrr(sStart:sEnd);
xr = x1 .* 2;
% ----- b, r -----
nFrames2 = floor((numel(xr) - w2) / tau);
if nFrames2 < 1
    error('Відрізок занадто короткий для частини 2','Помилка');
    return;
end
b = zeros(1,nFrames2);
r = zeros(1,nFrames2);
for k=0:nFrames2-1
    i1 = k*tau + 1;
    i2 = i1 + w2 - 1;
    x12 = xr(i1:i2);
    z1 = xcorr(x12);
    N = length(z1);
    y = abs(fft(z1, N))/N*2;
    df = fs / N;
    n1 = max(1, fix(fLow/df));
    n2 = min(N, fix(fHigh/df));
    if n2 <= n1
        hh = mean(y(n1:min(N,n1+1)));
    else
        hh = mean(y(n1:n2));
    end
    r(k+1) = hh;
    b(k+1) = (hh >= thr2);
end
% ----- ПЛОТИ -----
axes(S.ax1); cla(S.ax1); plot(S.ax1, xrr); grid(S.ax1,'on'); title(S.ax1,'xrr (сирий сигнал)');
axes(S.ax2); cla(S.ax2); plot(S.ax2, p); grid(S.ax2,'on');
title(S.ax2, sprintf('p (thr1=%0.4g, ta=%0d, w1=%0d)',thr1,ta,w1));
ylim(S.ax2,[-0.2 1.2]);
t11 = (0:numel(xr)-1) ./ fs;
axes(S.ax3); cla(S.ax3);
% 1) xrr
plot(S.ax3, xrr, 'b'); grid(S.ax3,'on'); hold(S.ax3,'on');
% 2) p, але розтягнути по x (по кадрах) і масштабувати по амплітуді

```

```

A = max(abs(xrr));
if A == 0, A = 1; end
idx_p = (0:numel(p)-1)*ta + 1;      % позиції кадрів у семплах
p_vis = (p*2 - 1) * (0.8*A);        % 0->-0.8A, 1->+0.8A
stairs(S.ax3, idx_p, p_vis, 'r', 'LineWidth', 1.5);
title(S.ax3, 'xrr + p (накладено)');
xlim(S.ax3, [1 numel(xrr)]);
hold(S.ax3, 'off');
    axes(S.ax4); cla(S.ax4); plot(S.ax4, b); grid(S.ax4, 'on');
    title(S.ax4, sprintf('b (thr2=%.4g, tau=%d, w2=%d)', thr2, tau, w2));
    ylim(S.ax4, [-0.2 1.2]);
    axes(S.ax5); cla(S.ax5); plot(S.ax5, r); grid(S.ax5, 'on');
    title(S.ax5, sprintf('r (середнє %.0f..%.0f Гц)', fLow, fHigh));
    setStatus(sprintf('готово. p=%d кадрів, b=%d кадрів', numel(p), numel(b)));
catch ME
    errordlg(ME.message, 'Помилка аналізу');
    setStatus('помилка');
end
end
end
end

```

Додаток Б

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ**

МАТЕРІАЛИ

ХІІ НАУКОВО-ТЕХНІЧНОЇ КОНФЕРЕНЦІЇ

**«ІНФОРМАЦІЙНІ МОДЕЛІ,
СИСТЕМИ ТА ТЕХНОЛОГІЇ»**



17-18 грудня 2025 року

**ТЕРНОПІЛЬ
2025**

УДК 621.372.852.15

М. Шастків; В. Березовський; А. Федунь; І. Дедів, к. т. н., доц.

(Тернопільський національний технічний університет імені Івана Пулюя, Україна)

ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ МЕТОДУ АДАПТИВНОЇ ФІЛЬТРАЦІЇ АУДІОСИГНАЛІВ НА ОСНОВІ СПЕКТРАЛЬНОГО ВІДНІМАННЯ

UDC 621.372.852.15

M. Shastkiv; V. Berezovsky; A. Fedun; Iryna Dediv, PhD., Assoc. Prof.

INCREASING THE EFFICIENCY OF THE METHOD OF ADAPTIVE AUDIO SIGNALS FILTERING BASED ON SPECTRAL SUBTRACTION

Центральною задачею обробки сигналів є розроблення ефективних методів виділення їх на фоні завад. Зокрема багато уваги приділяється питанням шумоподавлення в різних сумішах сигналів та шумів, зокрема при адитивних шумах в результаті впливів різних факторів як навколишнього середовища, так і каналу зв'язку. В якості останнього часто розглядається радіоканал.

З цією метою була розвинута теорія оптимальної фільтрації, зокрема із застосуванням адаптивних фільтрів. Такі фільтри використовуються для опрацювання нестационарних сигналів та середовищ, або в застосунках, де потрібна адаптація процесу по кожній вибірці або низька затримка часу обробки. Застосування адаптивних фільтрів включають багатоканальне шумоподавлення, вирівнювання каналів для стільникових мобільних систем зв'язку, придушення відлуння та кодування мовлення з малою затримкою тощо. При цьому домінуючим є напрямок розроблення адаптивних алгоритмів шумоподавлення, які можуть застосовуватись до опрацювання як радіосигналів, так і прийнятих і демодульованих/декодованих аудіосигналів для радіозв'язку та телебачення. Адаптивне шумоподавлення є методом обробки сигналів для високоефективного подавлення адитивних перешкод або шумів, що спотворюють прийнятий корисний сигнал, коли відомі параметри та характеристики перешкод, і коли корисний сигнал та перешкоди не корельовані. До таких методів відносять зокрема методи Вінерівської фільтрації, Калманівської фільтрації, спектрального віднімання тощо.

Останній метод є перспективним в плані простоти технічної реалізації, малої часової затримки опрацювання тощо. Основною ж проблемою методу спектрального віднімання є наявність спотворень обробки, спричинених випадковими варіаціями шуму. Тому залишається актуальним завдання розроблення більш ефективного методу адаптивної фільтрації радіосигналів, зокрема аудіосигналів, на основі спектрального віднімання.

Власне для підвищення ефективності методу спектрального віднімання пропонується забезпечити можливість адаптивного оцінювання параметрів шуму. Пропонований алгоритм автоматично відстежує моменти, коли корисний сигнал відсутній, і оновлює параметри спектра шуму в режимі реального часу. Власне метод базується на адаптації до ділянок сигналу, на яких відсутній корисний сигнал. Проведено експериментальну верифікацію пропонованого методу спектрального віднімання із адаптивним визначенням ділянок шуму. Показано, що пропонований метод адаптивної фільтрації на основі спектрального віднімання є більш ефективним в порівнянні із класичним методом.

М. Шастків, В. Березовський, А. Федунь, І. Дедів ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ МЕТОДУ АДАПТИВНОЇ ФІЛЬТРАЦІЇ АУДІОСИГНАЛІВ НА ОСНОВІ СПЕКТРАЛЬНОГО ВІДНІМАННЯ M. Shastkiv, V. Berezovsky, A. Fedun, Iryna Dediv INCREASING THE EFFICIENCY OF THE METHOD OF ADAPTIVE AUDIO SIGNALS FILTERING BASED ON SPECTRAL SUBTRACTION	15
Г. Шимчук, Я. Литвиненко ПОРІВНЯЛЬНИЙ АНАЛІЗ ЦИКЛІЧНОГО ВИПАДКОВОГО ПРОЦЕСУ ТА ARIMA/SARIMA У МОДЕЛЯХ ГАЗОСПОЖИВАННЯ G. Shymchuk, I. Lytvynenko COMPARATIVE ANALYSIS OF CYCLIC RANDOM PROCESS AND ARIMA/SARIMA IN GAS CONSUMPTION MODELS	16
СЕКЦІЯ 2. ІНФОРМАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ, КІБЕРБЕЗПЕКА	
О. Бідюк, С. Марценко АНАЛІЗ ЗАСОБІВ ЗАХИСТУ ХМАРНИХ ІТ СЕРИДОВИЩ O. Bidiuk, S. Martsenko ANALYSIS OF CLOUD IT ENVIRONMENT PROTECTION TOOLS	18
Б. Білоус, М. Паламар АНАЛІЗ ВПЛИВУ ПАРАМЕТРІВ СТЕРЕОСИСТЕМ З ГЛИБИННОЮ КАРТОЮ НА ЯКІСТЬ РОЗПІЗНАВАННЯ ДИНАМІЧНИХ ОБ'ЄКТІВ У СИСТЕМАХ ТЕХНІЧНОГО ЗОРУ B. Bilous, M. Palamar ANALYSIS OF THE IMPACT OF STEREO VISION SYSTEM PARAMETERS WITH DEPTH MAPS ON DYNAMIC OBJECT RECOGNITION PERFORMANCE IN MACHINE VISION	19
Б. Білоус, М. Паламар ПОРІВНЯННЯ АПАРАТНИХ ПЛАТФОРМ ДЛЯ РОЗПІЗНАВАННЯ ДИНАМІЧНИХ ОБ'ЄКТІВ З ОБМЕЖЕННЯМИ ПО ЕНЕРГОСПОЖИВАННЮ ТА МАСІ B. Bilous, M. Palamar COMPARISON OF HARDWARE PLATFORMS FOR DYNAMIC OBJECT RECOGNITION UNDER POWER AND WEIGHT CONSTRAINTS	22
Д. Бойко УДОСКОНАЛЕННЯ ІНФОРМАЦІЙНОЇ СИСТЕМИ ІНТЕЛЕКТУАЛЬНОГО СОРТУВАННЯ ДАНИХ ДЛЯ ІНТЕГРАЦІЇ В «GOOGLE DRIVE API» НА ОСНОВІ КОНТЕНТУ D. Boiko IMPROVEMENT OF THE INFORMATION SYSTEM OF INTELLIGENT DATA SORTING FOR INTEGRATION INTO "GOOGLE DRIVE API" BASED ON CONTENT	24
В. Бонар, В. Готович МЕТОДИ МАШИННОГО НАВЧАННЯ ДЛЯ ЗАВЧАСНОГО ПЕРЕДБАЧЕННЯ НЕСПРАВНОСТЕЙ ТРАНСПОРТНОГО ЗАСОБУ НА ОСНОВІ OBD-2 ДАНИХ V. Bonar, V. Hotovych MACHINE LEARNING METHODS FOR EARLY PREDICTION OF VEHICLE MALFUNCTIONS BASED ON OBD-2 DATA	25
А. Боруха, А. Бадалян ВЕБОРІЄНТОВАНА ІНФОРМАЦІЙНА СИСТЕМА ЕЛЕКТРОННОЇ КОМЕРЦІЇ: АРХІТЕКТУРА, ФУНКЦІОНАЛЬНІСТЬ ТА ПЕРСПЕКТИВИ РОЗВИТКУ A. Borukha, A. Badalyan WEB-ORIENTED INFORMATION SYSTEM FOR E-COMMERCE: ARCHITECTURE,	27