

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Електронної та біомедичної інженерії
(повна назва)

Кафедра Мікроелектроніки, електронних приладів та пристроїв
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

«Розробка та аналіз параметрів фільтрувальних ланцюгів для обробки сигналів у цифрових системах»

(тема)

Виконав:

студент 4 курсу, групи ЕЕПС-21-1

Бібік Сергій Миколайович

(прізвище, ініціали)

Спеціальність 171 Електроніка

(код і повна назва спеціальності)

Тип програми освітньо-професійна

Освітня програма «Електронні пристрої та системи»

(повна назва освітньої програми)

Керівник Свідерська Л.І.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри МЕЕПП

(підпис)

Бондаренко І. М.

(прізвище, ініціали)

2025 р.

Харківський національний університет радіоелектроніки

Факультет Електронної та біомедичної інженеріїКафедра Мікроелектроніки, електронних приладів та пристроївРівень вищої освіти перший (бакалаврський)Спеціальність 171 Електроніка

(код і повна назва)

Тип програми освітньо-професійнаОсвітня програма «Електронні пристрої та системи»

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

«___» _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

Студентові Бібіку Сергію Миколайовичу

(прізвище, ім'я, по батькові)

1. Тема роботи «Розробка та аналіз параметрів фільтрувальних ланцюгів для обробки сигналів у цифрових системах»затверджена наказом університету від 26.05 2025 р. № 415 Ст.2. Термін подання студентом роботи до екзаменаційної комісії 6.05 2025 р.

3. Вихідні дані до роботи: _____

Аудіозаписи мовлення, зашумлені змінним стаціонарним шумом, що дозволило перевірити адаптивність розробленого алгоритму шумоподавлення. Обробка та аналіз цих даних здійснювалися за допомогою мови Python та її спеціалізованих бібліотек для обробки сигналів.

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Вибір способу вводу інформації _____

2. Вибір методу оцінки шуму та оновлення оцінки шуму _____

3. Реалізація детекції голосу _____

4. Вибір способу спектарального віднімання _____

5. Методи усунення проблем при спектральному відніманні

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) _____

Додаток А Відомість кваліфікаційної роботи бакалавра

Додаток Б Схема алгоритму шумоподавлення

Додаток В Код програми

Додаток Г Презентація

6. Консультанти розділів роботи (п.6 включається до завдання за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1.	Аналіз отриманого завдання на кваліфікаційну	1.04.2025р.	виконано
2.	Аналітичний огляд медичної та технічної літератури	15.04.2025р.	виконано
3.	Структурування матеріалу	27.04.2025р.	виконано
4.	Вибір програмних засобів	02.05.2025р.	виконано
5.	Опис обраного функціоналу, функціональна схема	05.05.2025р.	виконано
6.	Реалізація програмного забезпечення	18.05.2025р.	виконано
7.	Оформлення пояснювальної записки	22.05.2025р.	виконано
8.	Оформлення презентації доповіді	24.05.2025р.	виконано
9.	Рецензування, нормоконтроль	6.06.2025р.	виконано
10.	Дистанційний захист кваліфікаційної роботи	11.06.2025р.	виконано

Дата видачі завдання 19.04_2025 р.

Студент _____

(підпис)

Керівник роботи _____

(підпис)

доц. каф. МЕЕПШ Свідерська Л.І.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка містить: 51 с., 19 рис., 4 додатки, 10 джерел, 3 табл.

СПЕКТРАЛЬНЕ, ВІДНІМАННЯ, ШУМОЗАГЛУШЕННЯ, ЦИФРОВА ОБРОБКА СИГНАЛІВ, АМПЛІТУДНИЙ СПЕКТР, СТАЦІОНАРНИЙ ШУМ, ФІЛЬТРАЦІЯ, МОВНИЙ СИГНАЛ, АЛГОРИТМ, ПЕРЕТВОРЕННЯ ФУР'Є, ВІДНОВЛЕННЯ СИГНАЛУ

У роботі розглядається проблема шумозаглушення у цифрових аудіосигналах, зокрема мовних. Основну увагу приділено методу спектрального віднімання ефективному підходу до фільтрації стаціонарних та квазістаціонарних шумів.

Об'єкт дослідження: алгоритми цифрової фільтрації та шумозаглушення.

Мета: оцінити ефективність різних методів обробки сигналів (спектральне віднімання, адаптивні фільтри, фільтр Вінера) для підвищення якості аудіо в умовах шуму та розробити й вдосконалити алгоритм фільтрації для підвищення ефективності шумозаглушення.

У рамках роботи:

- досліджено методи фільтрації шуму;
- реалізовано та порівняно кілька алгоритмів;
- обрано оптимальний підхід та вдосконалено його;
- оцінено покращення якості сигналу до і після обробки.

Після порівняльного аналізу планується обрати один з алгоритмів фільтрації та провести його вдосконалення. Це дозволить продемонструвати вплив оптимізації на ефективність обробки сигналів — до та після покращення алгоритму.

Робота спрямована на покращення методів цифрової фільтрації та їх впровадження у сучасні цифрові системи, що використовуються в аудіотехніці, системах зв'язку та обробки мовлення. Результати роботи можуть бути застосовані для створення програмного забезпечення для поліпшення якості аудіосигналів, шумопригнічення в реальному часі та розробки більш ефективних цифрових фільтрів.

ABSTRACT

The explanatory note contains: 51 p., 19 fig., 4 appendices, 10 sources, 3 tables.

SPECTRAL, SUBTRACTION, NOISE REDUCTION, DIGITAL SIGNAL PROCESSING, AMPLITUDE SPECTRUM, STATIONARY NOISE, FILTERING, SPEECH SIGNAL, ALGORITHM, FOURIER TRANSFORM, SIGNAL RECOVERY

This work addresses the problem of noise reduction in digital audio signals, particularly speech signals. The main focus is on the spectral subtraction method, an effective approach for filtering stationary and quasi-stationary noise.

Object of study: Digital filtering and noise reduction algorithms.

Goal: To evaluate the effectiveness of various signal processing methods (spectral subtraction, adaptive filters, Wiener filter) for improving audio quality in noisy environments, and to develop and enhance a filtering algorithm to increase the efficiency of noise reduction.

Within the scope of this work:

- Noise filtering methods were investigated.
- Several algorithms were implemented and compared.
- The optimal approach was selected and improved.
- The improvement in signal quality before and after processing was evaluated.

After a comparative analysis, one of the filtering algorithms is planned to be selected and improved. This will demonstrate the impact of optimization on the effectiveness of signal processing—before and after the algorithm enhancement.

The work aims to improve methods of digital filtering and their implementation in modern digital systems used in audio technology, communication systems, and speech processing. The results of this work can be applied to create

software for improving audio signal quality, real-time noise suppression, and developing more effective digital filters.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ.....	10
ВСТУП.....	11
1 АНАЛІЗ ОСНОВНИХ ПРИНЦИПІВ ШУМОЗАГЛУШЕННЯ.....	12
1.1 Основи аналізу мовного сигналу та проблеми спотворень	12
1.2 Сприйняття сигналів і завад людським слухом.....	13
1.3 Класифікація методів шумозаглушення	15
1.3.1 Класифікація шумів.	15
1.3.2 Принципи шумоподавлення	16
1.3.3 Методи шумопригнічення.....	17
1.4 Загальний принцип роботи активного шумозаглушення	19
1.5 Цифрове представлення звукових сигналів	21
1.5.1 Імпульсно-кодова модуляція (ІКМ).....	22
1.5.2 Диференціальна імпульсно-кодова модуляція (ДІКМ)	23
1.5.3 Адаптивна диференціальна імпульсно-кодова модуляція (АДІКМ)....	23
1.5.4 Дельта-модуляція.....	23
1.5.5 Дельта-модуляція з безперервно змінюваною крутизною (CVSD).....	25
1.6 Аналіз цифрових методів пригнічення шуму.....	26
1.6.1 Алгоритми віднімання амплітудних спектрів.....	27
Найбільш популярними методами, заснованими на використанні спектральних характеристик шуму, є методи, що реалізують різні модифікації алгоритму вирахування амплітудних спектрів.	27
1.6.2 Метод оцінювання мінімальної середньоквадратичної помилки.....	28
1.6.3 Адаптивні компенсатори завад.....	29
1.6.4 Методи, які використовують статистичні властивості мовних сигналів у часовій області.....	30
1.6.5 Методи з використанням апарату прихованих марківських моделей..	30
1.6.6 Методи, які використовують властивості сприйняття мови людиною	31
2 АНАЛІЗ АЛГОРИТМУ ПРИГЛУШЕННЯ ШУМУ В МОВНОМУ СИГНАЛІ ЗА ДОПОМОГОЮ СПЕКТРАЛЬНОГО ВІДНІМАННЯ	34
2.1 Основи роботи алгоритму основоного на спектральному аналізі.	34
2.2 Перетворення Фур'є.....	35
2.3 Оцінка спектру шуму(Шумового профілю)	38

3 РЕАЛІЗАЦІЯ АЛГОРИТМУ	43
3.1 Функціонал вводу вивода аудіо	43
3.2 Підготовка аудіо до дитекції голосу	45
3.3 Дитекція голосу	46
3.4 Алгоритм початкової оцінки профілю шуму та алгоритм оновлення профілю шуму.....	47
3.5 Спектральне віднімання	48
3.6 Основна функція, що керування алгоритмом обробку аудіо	48
ВИСНОВКИ.....	50
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	51
ДОДАТОК А	Ошибка! Закладка не определена.
ДОДАТОК Б	Ошибка! Закладка не определена.
ДОДАТОК В	Ошибка! Закладка не определена.
ДОДАТОК Г	Ошибка! Закладка не определена.

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

DSP - digital signal processor

АМ – амплітудна модуляція

ЧМ – частотна модуляція

LPF – low pass filter

HPF – high pass filter

BPF – bandpass filter

АШП – активне шумопригнічення

ІКМ – імпульсно-кодова модуляція

ДІКМ – диференціальна імпульсно-кодова модуляція

АДІКМ – адаптивна диференціальна імпульсно-кодова модуляція

ДМ – дельта-модуляція

MMSE – метод оцінювання мінімальної середньоквадратичної помилки

SNR – signal-to-noise ratio

LMS – метод найменших середніх квадратів

ШПФ – швидке перетворення Фур'є

DFT - discrete Fourier transform

КПФ – короткочасне перетворення Фур'є

VAD – voice activity detection

ВСТУП

Шум є невід’ємною складовою будь-якої системи вимірювання та передачі сигналів. У процесі збору даних фізичні обмеження часто заважають отримати чистий сигнал, що призводить до виникнення небажаних збурень. У сфері інженерії шум визначається як небажане доповнення до сигналу, яке може негативно впливати на точність результатів обробки та подальше використання даних.

Актуальність дослідження методів шумозаглушення обумовлена зростаючою потребою у високоякісній цифровій обробці сигналів. Це особливо важливо для таких сфер, як телекомунікації, аудіообробка, системи розпізнавання мови, медична діагностика та інші галузі, де точність сигналу є критичною. В таких випадках ефективне пригнічення шуму є необхідною умовою забезпечення достовірності інформації.

Сучасна аудіотехніка та цифрові системи обробки сигналів (DSP) охоплюють широкий спектр застосувань — від професійного звукозапису до мультимедійних інтерфейсів у мобільних телефонах, камерах, персональних комп’ютерах та інтелектуальних пристроях. Із розвитком апаратного забезпечення на перший план виходить не швидкодія, а якість обробки аудіосигналу. Це відкриває нові можливості для впровадження інтелектуальних алгоритмів фільтрації.

У даній кваліфікаційній роботі досліджуються методи пригнічення адитивних стаціонарних шумів, з якими часто стикаються користувачі: шипіння мікрофона, гул електроприладів, вентиляційний шум тощо. Особливу увагу приділено алгоритмам спектрального віднімання.

Крім того, наводиться короткий огляд альтернативних методів фільтрації, зокрема фільтра Вінера та адаптивних фільтрів, із метою порівняльного аналізу.

1 АНАЛІЗ ОСНОВНИХ ПРИНЦИПІВ ШУМОЗАГЛУШЕННЯ

1.1 Основи аналізу мовного сигналу та проблеми спотворень

Голос і мова є основними засобами комунікації у всіх сферах людської діяльності, забезпечуючи передачу важливої інформації, яка може мати виняткову цінність у таких областях, як медіа, освіта, реклама та маркетинг, а також у наукових дослідженнях та технічних комунікаціях. Мова, зокрема, є складним акустичним сигналом, який містить численні характеристики, що забезпечують її розбірливість і точність сприйняття.

Людська мова являє собою шумоподібний акустичний сигнал, який включає амплітудну (АМ) і частотну модуляцію (ЧМ). Амплітудна модуляція відбувається, коли коливання сигналу змінюються по амплітуді відповідно до коливань низької частоти, а частотна модуляція — коли несуча частота змінюється згідно з модулюючими коливаннями.

Основна енергія акустичних коливань мовного сигналу зосереджена в діапазоні від 70 Гц до 7 кГц, при цьому понад 95% смислової інформації міститься в більш вузькому діапазоні від 200 Гц до 5 кГц. Вищі і нижчі частоти зазвичай несуть інформацію, що стосується емоційного забарвлення мови та особистісних характеристик мовця, а також можуть допомогти покращити розбірливість мови в умовах зашумленого середовища.

Основною причиною низької якості та розбірливості записаних мовних сигналів є наявність спотворень і завад. Спотворення — це зміни в структурі мовного сигналу, які призводять до зниження його якості. Вони можуть змінювати складові частини сигналу, що веде до втрати або спотворення інформації.

Завада — це стороннє обурення, що заважає коректному прийому сигналу і може спотворювати передану інформацію. Джерела завад можуть бути як зовнішні (наприклад, сторонні звуки), так і внутрішні (електронні або технічні

шуми у системах запису та передачі сигналу). Завади можуть проявлятися у вигляді таких явищ, як шерехи, тріск, погана розбірливість мови, а також чутність розмов, що ведуться через сусідні канали.

1.2 Сприйняття сигналів і завад людським слухом

Для вирішення практичних завдань шумоподавлення важливими є специфічні властивості слуху людини, зокрема частотне і часове маскування. Ці феномени описують, як певні звукові сигнали можуть ставати нечутними або маскованими іншими звуками, що має важливе значення для ефективного обробки аудіо та покращення якості сприйняття сигналу.

Частотне маскування — це явище, яке виникає, коли сигнал з низьким рівнем (маскований сигнал) стає нечутним при одночасному звучанні з більш сильним сигналом (маскером), якщо вони перебувають у близьких частотних діапазонах. Це явище широко використовується в обробці аудіо, зокрема при стисненні звуку, де шум або фонові перешкоди маскуються основним сигналом.

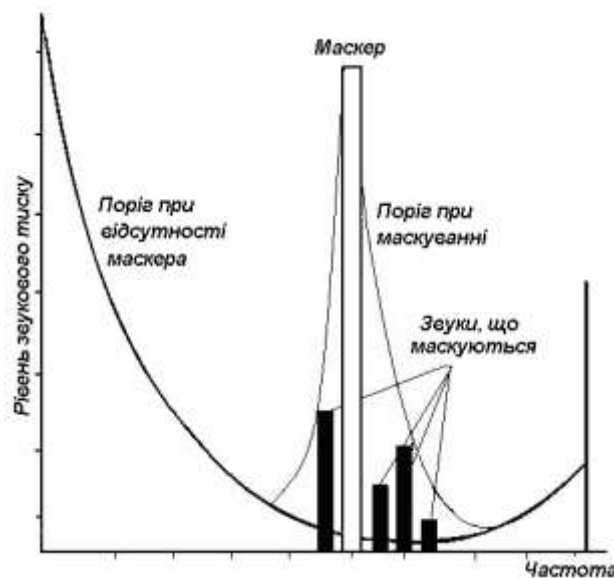


Рисунок 1.1 – Миттєве маскування в частотній області
для слухової системи людини

На рис. 1.1 зображено, як сильний вузькосмуговий сигнал (маскер) перебиває чотири слабших сигналу, розташованих поблизу. Вони виявляються нечутними через ефект маскування, бо їх рівень знаходиться нижче порога чутності. Це означає, що навіть якщо в сигнал додаються шумові компоненти на цих частотах, слухач може не сприйняти їх, оскільки маскуються потужним основним сигналом.

Часове маскування, на відміну від частотного, пов'язане з тим, як сигнали маскуються в часі. Після припинення подачі маскуючого звуку, слабший сигнал ще деякий час залишається нечутним, навіть якщо його рівень підвищується. Це явище може тривати до 100 мс, залежно від частоти та амплітуди сигналу. Наприклад, звуковий вибух або різкий звук може тимчасово «закрити» сприйняття інших сигналів, навіть якщо вони розташовані на іншій частоті.

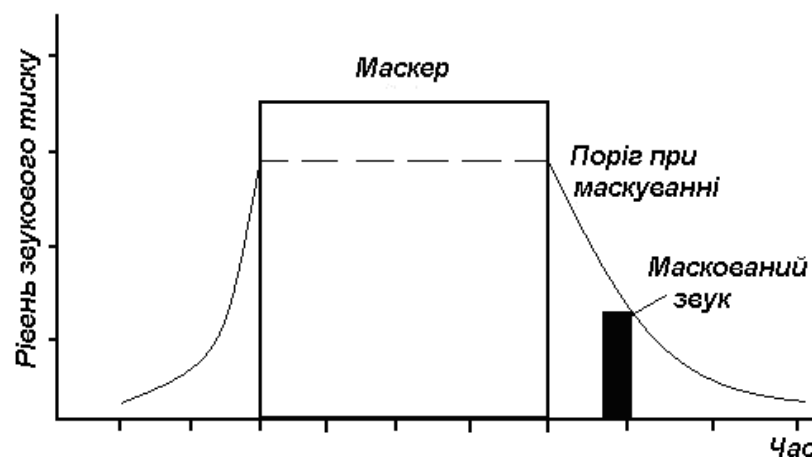


Рисунок 1.2 – Приклад часового маскування

На рис. 1.2 показано, як сигнал маскера продовжує впливати на інші звуки після свого припинення, залишаючи їх нечутними, якщо їх рівень знаходиться нижче порога чутності.

У зв'язку з вищесказаним, досить часто виникає потреба в проведенні спеціальної обробки або шумоподавлення звукового сигналу, головною метою якого є підвищення якості і розбірливості записаної на фонограмі мови, для її подальшого комфортного прослуховування і сприйняття.

1.3 Класифікація методів шумозаглушення

1.3.1 Класифікація шумів.

Існує безліч різних класифікацій шумів, наприклад, за характером спектра або частотою хвиль. Однак, якщо є потреба позбавитися шумів у записі, варто в першу чергу враховувати категоризацію шумів за часовими характеристиками.

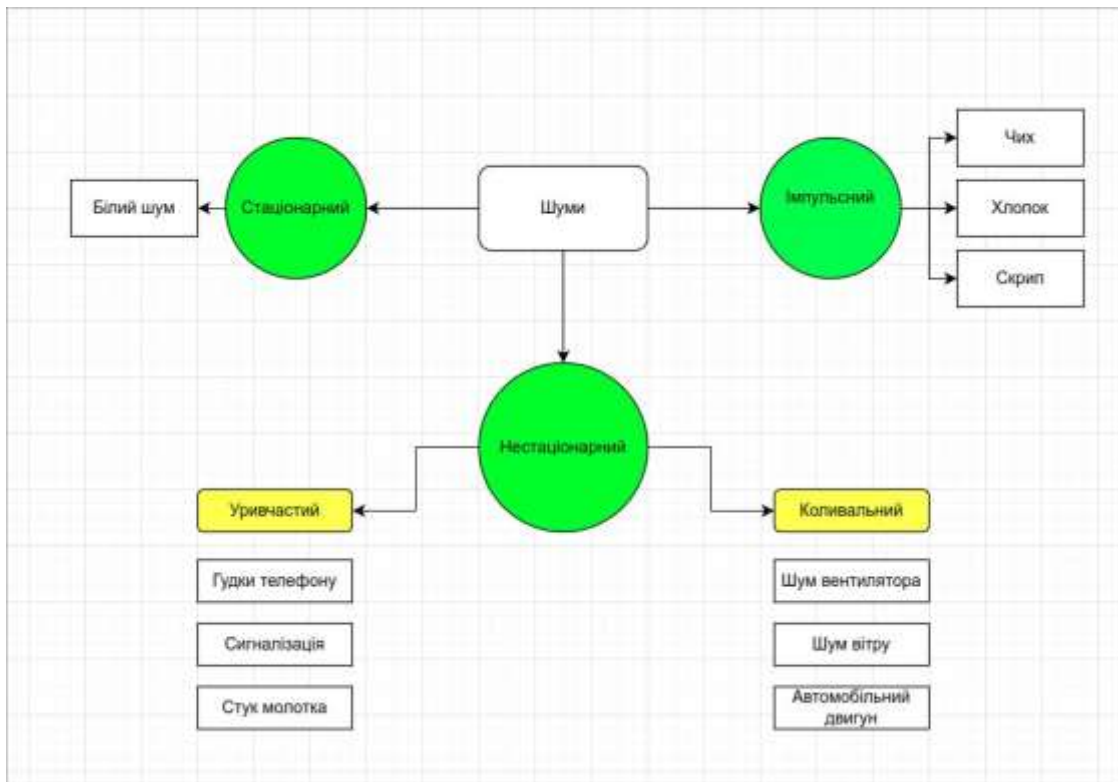


Рисунок 1.3 – Класифікація шумів

Часові характеристики шуму тісно пов'язані зі способом його утворення. Стаціонарний і коливальний шум, як правило, виникають у результаті постійних процесів (природних або штучних), тоді як уривчастий і імпульсний шум – через різкі одноразові процеси. Уривчастий шум для простоти можна сприймати як імпульсний шум, що повторюється з певною періодичністю.

Категоризація шумів необхідна для визначення складності їхнього придушення. Найлегше усунути шум, якщо нам заздалегідь відомо, який саме тип шуму присутній у сигналі та де він знаходиться. Зокрема, стаціонарний шум можна усунути досить просто, оскільки його поріг гучності у спектрі

залишається незмінним. Наприклад, білий шум рівномірно розподілений по всьому сигналу, що дозволяє легко його ідентифікувати у фрагментах тиші.

1.3.2 Принципи шумоподавлення.

Основним принципом шумоподавлення є підкреслення корисного сигналу в тих частотних ділянках, де завада є сильною, а корисний сигнал — слабким. Це дозволяє зменшити рівень шуму та одночасно зберегти або посилити потрібні компоненти сигналу. Цей процес найкраще реалізується за допомогою цифрової фільтрації, оскільки вона дозволяє точніше налаштувати параметри обробки, на відміну від аналогової, яка має обмеження в цьому аспекті.

Найефективніші методи шумоподавлення базуються на використанні різних фільтраційних технік. Ці методи можуть варіюватися залежно від типу шуму. Наприклад, стаціонарний шум часто усувається за допомогою фільтрів, які зменшують рівень шуму на певних частотах, в той час як для імпульсних шумів використовуються алгоритми, які розпізнають і видаляють різкі стрибки в амплітуді сигналу.

Таблиця 1.1 – Методи шумоподавлення

Метод 1	Опис 2	Застосування 3
Розмаскування мовного сигналу в частотній та часовій областях	Усунення ефекту слухового маскування, перерозподіл енергії сигналу для кращого сприйняття.	Використовується для підкреслення корисного мовного сигналу, зменшення впливу завад.
Усунення фонових шумів	Зниження інтенсивності фонових шумів, що зменшує стомлюваність слухача.	Застосовується в шумних середовищах, для зниження зусиль слухача при прослуховуванні.
Зниження смуги частот сигналу	Скорочення смуги частот для зменшення стомлення, видалення низькочастотного гулу та високочастотного шипіння.	Підходить для ситуацій, де важливо зберегти лише корисний частотний діапазон.

Продовження табл.1.1

1	2	3
Вирівнювання амплітудних піків	Зменшення різких перепадів гучності сигналу для збереження рівного звукового потоку.	Застосовується при значних перепадах гучності сигналу, особливо в записах.
Видалення імпульсних завад	Усунення короточасних та інтенсивних перешкод, таких як електричні сплески або удари.	Корисно для видалення короточасних перешкод в аудіо.
Згладжування спектра сигналу	Вирівнювання амплітудно-частотної характеристики для усунення спотворень.	Застосовується для корекції спотвореного спектра сигналу, наприклад, після передачі через канал.
Формування оптимальної частотної характеристики для слуху користувача	Корекція частотної характеристики для покращення слухової сприйнятливості користувача.	Використовується для персоналізації сигналу в залежності від індивідуальних потреб слухача.

Загалом, вибір методу обробки сигналу залежить від характеру шуму та наявних вимог до якості сигналу.

Найбільш загальним підходом у визначенні типу спотворень є аналіз спектра і осцилограми, автокорелограми і інших сигналів. На основі знань про характер шумів і спотворень можна вибрати найкращий метод або ряд послідовно (в певному порядку) реалізованих методів обробки мовних сигналів.

Якщо характеристики завади змінюються в часі, то необхідно використовувати алгоритми обробки, що автоматично налаштовують (тобто адаптують) параметри фільтра до характеристик завади. Це може бути забезпечено алгоритмами адаптивної фільтрації.

1.3.3 Методи шумопригнічення.

Методи шумопригнічення можна умовно поділити на три основні групи: фільтри, комбіновані фільтри та нейромережі. Кожна з цих груп має свої особливості та переваги, залежно від характеру шуму та конкретних вимог до якості сигналу.

Фільтрація є основним методом обробки звукових сигналів і включає як традиційні аналогові, так і сучасні цифрові фільтри.

Аналогові фільтри застосовуються рідше через свою обмежену точність і складність налаштувань. Вони можуть бути корисними для базових завдань у реальному часі.

Цифрові фільтри є основними для сучасних приладів шумопригнічення. Серед них:

- фільтр низьких частот (LPF) — усуває високочастотний шум, пропускаючи тільки сигнали нижче певної частоти;
- фільтр високих частот (HPF) — видаляє низькочастотні шуми, наприклад, гул, залишаючи тільки високочастотний корисний сигнал;
- фільтри смугового пропускання (BPF) — пропускають сигнали в певному діапазоні частот, блокуючи як низькочастотні, так і високочастотні шуми;
- адаптивні фільтри — використовують методи адаптації для динамічного коригування фільтрації залежно від рівня і типу шуму.

Важливі методи шумопригнічення в цифрових фільтрах:

Фільтр Вінера — це адаптивний фільтр, який є одним з найефективніших методів для зменшення шуму в сигналах з відомими статистичними властивостями. Він працює на основі оцінки апостеріорної ймовірності корисного сигналу та шуму, що дозволяє оптимізувати співвідношення сигнал/шум.

Спектральне віднімання — метод, що базується на аналізі спектра сигналу. Він передбачає визначення спектральної компоненти шуму та її подальше віднімання від спектра зашумленого сигналу. Це ефективний метод для усунення білого шуму.

Спектральне гейтування — цей метод передбачає застосування "гейту" для відсіювання частотних компонент, рівень яких нижчий за певний поріг. Це

дозволяє зменшити фонові шуми та залишити лише корисні компоненти сигналу.

Важливою складовою багатьох методів фільтрації є використання вікон, що допомагає знизити ефекти розривів сигналу при його обробці. Вікна дозволяють обробляти сигнал у часткових інтервалах часу та зменшувати спотворення, що виникають через обмеження фіксованих ділянок сигналу.

Для досягнення кращих результатів також використовуються комбіновані методи, що поєднують різні види фільтрації та застосовують різні вікна для обробки сигналів на етапах спектрального аналізу.

Нейромережі стали основою для сучасних методів шумопригнічення завдяки здатності адаптуватися до різних умов шуму і розпізнавати складні патерни в даних. Використання алгоритмів машинного навчання дозволяє нейромережам автоматично розпізнавати зашумлені і чисті сигнали, а також адаптивно налаштовувати параметри фільтрації відповідно до характеристик конкретного запису.

Комбінування різних фільтрів дозволяє значно покращити якість обробки звуку. Використання нейромереж є більш сучасним підходом, який дозволяє досягти ефективного шумозаглушення в реальному часі, адаптуючись до різних типів шуму.

1.4 Загальний принцип роботи активного шумозаглушення

Активне шумопригнічення (АШП) засноване на принципі протифазної компенсації. Суть цього методу полягає в тому, що генерується компенсаційний сигнал, який має амплітуду, подібну до шуму, але з інвертованою фазою. Завдяки цьому, сигнал пригнічення скасовує шум, і в результаті отримуємо значно зменшену завадність в аудіосигналі.

У ідеальних умовах компенсаційний сигнал повинен точно відповідати шуму за амплітудою та фазою, але так, щоб їхні коливання були зворотними. Це дозволяє досягти максимального ефекту пригнічення шуму. Якщо ж амплітуда або фаза компенсаційного сигналу відрізняються від шуму, це призводить до неефективності системи, оскільки замість придушення шуму може виникнути додаткове посилення.

Основне завдання в системах АШП — це створення точного сигналу пригнічення. Однак для цього необхідно враховувати кілька важливих факторів, включаючи мінімізацію затримок в обробці сигналу. Вся затримка, яка виникає між отриманням опорного сигналу та генеруванням сигналу пригнічення, може негативно впливати на якість системи і призвести до погіршення результатів, оскільки частина шуму може залишатися незаглушеною.

Для ефективної роботи активного шумопригнічення важливо, щоб система працювала в реальному часі з мінімальними затримками. Вимоги до таких систем є надзвичайно складними, адже необхідно забезпечити високошвидкісну обробку сигналу з високою точністю в режимі реального часу.

Аналогові системи шумопригнічення з часом стикалися з технічними обмеженнями. Зазвичай такі схеми були реалізовані в одному каналі, і їх здатність до зменшення шуму була обмежена можливостями аналогових фільтрів. Часто такі системи мали обмеження в регулюванні коефіцієнта посилення, і шум можна було приглушити лише в залежності від вхідного рівня сигналу, без можливості точного налаштування.

Однією з основних проблем аналогових систем було те, що зменшення коефіцієнта посилення не завжди могло бути динамічно адаптованим до змінного рівня шуму. В результаті, навіть незначні зміни в умовах акустичного середовища могли призвести до зниження ефективності системи.

З появою цифрових методів обробки сигналу почалося використання більш складних алгоритмів, що здатні динамічно адаптуватися до змін у рівнях шуму та вимогах до зменшення посилення. Такі системи можуть аналізувати

характер шуму, визначати найбільш відповідні частотні діапазони, де потрібно зменшити амплітуду, а також адаптувати коефіцієнт посилення в реальному часі.

Цифрові системи також дозволяють використовувати більш складні методи, такі як фільтрація за допомогою адаптивних фільтрів, спектральне віднімання або алгоритми на основі машинного навчання, що дозволяють ефективно визначати, який сигнал є шумом, а який — корисним. Завдяки цим методам можна значно покращити точність шумопригнічення, враховуючи мінливість умов шуму.

Важливу роль у роботі активного шумопригнічення відіграє розташування мікрофонів та гучномовців в акустичному середовищі. Правильне розташування цих компонентів має вирішальне значення для ефективності системи, оскільки навіть незначні зміни в розташуванні можуть суттєво вплинути на точність виявлення та компенсації шуму. Наприклад, неправильно налаштовані мікрофони можуть спотворювати отриманий сигнал, що ускладнить генерування компенсаційного сигналу.

Активне шумопригнічення є складною і комплексною системою, яка об'єднує кілька галузей знань: обробка сигналів, техніка управління, акустичні вимірювання та системи реального часу. Для її ефективною реалізації необхідні спеціалізовані знання, які дозволяють налаштувати кожен етап процесу, від отримання сигналу до випромінювання сигналу пригнічення.

Розвиток цифрових технологій і новітніх алгоритмів дозволяє значно підвищити ефективність шумопригнічення в реальному часі, що відкриває нові можливості для застосування АШП у різних галузях, від мобільних пристроїв до великих акустичних систем.

1.5 Цифрове представлення звукових сигналів

Цифрове представлення звукових сигналів передбачає перетворення аналогового аудіосигналу на цифрову форму для подальшої обробки, зберігання

або передавання. Це досягається шляхом дискретизації за часом і квантування амплітуди сигналу. Такий підхід забезпечує високу точність відтворення звуку, сумісність з комп'ютерними системами, цифровими носіями та мережевими технологіями.

1.5.1 Імпульсно-кодова модуляція (ІКМ).

Імпульсно-кодова модуляція – базовий метод перетворення первинного аналогового сигналу в цифрову форму. При ІКМ значення дискретних відліків піддають квантуванню за рівнем, потім квантовані відліки кодують в двійковій системі числення. Особливістю ІКМ є незалежне кодування кожного відліку. Частоту дискретизації необхідно вибирати згідно теореми В. А. Котельникова, відповідно до якої сигнал з обмеженим спектром частот f_s можна представити сукупністю відліків з тактовою частотою, що перевищує не менш, ніж у 2 рази найвищу частоту спектра сигналу.

Квантування призводить до шуму квантування, величина якого зменшується зі збільшенням кількості бітів. Наприклад, 16-бітне кодування забезпечує якість CD-аудіо (~98 дБ), тоді як 24-бітне — студійну якість з більш точним збереженням звукових нюансів.

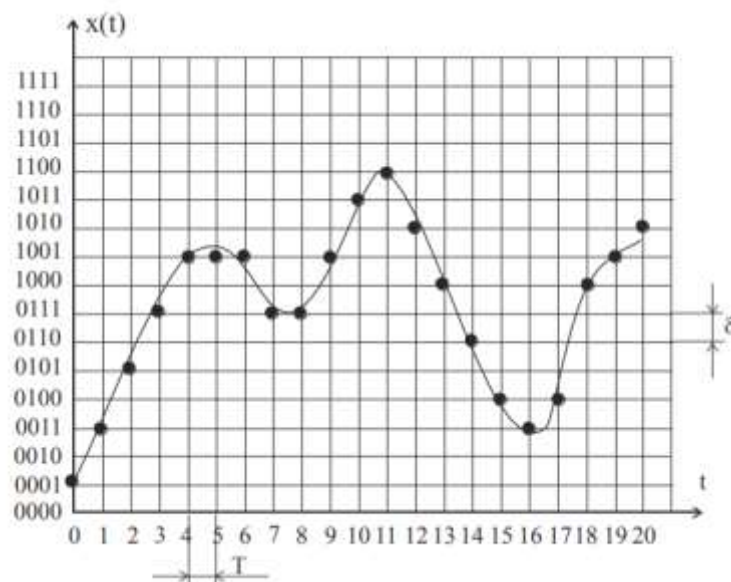


Рисунок 1.4 – Приклад аналого-цифрового перетворення: δ – крок квантування сигналу по амплітуді; T – крок дискретизації

1.5.2 Диференціальна імпульсно-кодова модуляція (ДІКМ).

Диференціальна ІКМ (ДІКМ, англ. DPCM — Differential PCM) замість абсолютних значень сигналу кодує різницю між поточним і попереднім відліком. Це дозволяє зменшити обсяг даних, особливо якщо сигнал змінюється повільно. Однак точність залежить від здатності точно прогнозувати значення сигналу, і накопичення помилок може знижувати якість.

1.5.3 Адаптивна диференціальна імпульсно-кодова модуляція (АДІКМ).

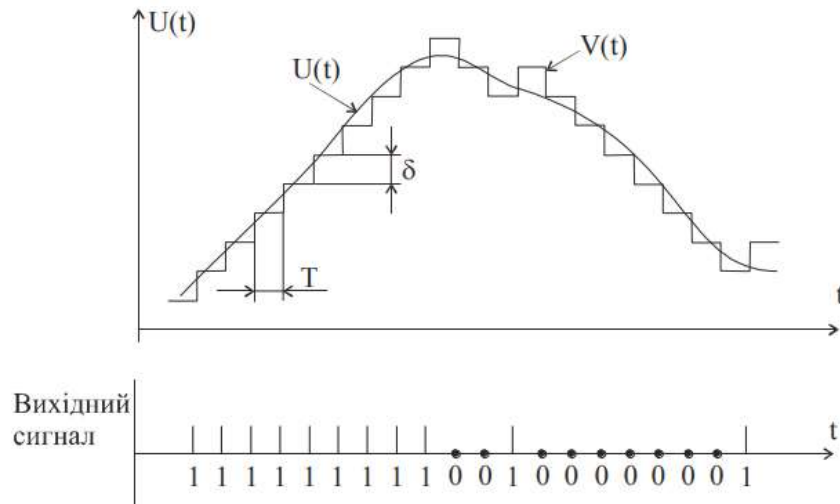
Адаптивна диференціальна ІКМ (АДІКМ, англ. ADPCM) розвиває ідею ДІКМ шляхом адаптації кроку квантування. Якщо сигнал змінюється швидко — крок збільшується; якщо повільно — зменшується. Такий підхід дозволяє ефективніше передавати змінні сигнали при збереженні низької швидкості передавання даних.

1.5.4 Дельта-модуляція.

Дельта-модуляція (ДМ, англ. Δ -modulation) — спрощений варіант ДІКМ, де використовується всього один біт на кожен відлік. Він показує, чи збільшився сигнал відносно попереднього значення, чи зменшився. ДМ проста в реалізації та добре підходить для передачі з низькою швидкістю, однак погано справляється з швидкозмінними сигналами, що може спричинити перевантаження по крутизні (сигнал змінюється швидше, ніж система може його відстежити). У порівнянні зі своїми конкурентами, ІКМ і АДІКМ, дельта-модуляція характеризується меншою складністю технічної реалізації, більш високою перешкодозахищеністю.

Використання одного розряду для представлення помилки передбачення приводить до специфічних особливостей (рис. 1.5):

- при кодуванні сигналів постійного рівня апроксимуючий сигнал "скаче" відносно кодованого рівня (гранулярний шум);
- при кодуванні швидкозмінних сигналів з'являються помилки, обумовлені неможливістю зміни апроксимуючого сигналу більш, ніж на один крок квантування (перевантаження по крутизні) (рис. 1.6).



T — крок дискретизації, δ — крок квантування,

$U(t)$ — вхідний аналоговий сигнал, $V(t)$ — апроксимуюча напруга

Рисунок 1.5 – Форма сигналу при дельта-модуляції

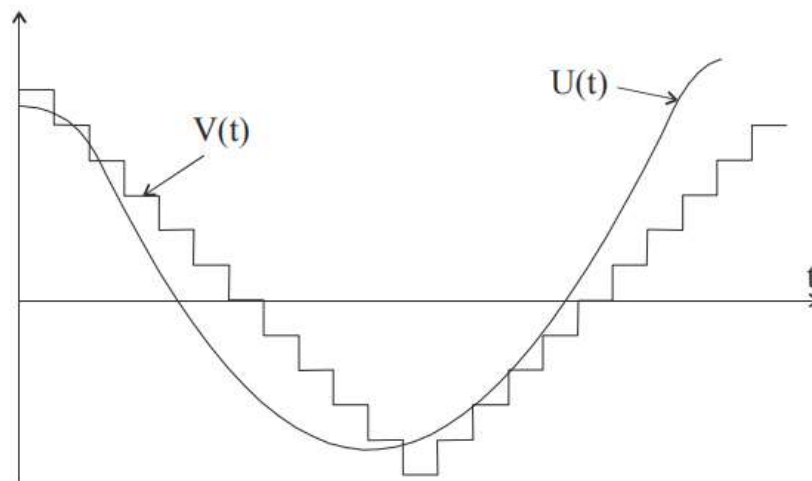


Рисунок 1.6 – Перевантаження по крутизні гармонійним сигналом при ДМ із постійним кроком збільшення апроксимуючої напруги: $U(t)$ — вхідний аналоговий сигнал, $V(t)$ — апроксимуюча напруга

1.5.5 Дельта-модуляція з безперервно змінюваною крутизною (CVSD).

Метод CVSD (Continuous Variable-Slope Delta-modulation) усуває основний недолік звичайної дельта-модуляції (ДМ) — перевантаження при швидких змінах сигналу, адаптуючи крок апроксимації до швидкості зміни вхідного сигналу: збільшуючи його при швидких змінах і зменшуючи при повільних. Це дозволяє досягти кращого співвідношення сигнал/шум, і тому метод часто використовується в Bluetooth, військових радіозасобах та голосових рекордерах. Більшість варіантів ДМ, що усувають перевантаження, базуються на миттєвому чи інерційному компандуванні сигналу або адаптивній зміні сходинок апроксимуючої напруги відповідно до швидкості зміни сигналу.

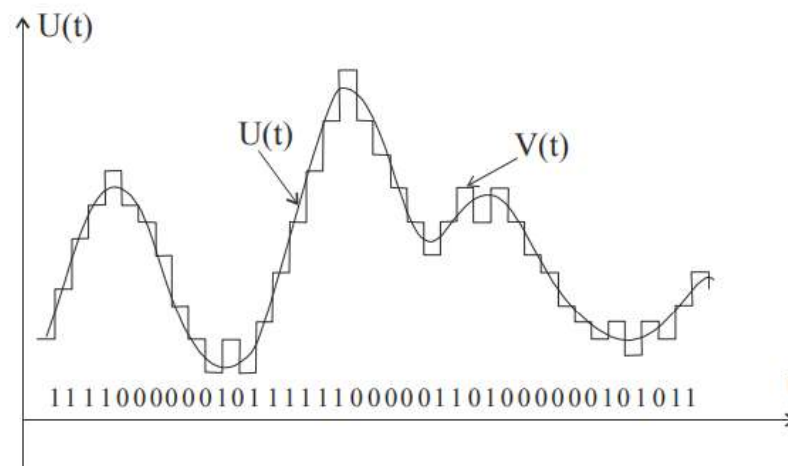


Рисунок 1.7 – Дельта-модуляція з безперервно змінюваною крутизною

Таблиця 1.2 – Методи цифрового представлення звукових сигналів

Методи	Переваги	Недоліки	Типове застосування
1	2	3	4
ІКМ (Імпульсно-кодова модуляція)	Висока якість, проста реалізація, універсальність	Великий обсяг даних	CD-аудіо, РСМ у комп'ютерах, студійний звук
ДІКМ (Диференціальна ІКМ)	Менше бітів, ефективне стискання сигнали з повільними змінами	Чутлива до накопичення помилок у випадку неточного прогнозування	Телефонія, системи передачі мовлення

Продовження табл.1.2

1	2	3	4
АДІКМ (Адаптивна ДІКМ)	Адаптується до змін сигналу, ефективна при кодуванні змінних сигналів	Складніша в реалізації, можливі артефакти при різких стрибках сигналу	VoIP, мобільний зв'язок, аудіо кодеки
Дельта-модуляція	Проста реалізація, мінімум бітів	Велике спотворення при швидких змінах	Низькошвидкісні передачі, прості аудіо кодеки
CVSD (Delta-модуляція з безперервно змінюваною крутизною)	Покращене слідкування за швидкими змінами, вища якість	Більш складна логіка адаптації	Bluetooth, військовий зв'язок

У більшості сучасних цифрових звукових систем використовуються стандартні частоти дискретизації 44.1 і 48 кГц, проте частотний діапазон сигналу звичайно обмежується біля 20 кГц для залишення запасу по відношенню до теоретичної межі. Також найбільш поширене 16-розрядне квантування по рівню, що дає граничне співвідношення сигнал/шум близько 98 дБ. У студійній апаратурі використовуються вищі рівні 18, 20 і 24-розрядне квантування при частотах дискретизації 56, 96 і 192 кГц. Це робиться для того, щоб зберегти вищі гармоніки звукового сигналу, які безпосередньо не сприймаються слухом, але впливають на формування загальної звукової картини.

Для менш якісних сигналів, наприклад, у телефонії, застосовуються 7-8 розрядів та частоти 8-12 кГц. Вибір методу залежить від вимог до якості, швидкості передачі та складності реалізації.

1.6 Аналіз цифрових методів пригнічення шуму

Цифрове пригнічення шуму є ключовим етапом у покращенні якості звукових сигналів. Існує кілька основних алгоритмів та методів основаних на них, які використовуються для цієї мети.

Основні алгоритми пригнічення шуму:

- алгоритми віднімання амплітудних спектрів;
- метод оцінювання мінімальної середньоквадратичної помилки;
- адаптивні компенсатори завад;
- методи, які використовують статистичні властивості мовних сигналів у часовій області;
- методи з використанням апарату прихованих марківських моделей;
- методи, які використовують характерні властивості мовного сигналу.

Кожен з цих алгоритмів має свої особливості реалізації, переваги та недоліки. Нижче розглянемо їх детальніше.

1.6.1 Алгоритми віднімання амплітудних спектрів.

Найбільш популярними методами, заснованими на використанні спектральних характеристик шуму, є методи, що реалізують різні модифікації алгоритму вирахування амплітудних спектрів.

Як обґрунтування цих методів наводяться такі міркування. Якщо стаціонарний сигнал $s(t), t = \dots -1, 0, 1, \dots$ зі спектральною щільністю потужності $P_{ss}(j\omega)$ спотворений адитивним стаціонарним шумом $n(t)$ зі спектральною щільністю потужності $P_{nn}(j\omega)$, який приймається некорельованим із $x(t)$, то спектральна щільність потужності зашумленого сигналу $P_{xx}(j\omega)$ дорівнює:

$$P_{xx}(j\omega) = P_{ss}(j\omega) + P_{nn}(j\omega) \quad (1.1)$$

отже спектральна щільність потужності корисного сигналу $s(n)$ може бути оцінена як:

$$P_{ss}(j\omega) = P_{xx}(j\omega) - P_{nn}(j\omega) \quad (1.2)$$

В силу нестационарності мовних сигналів використовувати співвідношення (1.1) безпосередньо не можна. На практиці, при обробці мови на досить коротких ділянках, наприклад, квазістаціонарних ділянках голосних звуків, величини $P_{xx}(j\omega)$, $P_{nn}(j\omega)$ апроксимують за допомогою усереднених квадратів короточасних амплітудних спектрів досліджуваного сигналу і шуму. Спектр шуму при цьому повинен оцінюватися в моменти пауз. Отримана таким чином оцінка відповідає квадрату амплітудного спектра сигналу. Відновлення

мовного сигналу у часовій області здійснюється за допомогою зворотного перетворення Фур'є, причому фазовий спектр, для відновленого сигналу береться таким же, як і у досліджуваного сигналу.

У найбільш загальному вигляді операція спектрального віднімання може бути виражена співвідношенням:

$$|S(t, j\omega)|^2 = \begin{cases} |X_i(t, j\omega)|^2 - A(t)|N(t, \nu)|^2 \\ B|N(t, j\omega)|^2 \end{cases},$$

якщо

$$|X_i(t, \nu)|^2 \geq (A(t)+B)|N(t, j\omega)|^2.$$

(1.3)

Тут $A(t)$ – фактор переоцінювання, який залежить від відношення сигнал-шум на сегменті аналізу, і має типові значення близько 0,7 – 0,95, а B – спектральний поріг, що обирається в діапазоні від 0,01 до 0,1.

Дослідження якості та чіткості мови, показали, що в тих випадках, коли шум або завада мають стаціонарний (або квазістаціонарний) характер і їх спектр має гармонійну структуру, досягається значне на слух підвищення як якості, так і розбірливості мови. Однак, в разі шумів з швидкозмінюваними спектральними характеристиками така обробка малоефективна.

Також використовуючи метод віднімання амплітудних спектрів можна зробити подавлення адитивного квазістаціонарного шуму. Як типовий приклад можна привести шуми кондиціонерів, відеокамери, трансформаторів і підсилювачів. Поведінка алгоритму контролюється набором параметрів, включаючи попередньо виміряні характеристики шуму.

1.6.2 Метод оцінювання мінімальної середньоквадратичної помилки.

Метод оцінювання мінімальної середньоквадратичної помилки (MMSE) — це алгоритм, спрямований на оцінку амплітудного спектра сигналу для шумоподавлення. Метод MMSE вважається одним з найбільш ефективних для одного мікрофона, оскільки він значно зменшує шум без створення артефактів,

таких як "музичні" тони. Перевага MMSE над фільтром Вінера та відніманням спектрів зумовлена використанням апріорної оцінки відношення сигнал/шум (SNR) у кожній спектральній смузі. Модифікації методів, що також використовують апріорні SNR, показують поліпшення результатів фільтрації. MMSE застосовується для мінімізації середньоквадратичної помилки амплітудного спектра, що дозволяє знижувати шум без значних втрат якості сигналу.

Фільтр Вінера є класичним методом для мінімізації середньоквадратичної помилки між бажаним сигналом і вихідним сигналом фільтра. Він використовує статистичні властивості сигналу та шуму, розраховуючи оптимальну передавальну функцію, яка залежить від SNR на кожній частоті. Однак, фільтр Вінера не завжди дає найкращі результати для нестационарних шумів, і саме MMSE стає більш адаптивним для таких випадків.

Метод найменших середніх квадратів (LMS) — це адаптивний алгоритм фільтрації, який використовується для пошуку коефіцієнтів фільтра, що мінімізують середньоквадратичну помилку між бажаним сигналом та вихідним сигналом фільтра. На відміну від фільтра Вінера, який вимагає знання статистичних характеристик сигналу та шуму, LMS є градієнтним методом, який ітеративно коригує коефіцієнти фільтра на кожному кроці на основі миттєвої оцінки градієнта функції середньоквадратичної помилки, що дозволяє йому адаптуватися до змінних умов.

1.6.3 Адаптивні компенсатори завад.

Адаптивні компенсатори завад використовують один або кілька опорних сигналів, корельованих із шумом, але некорельованих з корисним сигналом. Сигнал завади оцінюється за допомогою опорного сигналу, після чого він віднімається від зашумленого сигналу, що дозволяє отримати оцінку незашумленого сигналу. Ключовим елементом є пристрій управління ваговими

коефіцієнтами (ПУВК), який оптимізує коефіцієнти фільтра для мінімізації енергії сигналу на виході і максимізації відношення сигнал/шум.

Адаптивні компенсатори завад можуть значно покращити якість сигналу, зменшуючи рівень шуму на кілька десятків децибел. Однак їх застосування обмежене через потребу в наявності опорного сигналу. У випадках, коли опорний сигнал відсутній, його можна отримати на основі властивостей мовного сигналу, що робить компенсатор частиною складніших алгоритмів виділення мовного сигналу.

Ці компенсатори застосовуються для фільтрації стаціонарних завад, тональних шумів і деяких нестаціонарних завад, але обмежені через необхідність наявності опорного сигналу.

1.6.4 Методи, які використовують статистичні властивості мовних сигналів у часовій області

Методи цифрової обробки зашумлених мовних сигналів, засновані на математичних моделях мовних сигналів, швидко розвиваються та дають успішні результати. Завдання виділення мовного сигналу із шуму зводиться до оцінки параметрів моделі та подальшого синтезу або фільтрації сигналу.

Одним з перспективних підходів є статистична фільтрація в часовій області, зокрема авторегресійні моделі. Фільтрація здійснюється за допомогою оптимальних методів, таких як фільтр Кальмана. Обчислювально ефективна реалізація фільтрації на основі марківського кола використовує алгоритм максимізації математичного очікування та фільтр Кальмана-Бьюсі.

1.6.5 Методи з використанням апарату прихованих марківських моделей.

Іншим класом методів обробки зашумлених мовних сигналів, заснованих на використанні статистичних моделей мовного сигналу, є методи, в яких мовний сигнал моделюється прихованим марківським колом. Тобто, для моделювання мовного сигналу використаний найбільш ефективний для розпізнавання мови підхід.

Традиційні методи фільтрації (віднімання спектрів або фільтр Вінера) не використовують фонетичну інформацію, що переноситься мовним сигналом. Недавні дослідження показали, що знання і застосування в процесі обробки фонетичної структури сигналу призводить до поліпшення якості фільтрації. Тому цілком природним є застосування в процесі очищення мовного сигналу від шумів його статистичної моделі у вигляді прихованого марківського кола, який пов'язаний з фонетичною структурою сигналу.

Ідея полягає у побудові статистичних моделей фонем або ширших класів звуків на основі незашумленого мовлення. Після цього для кожного стану моделі розраховується оптимальний фільтр Вінера.

При обробці зашумленого сигналу спочатку оцінюється поточний стан ПММ (на основі попередньо відфільтрованого сигналу), на основі чого обирається відповідний оптимальний фільтр для подальшої фільтрації.

Для ефективної обробки нестационарного шуму окремо моделюється марківська модель шуму з використанням набору гауссовських компонент. При виявленні пауз використовується алгоритм Вітербі для вибору найбільш ймовірної моделі шуму, яка потім застосовується для обробки. Для збереження плавності між ітераціями використовується інерційна схема фільтра Вінера.

В даний час методи, засновані на використанні статистичних моделей мовних сигналів, є найбільш перспективними і їх подальша розробка та удосконалення допоможе досягти нових результатів у сфері шумоподавлення мовних сигналів.

1.6.6 Методи, які використовують властивості сприйняття мови людиною.

Розвиток методів обробки мовних сигналів базується на принципі, що аналіз, заснований на моделі слуху людини, є більш ефективним, ніж абстрактні моделі мовотворення або статистичні марковські моделі. Один з методів ґрунтується на чутливості слуху до модуляцій у спектральному обвідному сигналі з частотою 2-3 Гц, що покращує розбірливість мовного сигналу до зашумлення, але не підвищує її після шуму.

Інший метод використовує принцип частотного маскування, де сигнал низького рівня стає нечутним при наявності сильнішого сигналу з близькою частотою. Також існує метод, де мовний сигнал спочатку піддається високочастотній фільтрації для посилення вищих формант, а потім — кліппінгу для збільшення амплітуди важливих для сприйняття приголосних звуків.

Один з перспективних методів — хвильовий (вейвлетний) аналіз, який моделює ефект маскування слуховою системою, коли сильніший сигнал пригнічує слабший. Використання цієї методики дозволяє досягти виграшу до 26 дБ при гауссовому шумі, що в реальних умовах дає виграш не менше 18 дБ.

Огляд методів підвищення якості мовних сигналів показує багато різних підходів, що вказує на важливість проблеми очищення сигналу, але також на відсутність універсальних і оптимальних рішень. Вибір методу залежить від конкретних завдань, таких як підвищення чіткості мовлення чи збереження натуральності голосу.

Таблиця 1.3 – Цифрові методи пригнічення шуму

Метод	Переваги	Недоліки
1	2	3
Алгоритм віднімання амплітудних спектрів	<ul style="list-style-type: none"> - Простота реалізації. - Добре працює для стаціонарних або квазістаціонарних шумів. - Швидка обробка сигналу. 	<ul style="list-style-type: none"> - Малоефективний для нестационарних шумів. - Може знижувати якість сигналу при високому рівні шуму.
Метод оцінювання мінімальної середньоквадратичної помилки (MMSE)	<ul style="list-style-type: none"> - Хороша ефективність при низьких значеннях SNR. - Зменшує шум без артефактів ("музичних" тонів). - Враховує апіорну інформацію про SNR. 	<ul style="list-style-type: none"> - Висока обчислювальна складність. - Потребує точної оцінки SNR для оптимальної роботи.

Продовження табл.1.3

1	2	3
Фільтр Вінера	<ul style="list-style-type: none"> - Хорошо працює при стабільному рівні шуму. - Простий у використанні для оцінки середньоквадратичної помилки. 	<ul style="list-style-type: none"> - Погано працює при нестационарних шумів. - Залежить від точності оцінки спектральних щільностей шуму та сигналу.
Адаптивні компенсатори завад	<ul style="list-style-type: none"> - Значно покращують якість сигналу. - Можливість застосування для різних типів шуму. - Може працювати з кількома опорними сигналами. 	<ul style="list-style-type: none"> - Потребує наявності опорного сигналу. - Не завжди ефективні для складних нестационарних шумів
Методи статистичної фільтрації в часовій області	<ul style="list-style-type: none"> - Може значно покращити SNR. - Ефективно використовує авторегресійні моделі та фільтри Кальмана для обробки сигналу. 	<ul style="list-style-type: none"> - Висока обчислювальна складність. - Потребує точних статистичних моделей для досягнення високих результатів.
Методи з використанням апарату прихованих марківських моделей	<ul style="list-style-type: none"> - Добре справляються з нестационарними шумами. - Враховують фонетичну структуру мовного сигналу для поліпшення фільтрації. 	<ul style="list-style-type: none"> - Складні в реалізації. - Висока обчислювальна складність. - Потребує точних моделей шуму і мовних сигналів.
Методи на основі слухового сприйняття	Покращення розбірливості, підвищення чіткості сигналу	Не завжди ефективні в умовах сильного шуму

2 АНАЛІЗ АЛГОРИТМУ ПРИГЛУШЕННЯ ШУМУ В МОВНОМУ СИГНАЛІ ЗА ДОПОМОГОЮ СПЕКТРАЛЬНОГО ВІДНІМАННЯ

2.1 Основи роботи алгоритму основоного на спектральному аналізі.

Алгоритм спектрального віднімання широко використовується як сам по собі, так і у складі складніших алгоритмів, у багатьох програмах – редакторах звуку:

- перетворення сигналу з часової в частотну область використовуючи дискретне перетворення Фур'є;
- оцінка діапазону шуму(профілю шуму);
- "віднімання" амплітудного спектра шуму з амплітудного спектра сигналу;
- зворотне перетворення Фур'є – синтез результуючого сигналу.

Дослідження якості та чіткості мови, що отримана в результаті застосування описаної методики, показали , що у випадках, коли шум або завада мають стаціонарний (або квазістаціонарний) характер і їх спектр має гармонійну структуру, досягається значне на слух підвищення як якості, так і розбірливості мови.

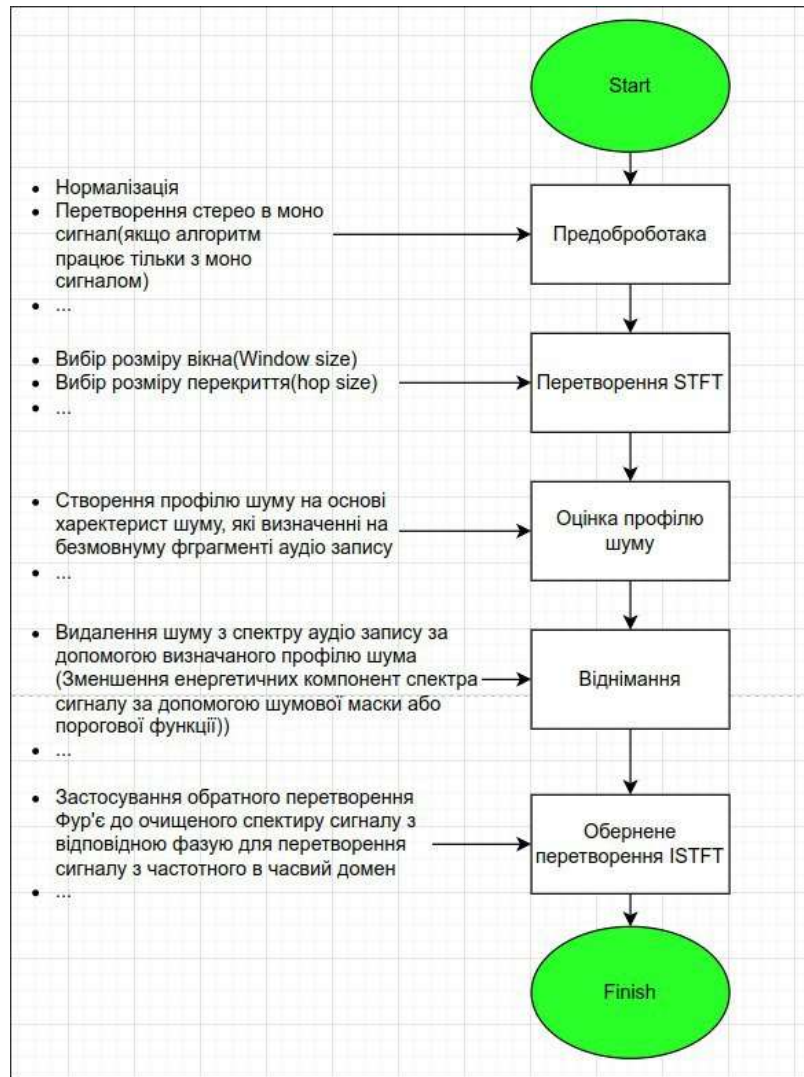


Рисунок 2.1 – Базова структура алгоритма спектального віднімання

2.2 Перетворення Фур'є

Швидке перетворення Фур'є (ШПФ) є одним з найважливіших алгоритмів обробки сигналів та аналізу даних. ШПФ — це швидкий алгоритм для обчислення дискретного перетворення Фур'є (DFT). DFT, як і більш знайома безперервна версія перетворення Фур'є, має пряму та зворотну форми, які визначаються таким чином:

Пряме дискретне перетворення Фур'є (DFT):

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-i2\pi kn}{N}}$$

Обернене дискретне перетворення Фур'є (IDFT):

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{\frac{i2\pi kn}{N}}.$$

Аналіз Фур'є заснований на ідеї, що будь-який часовий ряд можна розкласти на суму гармонійних хвиль різної частоти. Отже, теоретично можливо використовувати ряд гармонійних хвиль для створення будь-якого сигналу

Для відображення зв'язку між часом і частотою застосовується короткочасне перетворення Фур'є (КПФ). Ідея полягає у розбитті сигналу на короткі часові сегменти, для кожного з яких обчислюється спектр. Результатом такого аналізу є спектрограма — графічне зображення, де:

- вісь X відповідає часу,
- вісь Y — частоті,
- амплітуда спектра передається кольором або відтінками сірого.

Таким чином, спектрограма дозволяє спостерігати, як змінюється частотний склад сигналу з часом.

При застосуванні ДПФ важливо враховувати, що цей метод обробки сигналів передбачає їхню періодичність. Іншими словами, ДПФ інтерпретує обраний сегмент як один повний період нескінченного повторюваного сигналу. Однак на практиці таке припущення часто є неточним.

Найпоширенішою проблемою є розрив між початком і кінцем сегмента. Якщо ці точки не співпадають за значеннями та похідними, виникає штучний розрив, який у спектрі проявляється появою небажаних частотних компонентів. Це призводить до спотворення спектрального аналізу та додавання артефактів у результати.

Виток можна зменшити, згладжуючи розрив між початком і кінцем сегмента, і один із способів добитися цього - використання вікон.

Вікно - це функція, що перетворює аперіодичний сегмент у щось схоже на періодичне. На рис. 2.1 а) показано сегмент, де початок і кінець сигналу з'єднуються з розривом.

На рис. 2.1 б) показано вікно Хеммінга - одна з найбільш розповсюджених віконних функцій. Досконалих віконних функцій немає, але в тому чи іншому випадку деякі з них оптимальні.

На рис. 2.1 в) показано результат множення вікна на оригінальний сигнал. Там, де вікно близько до 1, сигнал не змінюється. Там, де вікно близько до 0, сигнал ослаблюється. Оскільки вікно мало на обох краях, кінець сегмента гладко стикується з початком.

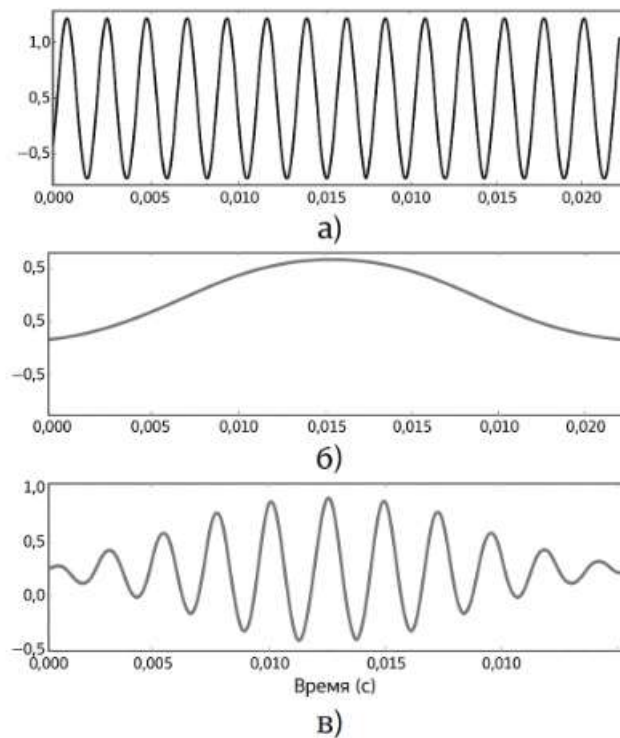


Рисунок 2.1 – а) Сегмент синусоїди, б) вікно Хеммінга та в) добуток сегмента та вікна

При обробці звуку у якості банк фільтрів використовують STFT з вікном Хана

$$w(t) = 1 - \cos \frac{2\pi(t+0,5)}{N}, 0 \leq t \leq N.$$

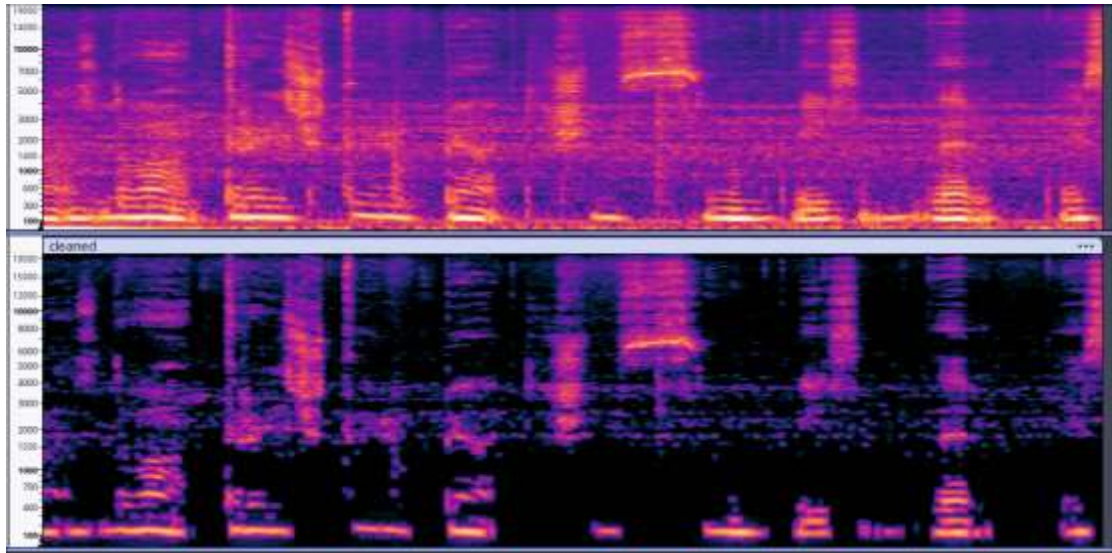


Рисунок 2.2 – Порівняння спектрограм зашумленого та очищеного аудіосигналів

2.3 Оцінка спектру шуму(Шумового профілю)

Оцінка спектру шуму є критично важливим етапом у багатьох алгоритмах обробки сигналів, особливо в задачах шумозаглушення та розпізнавання мови. Точна оцінка шумового профілю дозволяє ефективно відокремити корисний сигнал від небажаного шуму. Існує кілька підходів до оцінки спектру шуму:

Ручна оцінка. Користувач самостійно ідентифікує часовий сегмент сигналу, який містить лише шум і не містить корисного сигналу (наприклад, ділянки тиші перед початком мовлення). Шляхом аналізу спектра на часовому сегменті, який користувач ідентифікував як шум. Зліпок шуму отримуємо усередненням за часом амплітуд спектру, взятих з ділянки шуму, що не містить корисного сигналу:

$$foot(f) = \frac{\sum_{t=1}^K n(f,t)}{K}, \quad (2.1)$$

де $n(f,t)$ – спектр шуму;

f – індекс перетворення Фур'є, що відповідає частоті;

t – номер поточного STFT вікна;

K – кількість вікон на ділянці з шумом.

Автоматична оцінка. Для ситуацій зі змінним шумом, ручна оцінка профілю шуму є неефективною та непридатною. Постійні коливання шумового середовища вимагають динамічного підходу до його оцінки. З огляду на це, я зосереджуся на використанні автоматичних методів оцінки шуму.

Автоматичні методи прагнуть оцінити спектр шуму без втручання користувача, постійно аналізуючи вхідний сигнал. Існує кілька підходів до автоматичної оцінки:

Метод мінімальної статистики: Базується на припущенні, що в коротких часових інтервалах найменші значення енергетичних характеристик сигналу відповідають шуму. Під час обробки постійно відстежується мінімум спектральної амплітуди у кожній частотній смужі в певному часовому вікні.

Спектральне згладжування (Spectral Smoothing): Шумова оцінка формується шляхом згладжування спектральних коефіцієнтів у часі та/або частоті. Використовуються ковзні середні або експоненційні фільтри для приглушення короткочасних флуктуацій.

Метод виявлення немовних ділянок (Voice Activity Detection, VAD): Це саме той підхід, на якому я буду базуватися. Він передбачає, що спочатку використовується алгоритм VAD для визначення часових сегментів, які містять мовлення, і тих, які містять лише шум (ділянки "мовчання"). Алгоритм VAD аналізує вхідний сигнал і приймає рішення про наявність або відсутність мовлення в кожному часовому фреймі. Під час виявлення ділянок "мовчання" (коли VAD вважає, що мовлення відсутнє), спектр сигналу на цих ділянках використовується для оновлення поточної оцінки спектру шуму.

VAD-базова оцінка шуму передбачає:

- визначення мовних та шумових сегментів: Алгоритм VAD аналізує вхідний сигнал для ідентифікації часових фреймів, що містять мовлення, та тих, що містять виключно шум (так звані "ділянки мовчання");

- оновлення спектру шуму: На ділянках, де VAD виявляє лише шум, спектр сигналу використовується для оновлення поточної оцінки спектру шуму. Це дозволяє системі адаптуватися до змін у шумовому середовищі в режимі реального часу.

Такий підхід забезпечує більш точну та адаптивну оцінку шуму, що є критично важливим для ефективного шумозаглушення в умовах непередбачуваних змін у шумовому профілі.

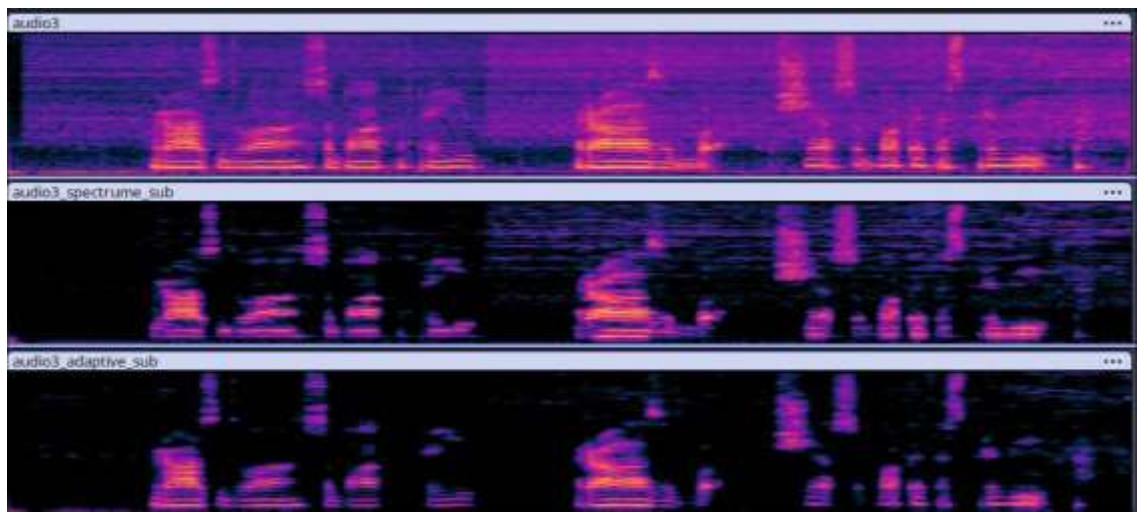


Рисунок 2.4 – Спектри зашумленого сигналу, очищеного методом спектрального віднімання та адаптивного спектрального віднімання

Віднімання амплітудних спектрів може здійснюватися за формулою

$$Y(f, t) = \max\{X(f, t) - kW(f, t), 0\} \quad (2.2)$$

де $X(f, t)$ и $W(f, t)$ – амплітудні спектри сигналу та шуму відповідно;

$Y(f, t)$ – амплітудний спектр результуючого очищеного сигналу;

k – коефіцієнт подавлення.

Коефіцієнт подавлення k залежить від відношення сигнал/шум на сегменті аналізу та має типові значення близькі до 0,7...0,95.

Фазовий спектр очищеного сигналу дорівнює фазовому спектру зашумленого сигналу, тому що через випадковість шуму немає алгоритму, що дозволяє «очистити» і фазовий спектр. Після такого віднімання очищеної спектрограми сигналу проводиться синтез результуючого сигналу.

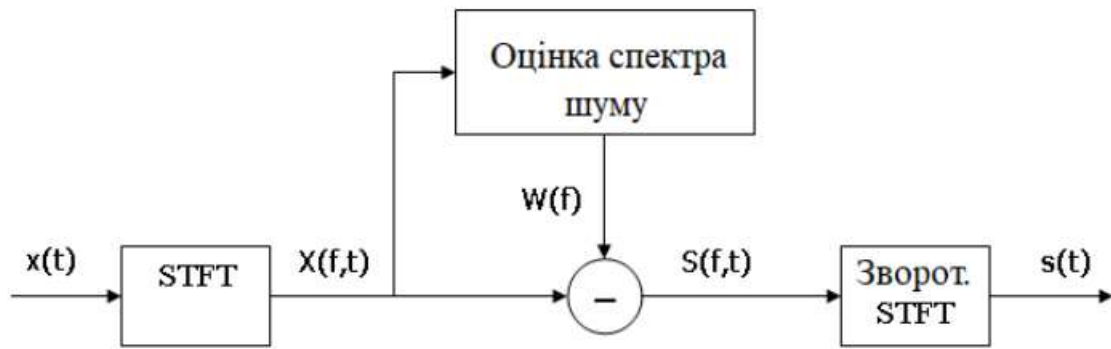


Рисунок 2.4 – Алгоритм віднімання амплітудних спектрів

Одна з проблем алгоритму спектрального віднімання – "музичний шум". Він утворюється внаслідок того, що коефіцієнти STFT шумових сигналів статистично випадкові. Це призводить до їхнього нерівномірного подавлення. В результаті, очищений сигнал містить короточасні та обмежені по частоті сплески енергії, які на слух сприймаються як "дзвіночки" або "вода, що ллється". Для подавлення артефакту як «музичного шуму» найчастіше застосовується алгоритм рекурсивного згладжування першого порядку за часом.

Існує безліч різних реалізацій алгоритму спектрального віднімання. Вони можуть відрізнятися типами передавальної характеристики у індивідуальних гейтів (правилом «віднімання» спектра шуму з спектра сигналу), числом частотних смуг, способами боротьби з різними артефактами.

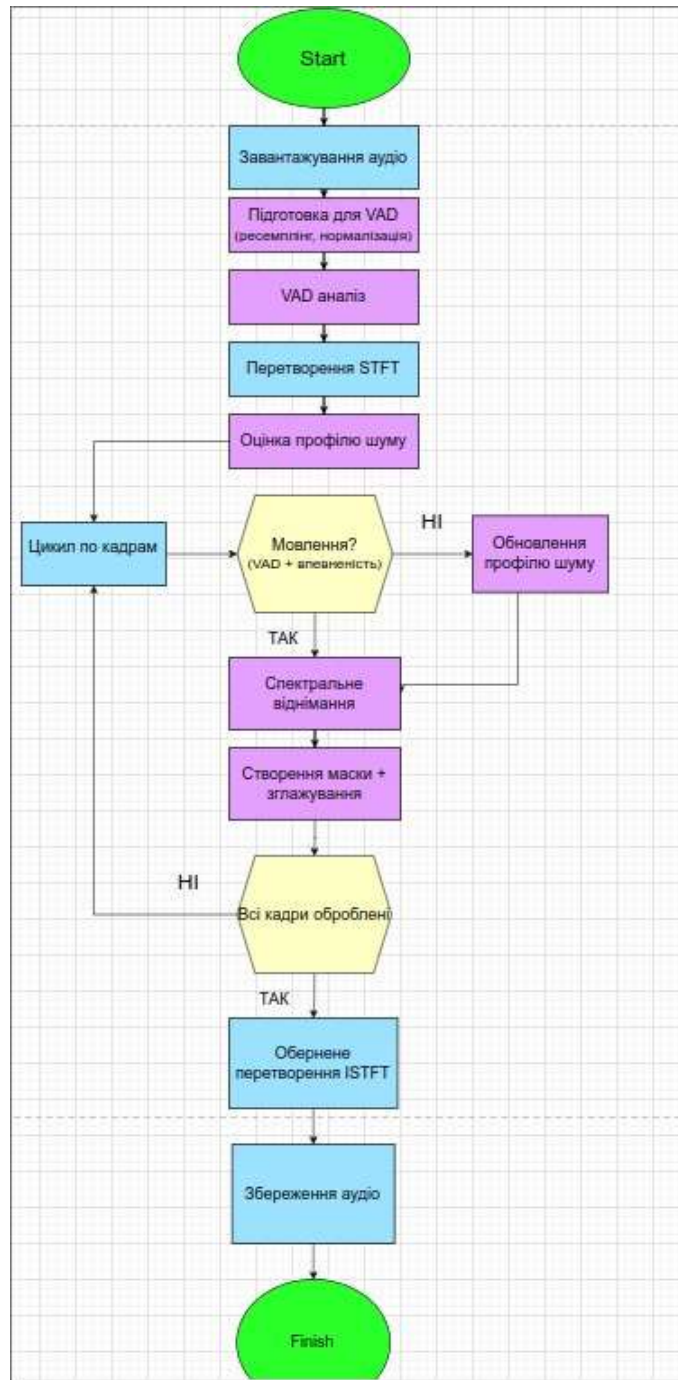


Рисунок 2.5 – Структура адаптивного алгоритма спектрального віднімання

3 РЕАЛІЗАЦІЯ АЛГОРИТМУ

Визначившись із метою розробки алгоритму спектрального віднімання для покращення якості аудіо, та маючи розуміння основних етапів його роботи, включаючи спектральний аналіз, оцінку шуму та спектральне віднімання, ми переходимо до детальної реалізації. У цьому розділі буде представлено розроблений алгоритм, що використовує бібліотеки `pumpy`, `librosa`, `soundfile`, `scipy` для обробки аудіосигналів та `webrtcvad` для виявлення голосової активності. Особливу увагу буде приділено реалізації адаптивного спектрального віднімання, оцінці профілю шуму на основі VAD, а також застосуванню м'якої маски для зменшення артефактів. Далі буде описано структуру розробленого класу `ImprovedSpectralSubtraction` та ключові методи, що забезпечують обробку аудіо.

3.1 Функціонал вводу вивода аудіо

```
def load_audio(self, file_path):
    """
    Гнучке завантаження аудіофайлу
    """
    try:
        audio, sr = librosa.load(file_path, sr=None, mono=True)
        print(f"Завантажено аудіо: {file_path}")
        print(f"Частота дискретизації: {sr} Hz")
        print(f"Тривалість: {len(audio)/sr:.2f} секунд")
        return audio, sr
    except Exception as e:
        print(f"Помилка завантаження файлу {file_path}: {e}")
        raise

def save_audio(self, audio, sr, output_path):
    """
    Збереження аудіофайлу
    """
    try:
        sf.write(output_path, audio, sr)
        print(f"Збережено оброблений аудіо: {output_path}")
    except Exception as e:
        print(f"Помилка збереження файлу {output_path}: {e}")
        raise
```

Рисунок 3.1 – Функції вводу вивода аудіо

Для забезпечення можливості завантаження вхідних аудіофайлів та збереження результатів обробки, клас `ImprovedSpectralSubtraction` містить два ключових методи: `load_audio` та `save_audio`. Ці методи інкапсують використання зовнішніх бібліотек `librosa` та `soundfile` для зручної роботи з аудіоданими.

Основне призначення методу `load_audio` полягає у зчитуванні аудіоданих з файлу, шлях до якого передається як аргумент `file_path`. Метод викликає функцію `librosa.load(file_path, sr=None, mono=True)`.

`file_path`: Вказує на розташування аудіофайлу у файловій системі. `librosa.load` підтримує широкий спектр аудіоформатів (наприклад, WAV, MP3, OGG).

`sr=None`: Інструктує `librosa` використовувати оригінальну частоту дискретизації, що міститься у файлі. Якщо вказати конкретне значення, аудіо буде автоматично ресемпльовано. У нашому випадку ми зберігаємо оригінальну частоту для подальшої обробки.

`mono=True`: Вказує на необхідність перетворення аудіо у моно, якщо воно є стерео або багатоканальним. Це спрощує подальшу обробку алгоритмом спектрального віднімання, який зазвичай застосовується до моноканальних сигналів.

Після успішного завантаження, метод виводить інформаційні повідомлення про завантажений файл, включаючи його шлях, частоту дискретизації (отриману з файлу) та тривалість аудіосигналу. Тривалість обчислюється як відношення кількості семплів до частоти дискретизації.

Метод повертає два об'єкта:

- `audio`: Масив , що містить амплітуди аудіосигналу у вигляді чисел з плаваючою комою (зазвичай в діапазоні від -1.0 до 1.0);
- `sr` : Ціле число, що представляє частоту дискретизації завантаженого аудіосигналу у герцах.

Метод `save_audio(self, audio, sr, output_path)`

Метод `save_audio` виконує зворотну операцію – записує оброблений аудіосигнал у файл за вказаним шляхом.

3.2 Підготовка аудіо до дитекції голоса

Наступний фрагмент коду представляє метод `prepare_audio_for_vad`, який є ключовим етапом підготовки аудіо для подальшого виявлення голосової активності (Voice Activity Detection). Оскільки наступним кроком буде використання вже готової та популярної бібліотеки WebRTC VAD від Google, аудіосигнал необхідно перетворити у відповідну форму, яку вимагає цей інструмент. WebRTC VAD відомий своєю компактністю та низькими вимогами до ресурсів. Метод спочатку перевіряє, чи частота дискретизації вхідного аудіо вже підтримується WebRTC VAD. У випадку невідповідності, виконується ресемплінг аудіосигналу до найближчої підтримуваної частоти за допомогою бібліотеки `librosa`. Далі, аудіосигнал обмежується діапазоном $[-1.0, 1.0]$ та конвертується до 16-бітного цілочисельного формату, оскільки саме такий формат часто очікує WebRTC VAD. Для подальшого узгодження часових міток, обчислюється коефіцієнт відношення між оригінальною та VAD-сумісною частотами дискретизації. Нарешті, метод повертає підготовлений аудіосигнал, його частоту дискретизації для VAD та обчислений коефіцієнт ресемплінгу, забезпечуючи сумісність з бібліотекою WebRTC VAD.

```
def prepare_audio_for_vad(self, audio_signal, sr):
    if sr in self.supported_sample_rates:
        vad_sr = sr
        vad_audio = audio_signal
    else:
        vad_sr = min(self.supported_sample_rates, key=lambda x: abs(x - sr))
        print(f"Ресемплінг з {sr} Hz до {vad_sr} Hz для VAD")
        vad_audio = librosa.resample(audio_signal, orig_sr=sr, target_sr=vad_sr)

    vad_audio = np.clip(vad_audio, -1.0, 1.0)
    vad_audio = (vad_audio * 32767).astype(np.int16)

    sr_ratio = sr / vad_sr
    return vad_audio, vad_sr, sr_ratio
```

Рисунок 3.2 – Функція підготовки аудіо до VAD

3.3 Дитекція голосу

Після попередньої обробки аудіосигнал подається на вхід функції `voice_activity_detection`, яка відповідає за виявлення голосової активності. Спочатку визначається розмір кадру VAD у семплах. Потім, підготовлений аудіосигнал подається на алгоритм VAD, який ітерується по його фреймах. На виході цього процесу формується список `vad_results`, що містить 0 (якщо у фреймі не виявлено голосу) або 1 (якщо голос присутній). Для згладжування результатів VAD застосовується медіанний фільтр. Насамкінець, результати VAD зіставляються з часовими рамками STFT оригінального аудіо, створюючи масив `vad_labels`, де кожному STFT-кадру відповідає мітка голосової активності

```
def voice_activity_detection(self, audio_signal, sr):
    """
    Поширене виявлення голосової активності
    """
    vad_audio, vad_sr, sr_ratio = self.prepare_audio_for_vad(audio_signal, sr)
    vad_frame_samples = int(vad_sr * self.vad_frame_duration / 1000)

    if len(vad_audio) < vad_frame_samples:
        print(f"Аудіо занадто коротке для VAD")
        stft_frames = int(np.ceil(len(audio_signal) / self.hop_length))
        return np.zeros(stft_frames, dtype=int)

    vad_results = []
    vad_times = []

    for i in range(0, len(vad_audio) - vad_frame_samples + 1, vad_frame_samples):
        frame = vad_audio[i:i + vad_frame_samples]
        try:
            is_speech = self.vad.is_speech(frame.tobytes(), sample_rate=vad_sr)
            vad_results.append(int(is_speech))
            center_time = (i + vad_frame_samples // 2) * sr_ratio / sr
            vad_times.append(center_time)
        except Exception as e:
            print(f"Помилка VAD: {e}")
            vad_results.append(0)
            vad_times.append((i + vad_frame_samples // 2) * sr_ratio / sr)

    if len(vad_results) > 3:
        vad_results = medfilt(np.array(vad_results), kernel_size=3)

    stft_frames = int(np.ceil(len(audio_signal) / self.hop_length))
    vad_labels = np.zeros(stft_frames, dtype=int)

    for stft_idx in range(stft_frames):
        stft_time = (stft_idx * self.hop_length + self.frame_length // 2) / sr
        if len(vad_times) > 0:
            closest_vad_idx = np.argmin(np.abs(np.array(vad_times) - stft_time))
            vad_labels[stft_idx] = vad_results[closest_vad_idx]

    return vad_labels
```

Рисунок 3.3 – Функція дитекції голосу

3.4 Алгоритм початкової оцінки профілю шуму та алгоритм оновлення профілю шуму

Для спектрального віднімання необхідно мати оцінку профілю шуму. У представленому коді реалізовано дві функції для цієї мети. Перша функція, `estimate_noise_spectrum`, визначає первинний профіль шуму, аналізуючи спектральні величини початкових кадрів аудіо, які вважаються переважно шумовими, та усереднюючи їх за формулою 2.1 .

Друга функція, `update_noise_spectrum`, відповідає за подальше оновлення профілю шуму протягом обробки аудіо. Вона використовує результати VAD для ідентифікації кадрів, де не було виявлено мовлення. Спектральні величини цих "тихих" кадрів накопичуються, і періодично, коли набирається достатньо шумових кадрів, профіль шуму перераховується, адаптуючись до змін шумового фону. Таким чином, алгоритм підтримує актуальну оцінку спектру шуму для більш якісного шумозаглушення.

```
def estimate_noise_spectrum(self, noisy_stft, sr, initial_duration=0.5):
    initial_frames = int(initial_duration * sr / self.hop_length)
    initial_frames = min(initial_frames, noisy_stft.shape[1])

    initial_magnitude = np.abs(noisy_stft[:, :initial_frames])
    noise_spectrum = np.mean(initial_magnitude, axis=1, keepdims=True)

    return noise_spectrum

def update_noise_spectrum(self, noisy_stft, sr, vad_labels, noise_frames_buffer):
    noise_duration_frames = int(0.2 * sr / self.hop_length)
    noise_spectrum = self.estimate_noise_spectrum(noisy_stft, sr, initial_duration=0.5)

    if not noise_frames_buffer:
        noise_frames_buffer = []

    for frame_idx in range(noisy_stft.shape[1]):
        is_speech = (frame_idx < len(vad_labels)) and (vad_labels[frame_idx] == 1)

        if not is_speech:
            noise_frames_buffer.append(np.abs(noisy_stft[:, frame_idx:frame_idx+1]))

            if len(noise_frames_buffer) >= noise_duration_frames:
                buffer_array = np.concatenate(noise_frames_buffer, axis=1)
                noise_spectrum = np.mean(buffer_array, axis=1, keepdims=True)
                noise_frames_buffer = []
            else:
                noise_frames_buffer = []

    return noise_spectrum, noise_frames_buffer
```

Рисунок 3.4 – Функції оцінки початкового шуму та оновлення профілю шуму

3.5 Спектральне віднімання

Після етапів оцінки та оновлення профілю шуму, наступним ключовим кроком є безпосередньо спектральне віднімання, яке реалізовано в методі `spectral_subtraction_adaptive`. Ця функція приймає спектральну величину зашумленого сигналу, поточну оцінку спектру шуму. У першу чергу, обчислюються енергії (квадрати спектральних величин) як шумного сигналу, так і оціненого шуму. Потім відбувається процес спектрального віднімання формула 2.2, де з енергії шумного сигналу віднімається енергія шуму, помножена на коефіцієнт перевіднімання α (alpha). Для запобігання виникненню небажаних артефактів, таких як "музичний шум", вводиться поняття спектральної підлоги, яка визначається як добуток коефіцієнта β (beta) на енергію шумного сигналу. Отримана після віднімання енергія спектрально-покращеного сигналу обмежується знизу значенням спектральної підлоги. На завершення, з отриманої покращеної енергії видобувається амплітуда спектру шляхом взяття квадратного кореня, яка і повертається як результат обробки поточного спектрального кадру.

```
def spectral_subtraction_adaptive(self, noisy_magnitude, noise_spectrum, is_speech):
    noise_power = noise_spectrum ** 2
    noisy_power = noisy_magnitude ** 2

    enhanced_power = noisy_power - self.alpha * noise_power
    spectral_floor = self.beta * noisy_power
    enhanced_power = np.maximum(enhanced_power, spectral_floor)

    enhanced_magnitude = np.sqrt(enhanced_power)

    return enhanced_magnitude
```

Рисунок 3.5 – Реалізація спектрального віднімання

3.6 Основна функція, що керування алгоритмом обробку аудіо

Продовжуючи опис алгоритму, ми підійшли до головної функції обробки аудіо - `process_audio`. На початку вона завантажує аудіофайл за допомогою

load_audio (описаної в розділі 3.1). Потім викликається voice_activity_detection (описаного в розділі 3.3) для виявлення мовних фрагментів. Після цього обчислюється спектрограма шумного аудіо за допомогою STFT.

Оцінка початкового профілю шуму здійснюється через estimate_noise_spectrum(описаного в розділі 3.4) . Далі відбувається ітерація по кожному кадру спектрограми, де для кожного кадру оновлюється профіль шуму за допомогою update_noise_spectrum та застосовується спектральне віднімання функцією spectral_subtraction_adaptive. На завершення, оброблена спектрограма перетворюється назад у часову область за допомогою inverse STFT, і результат зберігається у вихідний файл методом self.save_audio. Таким чином, process_audio координує всі кроки від завантаження аудіо до збереження очищеного від шуму результату.

```
def process_audio(self, audio_input_file, audio_output_file, noise_duration=1.0):
    print("Завантаження аудіо...")
    audio_signal, sr = self.load_audio(audio_input_file)

    print("Виявлення мовлення за допомогою VAD...")
    vad_labels = self.voice_activity_detection(audio_signal, sr)

    speech_ratio = np.mean(vad_labels) * 100
    print(f"Виявлено {speech_ratio:.1f}% кадрів з мовленням")

    print("Обчислення STFT...")
    noisy_stft = librosa.stft(audio_signal,
                              n_fft=self.frame_length,
                              hop_length=self.hop_length,
                              window=self.window_type)

    noisy_magnitude = np.abs(noisy_stft)
    noisy_phase = np.angle(noisy_stft)

    print("Оцінка профілю шуму...")
    noise_spectrum = self.estimate_noise_spectrum(noisy_stft, sr, noise_duration)
    noise_frames_buffer = []

    print("Спектральне віднімання...")
    enhanced_magnitude = np.zeros_like(noisy_magnitude)

    for frame_idx in range(noisy_magnitude.shape[1]):
        is_speech = (frame_idx < len(vad_labels)) and (vad_labels[frame_idx] == 1)

        # Update noise spectrum
        noise_spectrum, noise_frames_buffer = self.update_noise_spectrum(
            noisy_stft, sr, vad_labels, noise_frames_buffer)

        enhanced_magnitude[:, frame_idx] = self.spectral_subtraction_adaptive(
            noisy_magnitude[:, frame_idx:frame_idx+1],
            noise_spectrum,
            is_speech
        ).flatten()

    print("Реконструкція аудіо...")
    enhanced_stft = enhanced_magnitude * np.exp(1j * noisy_phase)
    enhanced_audio = librosa.istft(enhanced_stft,
                                   hop_length=self.hop_length,
                                   window=self.window_type,
                                   length=len(audio_signal))

    self.save_audio(enhanced_audio, sr, audio_output_file)

    return enhanced_audio, vad_labels
```

Рисунок 3.6 – Релізації функції входу

ВИСНОВКИ

У даній роботі було розроблено та реалізовано покращений алгоритм спектрального віднімання, спрямований на підвищення якості та розбірливості мовних сигналів, що є важливим завданням у багатьох сферах, включаючи комунікації та обробку мультимедіа. В основу алгоритму покладено спектральний аналіз із застосуванням короточасного перетворення Фур'є (STFT), що дозволяє розкласти часовий сигнал на частотні компоненти для подальшої обробки.

Для більш ефективного шумозаглушення було інтегровано детектор голосової активності WebRTC VAD, який допомагає автоматично ідентифікувати ділянки сигналу, що містять лише шум. Це дозволило реалізувати адаптивну оцінку спектру шуму, що є критично важливим для успішного застосування методів спектрального віднімання, особливо у випадках нестационарних шумів.

Розроблений алгоритм використовує принцип віднімання амплітудних спектрів, де оцінений спектр шуму вилучається зі спектра зашумленого мовного сигналу. Для синтезу очищеного сигналу застосовується обернене перетворення Фур'є. Особливу увагу було приділено методам боротьби з артефактами, такими як "музичний шум", що часто виникають при спектральному відніманні.

Результати дослідження показали, що застосування VAD для керування оновленням шумового профілю сприяє кращому пригніченню шумів порівняно зі стандартними підходами. Подальші напрямки роботи можуть включати дослідження більш просунутих методів оцінки шуму, оптимізацію параметрів алгоритму та порівняльний аналіз з іншими сучасними методами шумозаглушення, включаючи методи, що використовують статистичні моделі мовних сигналів та принципи слухового сприйняття людини.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. URL: <https://sudonull.com/post/101170-Analysis-of-audio-analytics-algorithms>
[Електронний ресурс. Режим доступу 25.04.2025]
2. URL: https://habr.com/ru/companies/ru_mts/articles/584308/ [Електронний ресурс. Режим доступу 25.04.2025]
3. URL: <https://www.iks.rwth-aachen.de/en/research/audio/active-noise-control/>
[Електронний ресурс. Режим доступу 25.05.2025]
4. Цифрове оброблення сигналів: Посібник для студентів напряму підготовки 6.050901 "Радіотехніка" усіх форм навчання [Електронний ресурс] / Авт.-укл. С. В.Заболотній ; За ред. проф. Ю. Г. Леги ; М-во освіти і науки України, Черкас. держ. технол. ун-т. – Черкаси: ЧДТУ, 2010. – 119 с.
5. URL: <https://medium.com/@theclickreader/webrtc-voice-activity-detection-using-python-the-click-reader-9ee3797adbea> [Електронний ресурс. Режим доступу 25.05.2025]
6. URL: <https://medium.com/@joshi prerak123/transform-your-audio-denoise-and-enhance-sound-quality-with-python-using-pedalboard-24da7c1df042> [Електронний ресурс. Режим доступу 22.05.2025]
7. URL: https://speechprocessingbook.aalto.fi/Recognition/Voice_activity_detection.htm [Електронний ресурс. Режим доступу 25.05.2025]
8. URL: <https://inria.hal.science/hal-01881425v1/document> [Електронний ресурс. Режим доступу 24.05.2025]
9. URL: <https://www.degruyterbrill.com/document/doi/10.1515/jisys-2022-0024/html>
[Електронний ресурс. Режим доступу 24.05.2025]
10. URL: <https://www.scirp.org/journal/paperinformation?paperid=45989>
[Електронний ресурс. Режим доступу 24.05.2025]