

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: Метод та алгоритм підвищення точності розпізнавання звукових
сигналів у закритому середовищі

Виконав: студент VI курсу, групи РРМ-61

спеціальності

172 Телекомунікації та
радіотехніка

(шифр і назва спеціальності)

(підпис)

Сторож С.І.

(прізвище та ініціали)

Керівник

(підпис)

Дунець В.Л.

(прізвище та ініціали)

Нормоконтроль

(підпис)

Хвостівська Л.В.

(прізвище та ініціали)

Завідувач кафедри

(підпис)

Дунець В.Л.

(прізвище та ініціали)

Рецензент

(підпис)

(прізвище та ініціали)

Тернопіль
2023

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)

«__» _____ 2023 р.

**ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ**

на здобуття освітнього ступеня Магістр
(назва освітнього ступеня)

за спеціальністю 172 Телекомунікації та радіотехніка
(шифр і назва спеціальності)

Студенту Сторожу Сергію Ігоровичу
(прізвище, ім'я, по батькові)

1. Тема роботи Метод та алгоритм підвищення точності розпізнавання звукових сигналів у закритому середовищі

Керівник роботи Дунець Василь Любомирович, к.т.н., завідувач кафедри РТ
(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

Затверджені наказом ректора від «__» _____ 2023 року № _____

2. Термін подання студентом завершеної роботи _____

3. Вихідні дані до роботи _____

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

Вступ.

Розділ 1. Аналітична частина

Розділ 2. Основна частина

Розділ 3. Науково-дослідна частина

Розділ 5. Охорона праці та безпека в надзвичайних ситуаціях

Висновки

Перелік використаних джерел

Додатки

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

1. Титульний аркуш. 2. Актуальність теми. 3. Мета і задачі дослідження. 4. Математична модель звукових сигналів у закритому середовищі. 5. Штучна генерація реверберації.

6. Імпульсна характеристики ревербаторів Мура та реального сигналу. 7. Структура генератора мережі зворотніх ліній затримок. 8. Імпульсна характеристики ревербатора мережі зворотніх ліній затримок 4 порядку. 9. Метод підвищення точності розпізнавання звукових сигналів у закритому середовищі. 10. Моделювання роботи алгоритму розпізнавання звукових сигналів у закритому середовищі. 11. Наукова новизна отриманих результатів. 12. Висновки

6. Консультанти розділів роботи

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	проректор з адміністративно-господарської роботи та будівництва Клепчик В.М		
Безпека в надзвичайних ситуаціях	ст. викл. кафедри ОХ, Стручок В.С.		

7. Дата видачі завдання _____

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Ознайомлення з завданням до кваліфікаційної роботи		<i>Виконано</i>
2.	Підбір джерел по темі кваліфікаційної роботи		<i>Виконано</i>
3.	Переклад та опрацювання джерел по темі кваліфікаційної роботи		<i>Виконано</i>
4.	Оформлення розділу «Аналітична частина»		<i>Виконано</i>
5.	Оформлення розділу «Основна частина»		<i>Виконано</i>
6.	Оформлення розділу «Науково-дослідна частина»		<i>Виконано</i>
7.	Оформлення розділу «Охорона праці та безпека в надзвичайних ситуаціях»		<i>Виконано</i>
8.	Підготовка графічного матеріалу		<i>Виконано</i>
9.	Оформлення кваліфікаційної роботи		<i>Виконано</i>
10.	Нормоконтроль		<i>Виконано</i>
11.	Перевірка на плагіат		<i>Виконано</i>
12.	Попередній захист кваліфікаційної роботи		<i>Виконано</i>
13.	Захист кваліфікаційної роботи		

Студент _____
(підпис)

Сторож С.І.
_____ (прізвище та ініціали)

Керівник роботи _____
(підпис)

Дунець В.Л.
_____ (прізвище та ініціали)

АНОТАЦІЯ

Тема кваліфікаційної роботи: Метод та алгоритм підвищення точності розпізнавання звукових сигналів у закритому середовищі // Кваліфікаційна робота освітнього рівня «Магістр» // Сторож Сергій Ігорович // ТНТУ, ФПТ, група РРм-61 // Тернопіль, 2023 // С. ____, рис. – ____, табл. – ____, кресл. – ____, додат. – ____, бібліогр. – ____.

Ключові слова: звукові сигнали, точність, розпізнавання, реверберація, середовище, програмне забезпечення.

У рамках даної кваліфікаційної роботи вдосконалено підходи до покращення точності виявлення акустичних сигналів у ЗС, а також метод моделювання реверберації у ЗС і створення неприродної реверберації. Ці техніки ґрунтуються на застосуванні МСВ, використовуючи модель, яка базується на концепції Полака і методі Дебніха.

Вдосконалений підхід для покращення точності розпізнавання мовних сигналів реалізовано у формі програмного забезпечення.

У першому розділі:

- Розглянуто труднощі пов'язані із визначенням акустичних сигналів у ЗС.
- Досліджено прямиий звук та компоненти реверберації.
- Проаналізовано методи вимірювання часу реверберації.
- Розглянуто типи оцінки критичної відстані і акустичні особливості ЗС залежно від частотного діапазону.

У другому розділі кваліфікаційної роботи:

- Представлено конфігурації існуючих генераторів акустичних сигналів, обладнаних ефектами реверберації.
- Надано висновки моделювання імпульсних характеристик з метою порівняння із фактичними параметрами ЗС.

У третьому розділі кваліфікаційної роботи на основі аналізу вже існуючих методів для подолання ефектів реверберації та вдосконалення точності виявлення акустичних сигналів у ЗС, визначено, що найоптимальнішим для застосування є МСВ.

Проведено аналіз підходів, які дозволяють оцінити спектр пізньої реверберації для МСВ. Враховуючи переваги та недоліки, для застосування у дослідженні був обраний метод Дебніха (МСВ).

Цей підхід застосовує час реверберації для виконання розрахунків, які сприяють підвищенню точності вимірювань відповідно до визначених параметрів ЗС. Застосування даного підходу дозволяє більш детально дослідити ефективність алгоритму, що базується на цьому підході. Застосування генераторів штучної реверберації дає можливість встановлювати час реверберації та досліджувати результати фільтрації.

Надалі виконана процедура моделювання функціонування алгоритму, як на синтетичних, так і на дійсних аудіосигналах. В результаті цього моделювання визначено, що впроваджений спосіб ефективно долає пізню реверберацію, залишаючи певні залишкові ефекти у фільтрованій реалізації сигналу ранньої реверберації. Проаналізовано ефективність алгоритму залежно від змін входних параметрів, які виявляються значущими для явища реверберації. Це дозволило здійснити оцінку функціонування алгоритму в різноманітних умовах.

ANNOTATION

Method and algorithm for improving the accuracy of sound signal recognition in a closed environment // Qualification work of the educational level «Master» // Storozh Serhii Ihorovych // TNTU, FPT, group RRm-61 // Ternopil, 2023 // P. ____, fig. – ____, tabl. – ____, chair. – ____, annexes. – ____, references – ____.

Keywords: sound signals, quality, recognition, reverberation, premises, software.

This qualification work develops approaches to improve the accuracy of sound signal recognition in an enclosed environment, as well as methods for modeling reverberation in the environment and creating artificial reverberation. These methods are based on the application of spectral subtraction, using a model based on the concept of Polak and the Debnich method.

The developed approach to improving the accuracy of sound signal recognition was implemented in the form of a computer program.

The first chapter of the qualification work:

- The difficulties associated with the determination of sound signals in a closed environment are considered.
- The study of direct sound and reverberation components is described.
- The methods of measuring the reverberation time are analyzed.
- Methods for estimating the critical distance and acoustic features of the environment depending on the frequency range are considered.

In the second chapter of the qualification work:

- Configurations of existing sound signal generators equipped with reverberation effects are presented.
- The analysis of their structures is carried out.
- The results of modeling their impulse characteristics are presented in order to compare them with real environmental parameters.

In the third chapter of the qualification work, based on the analysis of existing methods for overcoming reverberation effects and improving the accuracy of audio signal recognition in a closed environment, it was concluded that the most optimal method for use is the spectral subtraction method.

An analysis of approaches that make it possible to estimate the late reverberation spectrum for the spectral subtraction algorithm was carried out. Taking into account the advantages and disadvantages, the Debnich method (spectral subtraction filter) was chosen for use in the work. This method uses the reverberation time to perform calculations, which allows for improved measurement accuracy in accordance with the specified environmental parameters. Using this approach allows us to analyze the effectiveness of the algorithm based on this method in more detail. Thanks to the use of artificial reverberation generators, we can set the reverberation time and study the filtering results. Thus, in the future, it is planned to apply the spectral subtraction method using the Debnich method for estimating the late reverberation spectrum.

In the qualification work, the algorithm was modeled on both synthetic and real audio signals. Based on the results of this simulation, it was found that the implemented algorithm successfully eliminates late reverberation with subsequent residuals in the filtered realization of the early reverberation signal. In addition, we analyzed the effectiveness of the algorithm depending on changes in input parameters that are important for the reverberation phenomenon. This made it possible to evaluate the algorithm's functioning in various conditions.

ПЕРЕЛІК СКОРОЧЕНЬ І ТЕРМІНІВ

BEM (англ. Boundary Element Method) – Метод граничних елементів.

EDC (англ. Energy Decay Curve) – Крива затухання енергії.

FEM (англ. Finite Element Method) – Метод скінченних елементів.

АКФ – Функція автокореляції.

ЗС – Закрите середовище.

ІХС – Імпульсна характеристика середовища.

ЛП – Лінійне передбачення.

ЛПС – Кодування з лінійним предиктором. Спосіб перетворення аналогового сигналу в цифрову форму, який використовується в основному в обробці аудіосигналу (мови) для представлення спектральної огинаючої цифрового сигналу мовлення в стислому вигляді, з використанням інформації лінійної прогностичної моделі.

МСВ – Метод спектрального віднімання.

СКВ – Середнє квадратичне відхилення.

СЩП – Спектральна щільність потужності.

ЗМІСТ

ВСТУП	9
РОЗДІЛ 1. АНАЛІТИЧНА ЧАСТИНА	11
1.1 Прямий звук та компоненти реверберації	11
1.2 Час реверберації	13
1.3 Критична відстань.....	16
1.4 Аналіз акустики середовища в залежності від діапазону частот.....	16
1.5 Висновок до розділу 1	18
РОЗДІЛ 2. ОСНОВНА ЧАСТИНА.....	19
2.1 Хвильове моделювання	19
2.2 Траєкторія променів	20
2.3 Метод уявних джерел	21
2.4 Статистична модель середовища	23
2.5 Вибір методу моделювання	26
2.6 Штучна генерація реверберації	27
2.6.1 Структури на базі всепропускаючих та гребінчастих фільтрів ..	28
2.6.2 Мережі зворотніх ліній затримок.....	30
2.6.3 Порівняння отриманих результатів.....	32
2.7 Метод підвищення точності розпізнавання звукових сигналів у закритому середовищі	324
2.7.1. Приглушення реверберації	325
2.7.2. Спектральне віднімання.....	326
2.7.3. Спектральне віднімання.....	328
2.7.4. Просторова обробка.....	39
2.7.5. Гомоморфна деконволюція.....	40
2.7.6. Метод спектрального віднімання для подолання реверберації звукового сигналу	42
2.8. Висновок до розділу 2	53

РОЗДІЛ 3. НАУКОВО-ДОСЛІДНА ЧАСТИНА.....	54
3.1. МОДЕЛЮВАННЯ ГАРМОНІЧНОГО СИГНАЛУ	54
3.2.Моделювання з використанням звучання акорду на гітарі	56
3.3.Моделювання мовного сигналу.....	57
3.4.Аналіз роботи алгоритму в залежності від вхідних параметрів	58
3.5.Висновок до розділу 4	61
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ	62
4.1.Питання щодо охорони праці	62
4.2.Питання щодо безпеки в надзвичайних ситуаціях	64
4.3.Висновок до розділу 4	70
ВИСНОВКИ.....	71
ПЕРЕЛІК ДЖЕРЕЛ	72

Вступ

Актуальність теми. Під час аналізу акустичних сигналів у ЗС виникає трудність, пов'язана із паразитними адитивними ефектами, що проявляються у вигляді суми залишкового післязвучання та корисного прямого сигналу. Ці паразитні ефекти формуються через множинне відбиття звуку від поверхонь та одночасне поглиблювання звукових хвиль [1].

Для мінімізації негативних впливів, таких як ранні відбиття та кінцеве післязвучання, використовуються методи, що ґрунтуються на принципах ЛП та зворотної згортки, використовуючи заздалегідь відомі дані про параметри ЗС та просторовий розподіл.

Однак існуючі методи проявляють обмежену ефективність у прибиранні паразитарних впливів, оскільки їхнє пригнічення супроводжується утисканням корисної частини акустичного сигналу, що стає неприйнятним у даному контексті [2]. Результати аналізу наукових досліджень у сфері визначення звукових сигналів у ЗС підтверджують, що корисна інформація також знаходиться в ранніх відбиттях при низькій потужності безпосереднього сигналу.

Тому розробка методу покращення точності розпізнавання звукових сигналів у ЗС, шляхом ефективного приглушення їхньої реверберації, є сучасною науковою проблемою.

Мета і задачі дослідження. Метою даної кваліфікаційної роботи освітнього рівня «Магістр» є впровадження методу і алгоритму для покращення точності виявлення звукових сигналів у ЗС.

З метою досягнення мети необхідно виконати наступні завдання:

- Проаналізувати фізичні характеристики, параметри аудіо-сигналів і методи виявлення акустичних сигналів у ЗС.

- Розробити підхід для моделювання утворення реверберації в ЗС і створити алгоритм для фальшивої генерації реверберації із зазначеним періодом часу.

- Створити підхід для покращення точності виявлення акустичних сигналів у ЗС.

- Оцінити можливості запропонованого методу в плані поліпшення точності виявлення акустичних сигналів у ЗС.

- Виготовити ПЗ для здійснення досліджень процесу визначення акустичних сигналів у ЗС.

Об'єкт дослідження. Основним об'єктом є аудіо-сигнал із впливом реверберації.

Предмет дослідження: алгоритми і методи покращення точності виявлення акустичних сигналів у ЗС.

Наукова новизна отриманих результатів. Розроблено алгоритм та МСВ, ґрунтуючись на статистичному моделюванні ревербераційного сигналу за допомогою моделі Полака, методу Дебніха. Це призвело до покращення точності виявлення звукових сигналів у ЗС через ліквідацію затримки реверберації.

Практичне значення роботи полягає в розробці методу і алгоритму, спрямованих на покращення точності виявлення акустичних сигналів у ЗС, котрі в певному відношенні перевершують недоліки визначних методів та алгоритмів.

РОЗДІЛ 1

АНАЛІТИЧНА ЧАСТИНА

1.1 Прямий звук та компоненти реверберації

Реверберація представляє собою явище, яке може виникати при поширенні акустичних хвиль у ЗС. В разі неочікуваного відключення джерела звуку спостерігається явище залишкового звуку, проілюстровані. На рисунку 1.1 наведено можливі напрямки поширення звукових хвиль.

Ефект реверберації є бажаним у музичному контексті, але виявляється небажаним у випадку мовлення. Для ефективного розпізнавання мови важливо, щоб час реверберації був менший за дві секунди [3]. Оскільки розміри та інші фізичні характеристики ЗС визначають його ревербераційні властивості, виникають протилежні вимоги до рівня реверберації в таких середовищах. У випадку, коли людина виступає перед аудиторією, реверберація повинна бути елімінована або зведена до мінімуму. З іншого боку, коли розпочинається оркестрове виконання, реверберація повинна бути відновлена.

Рисунок 1.1 відтворює процес утворення реверберації шляхом відбиття хвильових фронтів, які поширюються від джерела. Ці хвильові фронти відбиваються від стін і накладаються на мікрофон. На рисунку 1.1 цей процес ілюструється за допомогою прямого проходження хвилі та трьох відбитих її копій. Кожен із цих хвильових фронтів досягає мікрофона з відмінною амплітудою і фазою. Це зумовлено різною довжиною шляхів поширення хвильових фронтів до мікрофона і різною кількістю енергії, яку поглинають стіни.

Отриманий сигнал складається з прямого звуку, ранніх відбиттів, що надходять негайно після прямого звуку (рання реверберація/відбиття), та подальших відбиттів, що приходять негайно після ранньої реверберації (пізня реверберація/відбиття).

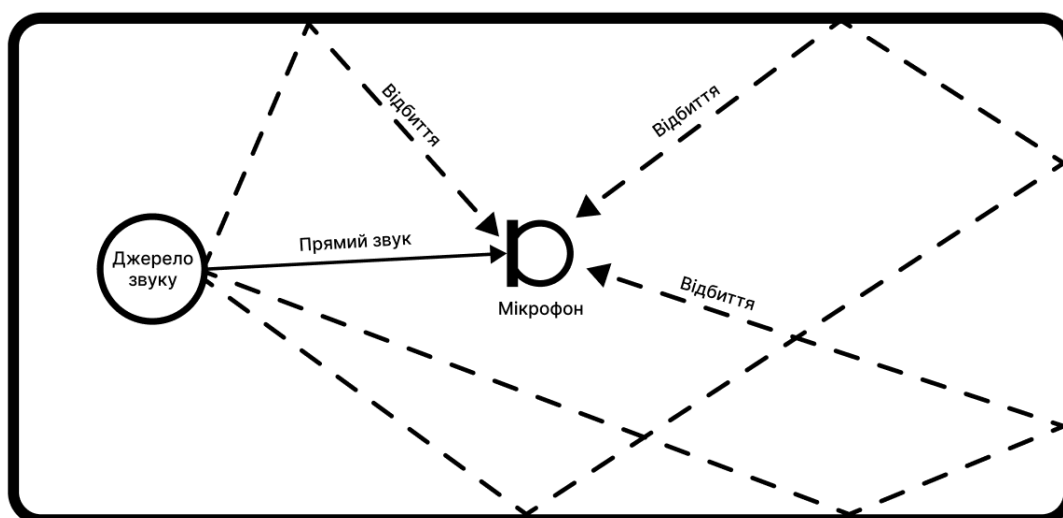


Рис. 1.1 Реверберація в ЗС

Рання реверберація є складовою частиною відбиттів, які надходять протягом невеликого часового інтервалу після приходу прямого звукового фронту. Ці компоненти досягають мікрофона з різних напрямків та в різний час в порівнянні з прямим звуком, при цьому мають меншу амплітуду. До того часу, поки затримка відбиттів не перевищить приблизно 80-100 мс в порівнянні з часом приходу прямого звуку, рання реверберація не розпізнається як самостійний звук. Використання ранньої реверберації полягає в підсиленні прямого звуку, і, отже, вона корисна для алгоритмів розпізнавання мови [4]. Рання реверберація сприяє кращому розпізнаванню мови в ЗС у порівнянні з відкритими. Особливо важливою є рання реверберація для акустики невеликих середовищ, де стіни, стеля і підлога розташовані на відносно короткій відстані одна від одної. З іншого боку, ранні відбиття викликають спотворення спектра прийнятого сигналу, відоме як забарвлення. Цей ефект виникає внаслідок короточасних кореляцій, введених сигналом внаслідок ранніх відбиттів. В результаті більшість алгоритмів усунення реверберації розглядають приглушення як ранніх, так і пізніх відбиттів. Важливо також відзначити, що алгоритми усунення реверберації були розроблені з урахуванням різноманітних застосувань, включаючи автоматичне розпізнавання мови, де пізні відбиття не розглядаються як корисні.

Пізня реверберація визначається як та частина ревербераційного ефекту, що виникає внаслідок затримок у відбиттях, які надходять після приходу прямого звуку. Людське сприйняття цих затриманих відбиттів виражається у вигляді окремих ехо-сигналів, що призводить до погіршення ясності мовлення.

1.2 Час реверберації

Широко використовуваною та важливою характеристикою акустичного простору є час реверберації [5]. Вимір часу реверберації можна здійснити шляхом введення широкосмугового сигналу в середовище до досягнення стійкого стану рівномірного розподілу звукової енергії. Після цього слід вимкнути джерело звуку і фіксувати EDC, що відома як EDC, відображаючи затухання квадрату звукового тиску у залежності від часу. Час реверберації, позначений як T_{60} , визначається для розсіяного звукового поля як період часу в секундах, необхідний для зменшення EDC на 60 дБ.

Ця концепція базується на досліджах Сабіна [6], який встановив, що період реверберації пропорційний об'єму середовища, V , та обернено пропорційний коефіцієнту поглиблення поверхонь у середовищі. Незважаючи на вплив ослаблення через поширення в повітрі, за методом Сабіна можна отримати оцінку часу реверберації:

$$T_{60} = \frac{24 \ln(10)}{c} \cdot \frac{V}{\alpha_{sabine} A} \quad (1.1)$$

У даному виразі α_{sabine} представляє повне поглинання, яке формується за допомогою суми добутків коефіцієнтів звукопоглинання Сабіна та відповідних їм площ поверхонь. Крім того, можна обчислити поглинання за допомогою середнього коефіцієнта поглинання c та загальної площі відповідних відбиваючих поверхонь.

Більше того, час реверберації для даного середовища не залежить від розташування джерела звуку всередині середовища та точки його вимірювання. Крім того, характеристика EDC може бути отримана шляхом обчислення інтегралу Шредера за відомою акустичною імпульсною характеристикою (AIR) середовища, позначеною як $h(t)$ [7].

$$EDC(t) = \int_t^{\infty} h^2(\tau) d\tau \quad (1.2)$$

На рисунку 1.2 наведено графік, що ілюструє EDC для реальної ІХС, яку було виміряно за допомогою бази даних імпульсних характеристик MARDY [7]. Зазначимо чотири відокремлені фази цього відгуку.

– Недалеко від початку відліку часу спостерігається сталий і майже максимальний рівень, який становить опорний рівень 0 дБ. Ця фаза відображає прямий компонент поширення звуку.

– Відразу після початку відліку часу відбувається різке зниження енергії, що вказує на перехід від області прямого поширення та ранніх відбиттів до області вільного затухання.

– У проміжку 50 і 300 мс графік показує нахил із практично постійним від'ємним градієнтом, відповідний зоні вільного затухання.

– На 350 мс помітно, що крива стає більш плавною. Це спостерігається, коли енергія енергетично-часового розподілу достатньо спадає, щоб наблизитися до енергії шуму, яка у цьому випадку складає приблизно -48 дБ.

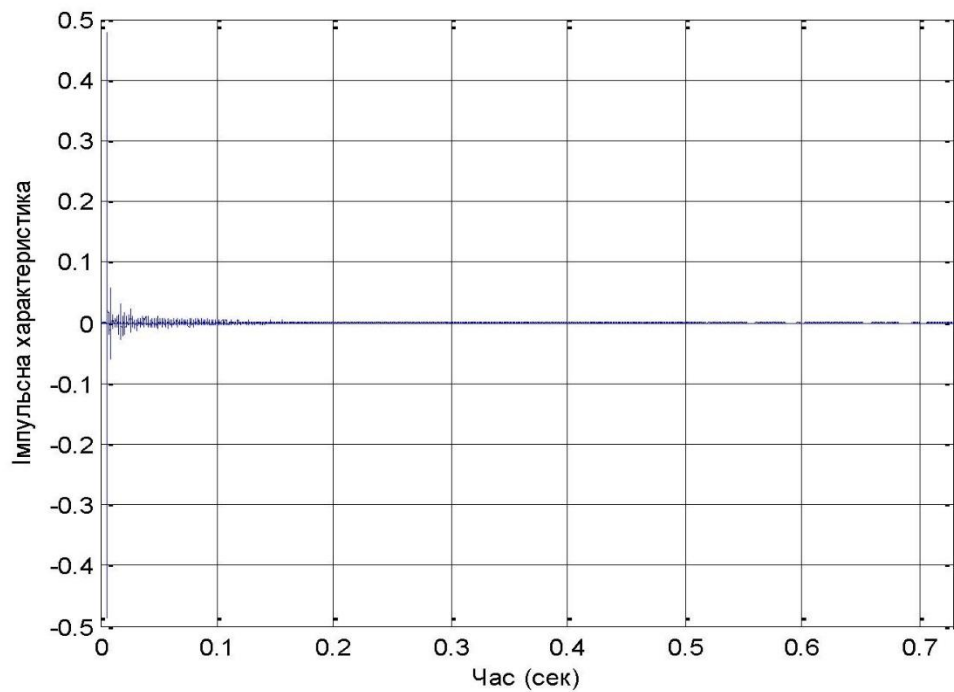


Рис. 1.2 ІХС

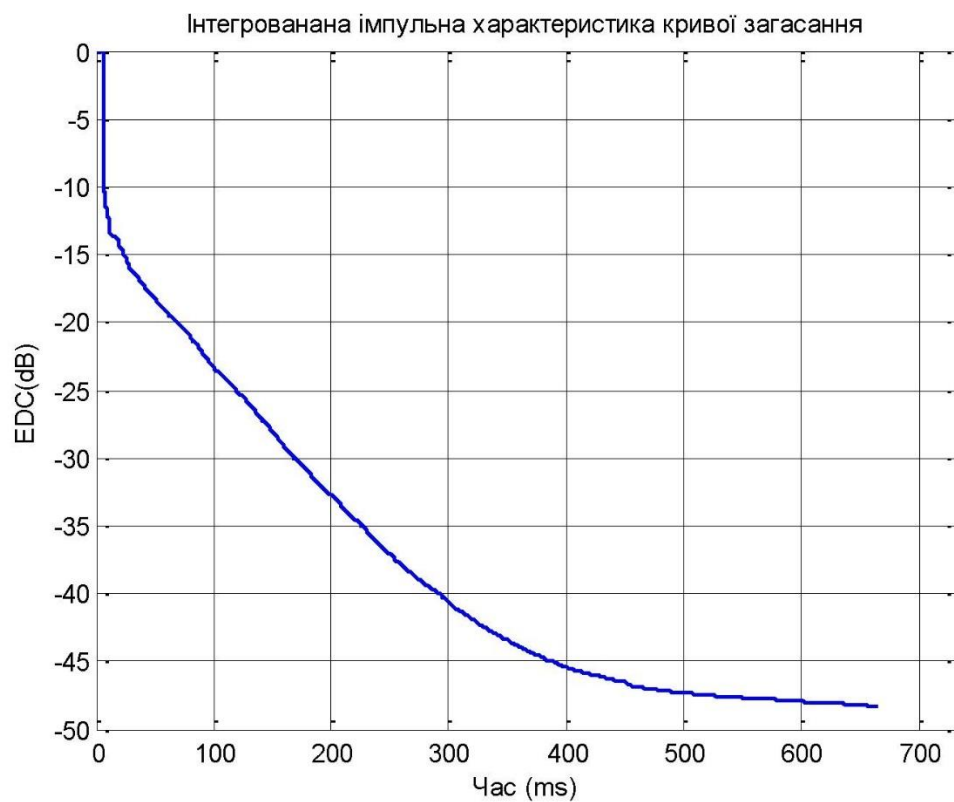


Рис. 1.3 EDC

Вищенаведений рисунок 1.3 відображає EDC.

1.3 Критична відстань

Критична відстань визначається як відстань D_c від джерела, на якій густина звукової енергії компонента прямого проходження сигналу E_d рівна густині звукової енергії ревербераційних компонент (ранньої та пізньої) E_r . Цей параметр розраховується за наступним рівнянням:

$$\frac{Q}{4\pi D_c^2} = \frac{4}{R} \quad (1.3)$$

де

$$D_c = \sqrt{\frac{QR}{16\pi}} \quad (1.4)$$

Було продемонстровано [8], що критичну відстань можна також виразити через параметри Q , V та час реверберації T_{60}

$$D_c \approx 0.1 \sqrt{\frac{QV}{\pi T_{60}}} \quad (1.5)$$

Критична відстань виступає важливим фактором реверберації при розповсюдженні звуку в ЗС.

1.4 Аналіз акустики середовища в залежності від діапазону частот

Існує декілька різноманітних методів дослідження акустики середовища, кожен з яких може бути використаний лише для обмеженого діапазону слухового спектру; немає єдиного аналітичного або числового інструменту, який здатний моделювати ефекти звукової реверберації на всьому звуковому спектрі від 20 Гц до практично 20 кГц. Таким чином, звуковий спектр можна

зручно розподілити на наступні чотири області (I-IV), що визначаються швидкістю звуку, z і найбільшим розміром середовища, v_{max} :

1. При дуже низьких частотах, для яких $F < c/2 v_{max}$, в середовищі відсутні звукові резонанси. Цей діапазон частот може бути розглянутий шляхом використання негармонійних розв'язків хвильового рівняння. Наприклад, для середовища з розмірами $3 \times 5 \times 7$ м та швидкістю звуку 344 м/с відсутні резонансні частоти нижче 25 Гц.

2. У наступному діапазоні частот, $F \sim c/v_{max}$, що простягається від низьких резонансних мод до частоти Шредера, виявляється значущою у акустичному вивченні середовищ [9]. Це концептуально виділяє діапазон частот, в межах якого можна спостерігати різні резонанси, від діапазону, де резонанси розташовані настільки близько один до одного, що їх важко розрізнити, оскільки при будь-якій частоті спостереження в цьому діапазоні ефекти кількох сусідніх резонансів накладаються. З метою більшої точності розглянемо середовище об'ємом V , з резонансами із середнім рівнем 3 дБ смуги пропускання (Δf). Частота Шредера – це значення, при якому резонансні частоти середовища розділені так, що три резонансні частоти знаходяться в межах однієї резонансної смуги пропускання, що вирішує рівняння для f :

$$\langle \Delta f \rangle = 3 \cdot \frac{c^3}{4\pi V f^2} \quad (1.6)$$

Отримана частота Шредера може бути виражена наступним чином:

$$f_g \approx \frac{5500}{\sqrt{V \bar{\xi}}} \quad (1.7)$$

де $\bar{\xi}$ представляє собою середнє значення константи згасання, пов'язаної з кожною резонансною частотою в середовищі.

Далі можна представити наближене вираження для частоти Шредера у такій формі:

$$f_g \approx 2000 \sqrt{\frac{T_{60}}{V}} \quad (1.8)$$

Для середовища з габаритами 3x5x7 метрів, швидкістю звуку 344 м/с та часом реверберації 0,5 секунди цей діапазон частот коливається від 25 до 138 Гц.

3. У діапазоні від частоти Шредера f_g до приблизно $4f_g$ взято довжини хвиль, які є надто короткими для точного моделювання за допомогою акустичних хвильових методів і надто довгими для геометричної акустики. Таким чином, у цьому частотному діапазоні зазвичай використовується статистичне моделювання. Даний діапазон охоплює значення від 138 до 552 Гц для середовища з розмірами 3x5x7 м та $T_{60} = 0,5$ с.

4. На високих частотах $F > C/L_{max}$, де розміри середовища значно перевищують довжину звукової хвилі, застосовується геометричне або променеве моделювання. Це охоплює широкий діапазон звукових частот у звичайних середовищах.

1.5 Висновок до розділу 1

В першому розділі кваліфікаційної роботи:

- Розглянуто аспекти проблем з виявлення звукових сигналів у ЗС.
- Висвітлено аспекти прямого звуку та компоненти реверберації.
- Проведено аналіз методів визначення часу реверберації, оцінювання критичної відстані.
- Досліджено акустичні характеристики середовища в залежності від частотного діапазону.

РОЗІДЛ 2

ОСНОВНА ЧАСТИНА

Внаслідок складних особливостей реверберації в середовищі виникає необхідність в розробці простих та точних моделей процесу реверберації, які можна використовувати, наприклад, для аналізу чи синтезу акустики середовища. Такі моделі можуть бути класифіковані в трьох категоріях [10]: хвильові, променеві та статистичні моделі. Як вже зазначено у розділі 1.3, різні методи аналізу є застосовними для різних частотних діапазонів звуку в середовищі. Отже, комбінування методів моделювання є необхідним для досягнення високої точності для всього спектру звукових сигналів. Проте через обмежений широкосмуговий характер сигналів мовлення, зазвичай використовуються більш прості моделі.

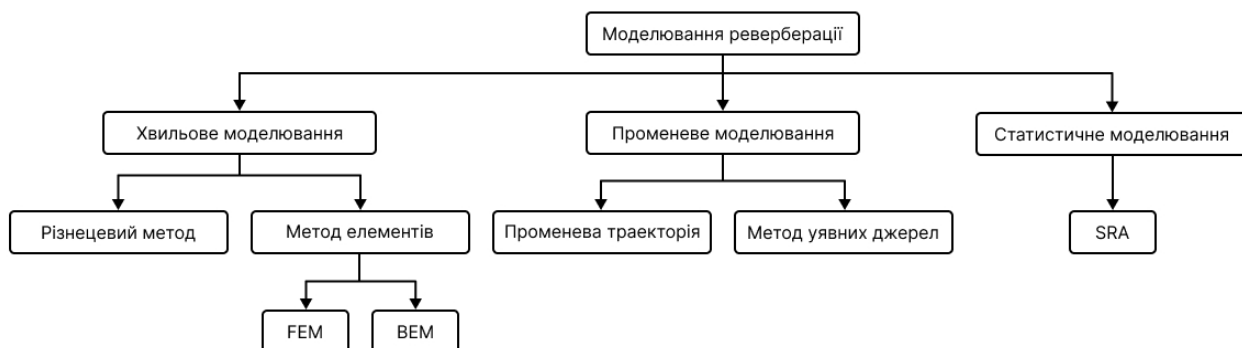


Рис. 2.1 Ревербераційні моделі

На вищенаведеному рисунку 2.1 наведено ревербераційні моделі.

2.1 Хвильове моделювання

Аналітичні рішення для хвильового рівняння зазвичай можливо знайти лише в простих випадках, таких як прямокутні середовища з жорсткими стінами. У інших випадках широко використовують чисельні методи для

вирішення хвильового рівняння. Таким чином, FEM і BEM можуть бути використані для моделювання акустики в середовищі. У обох цих чисельних методах елементи взаємодіють між собою для представлення поширення хвиль. Розмір елементів слід обирати так, щоб він був якнайменшим у порівнянні із значенням довжини хвилі для всіх частот. Висока частота вимагає великої кількості елементів, що призводить до значної обчислювальної складності [11], і, отже, ці методи найбільше підходять для низьких частот і невеликих середовищ.

Іншим методом моделювання акустики середовища є метод скінчених різниць у часовій області (FDTD). Основний принцип цього підходу полягає в тому, що похідні в хвильовому рівнянні замінюються відповідними різницевиими рівняннями.

Метод скінчених різниць формує імпульсні характеристики, які демонструють вищу придатність для сприйняття, ніж випадки, коли використовуються FEM та BEM. З іншого боку, головною перевагою FEM та BEM порівняно із методом SRA є можливість створення густої порожнистої структури, зокрема, в області розташування кутів чи інших акустично складних ділянок. У всіх хвильових методах зазвичай виникає значна складність моделювання, оскільки потрібно враховувати відповідні граничні умови та геометричний опис об'єктів всередині акустичного середовища. Таким чином, використання цих методів у науковому дослідженні було обмежено.

2.2 Траєкторія променів

Розповсюдження звуку у акустиці середовища можна апроксимувати до форми, де звукові хвилі подаються у вигляді променів, а відбиття розглядається як дзеркальне відображення (див. рис. 2.2). Це спрощення застосовується, коли впливи дифракції та інтерференції на поширення хвиль мінімальні, наприклад, при невеликих довжинах звукових хвиль у порівнянні з розмірами відбиваючих

поверхонь у середовищі та великих у порівнянні з будь-якими структурними деталями чи текстурами поверхні [13].

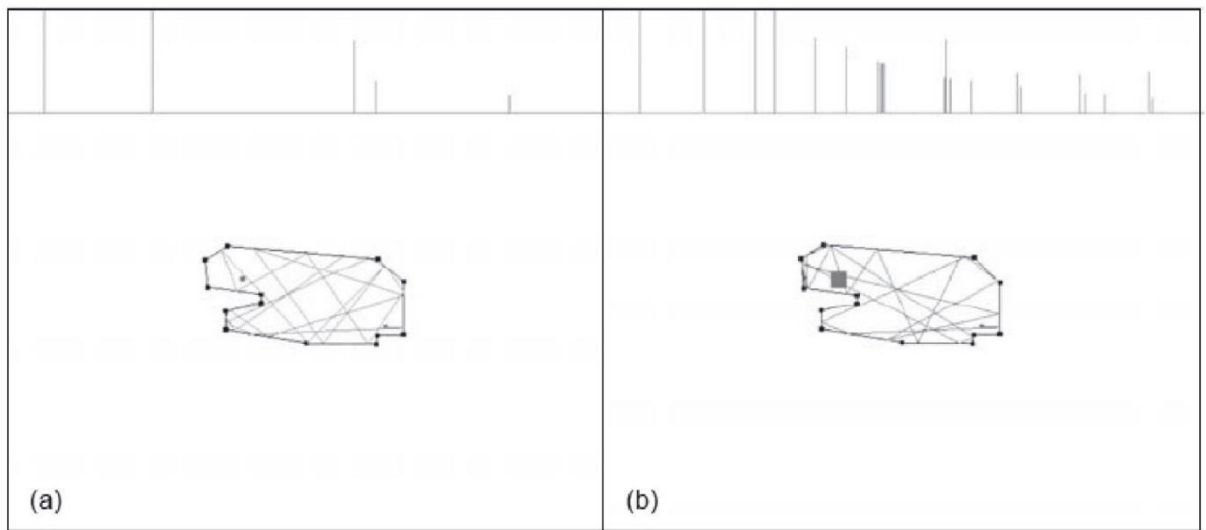


Рис. 2.2 Променевий метод: (a) – розсіяна реакція, (b) – у випадку збільшення площі приймача моделюється більше віддзеркалень

Було висунуто методи променевого моделювання [14], де звукові промені випромінюються з джерела та досягають точки вимірювання через одне чи кілька дзеркальних відбиттів. Вимірювальні дані накопичують промені для формування акустичної реакції від джерела до точки вимірювання. З метою контролю кількості використаних променів у моделюванні, яка може сягати порядку $\sim 10^5$ чи більше, було визначено обмеження моделі, зокрема включення в неї лише відбиттів першого та другого порядку, щоб забезпечити обчислювальну ефективність на низькому рівні.

2.3 Метод уявних джерел

Ще однією технікою променевого моделювання реверберації є метод уявних джерел, який був представлений Алленом та Берклі. Це є одним із найчастіше використовуваних підходів для емуляції акустичних характеристик. З урахуванням єдиного реального ненапрявленого джерела у ревербераційному

середовищі, ефект реверберації відтворюється за допомогою набору уявних вихідних джерел. Розташування зображених джерел визначається габаритами середовища, яке розглядається як прямолінійне. Усі джерела сигналу одночасно випромінюють той самий сигнал, що й реальне джерело. Сигнали, що емітуються віртуальними джерелами, приходять до точок вимірювання з визначеною часовою та інтенсивною характеристикою, що залежить від відстані між джерелами та місцями їх вимірювання. Коефіцієнт відбиття φ використовується для обчислення звуку, який відбивається від поверхонь та пов'язаний із коефіцієнтом поглинання.

Рисунок 2.3 демонструє застосування методу уявних джерел для двовимірного випадку, де середовище позначено виділеним прямокутником. Внаслідок різних відстаней від кожного зображення в області вимірювання сигнали досягають мікрофону у різні моменти часу та із різною інтенсивністю. Остаточні коефіцієнти відбиття позначені $\varphi_{x_1} + \varphi_{x_2} + \varphi_{y_1} + \varphi_{y_2} + \varphi_{z_1} + \varphi_{z_2} = \sqrt{1 - a}$. Отже, для середовища прямокутної форми з розмірами (L_x, L_y, L_z) , де приймач розташований у точці (x, y, z) , а джерело знаходиться в точці $(\bar{x}, \bar{y}, \bar{z})$, AIR обчислюється як

$$h_i = \sum_{\varepsilon=0}^1 \sum_{\rho=-\infty}^{\infty} \varphi_{x_1}^{|q-u|} \varphi_{x_2}^{|q|} \varphi_{y_1}^{|r-v|} \varphi_{y_2}^{|r|} \varphi_{z_1}^{|s-w|} \varphi_{z_2}^{|s|} \times \frac{\delta(n - (|D_\varepsilon + D_\rho|/c))}{4\pi|D_\varepsilon + D_\rho|} \quad (2.1)$$

для уявних джерел. Даний розрахунок імпульсної характеристики середовища впливає із розв'язку хвильового рівняння для прямокутної конфігурації середовища [16].

Де $D_\varepsilon = (x - \bar{x} + 2u\bar{x}, y - \bar{y} + 2v\bar{y}, z - \bar{z} + 2w\bar{z})$ та $D_\rho = (qL_x, rL_y, sL_z)$, де $|D_\varepsilon + D_\rho|$ представляє Евклідову відстань між приймачем та кожним уявним джерелом. Позначення $\varepsilon = (u, v, w)$ та $\rho = (q, r, s)$ вказують на те, що кожна сума фактично складається із трьох різних сум. Крім того, $\rho = (q, r, s)$ визначається на інтервалі $[-\infty; +\infty]$. На практиці, межа

обмежена та залежить від обраного порядку уявних джерел. Цей розрахунок ІХС впливає з розв'язку хвильового рівняння для прямокутного середовища [16].

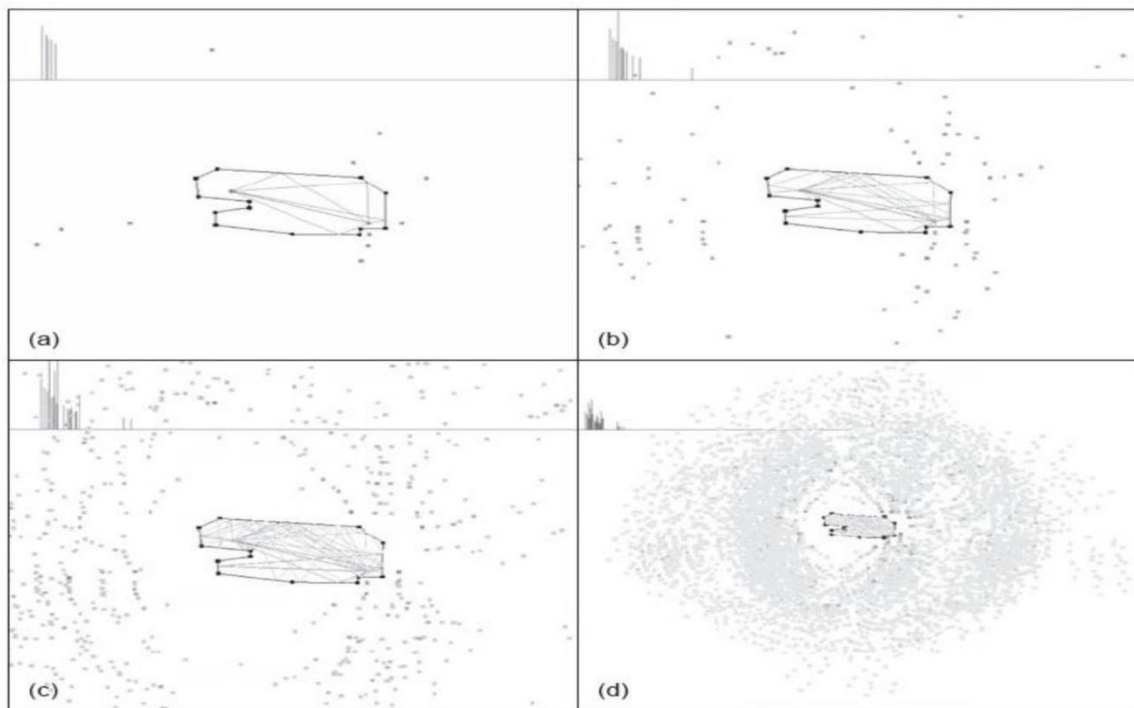


Рис. 2.3 Візуалізація методу уявних джерел

На вищенаведеному рисунку 2.3 (a-d) зображено представлення уявних джерел від 1-го до 4-го порядку.

2.4 Статистична модель середовища

В межах статистичної моделі середовища (SRA), амплітуди та фази усіх відбитих акустичних плоских хвиль в середовищі розглядаються як випадкові величини з рівномірним розподілом. Звукове поле у будь-якій точці середовища вважається розсіяним [17]. Подальше передбачається, що функцію передачі акустичного сигналу $H_m(e^{i\omega})$ (ATF) від джерела до мікрофона можна виразити як суму двох компонент: прямого проходження ($H_{d,m}(e^{i\omega})$), який

включає лише безпосередній шлях поширення, та ревербераційної компоненти ($H_{r,m}(e^{i\omega})$), що містить всі відбиття, де $m = 1, 2, \dots, M$.

У методі SRA для забезпечення належної ефективності враховуються наступні припущення, які також мають застосування в різноманітних практичних сценаріях [18] через широкий діапазон сигналу, що має вагоме значення для мовного зв'язку.

– Габарити середовища значно перевищують довжину хвилі на всіх корисних частотах.

– Середня відстань між резонансними частотами у середовищі менша, ніж на третину їх пропускної здатності. Це спостерігається на всіх частотах, що перевищують частоту Шредера.

– Звукові джерела та мікрофони розташовані у внутрішній частині середовища та розміщені принаймні на половині довжини хвилі від прилеглих стін.

Відповідно до наведених припущень та урахувавши різні напрямки поширення та випадкового фазового співвідношення між прямим проходженням та всіма відбитими хвилями можна припустити, що прямі та ревербераційні компоненти не корелюють між собою. Далі використовується середнє значення $E\{\cdot\}$, яке в цьому контексті визначається як середнє за всіма можливими розташуваннями джерел та мікрофонів у середовищі. Середнє значення щільності енергії спектра ATF тепер можна виразити у формі

$$E\{|H_m(e^{i\omega})|^2\} = E\{|H_{d,m}(e^{i\omega})|^2\} + E\{|H_{r,m}(e^{i\omega})|^2\} \quad (2.2)$$

де величина, що позначена як $E\{\cdot\}$, дає результат, який взагалі залежить від конфігурації джерел та мікрофонів і може бути обчислена за допомогою методів, які вже визначені в Радловіча та Густафсона.

Початкова геометрія визначається за вихідною позицією $q_{src}(0)$, яка служить початковим положенням для кожного мікрофона, де $q_{mic,m}(0)$. Для

створення i -тої координати $q_{src}(i)$ використовуються випадкові вектори перетворення, $\theta(I)$, та вектори обертання, $\Theta(I)$, за допомогою формули $q_{src}(i) = \Theta(i)q_{src}(0) + \theta(i)$.

$$q_{mic,m}(i) = \Theta(i)q_{mic,m}(0) + \theta(i), \quad i = 1, 2, \dots, N \quad (2.3)$$

Отже, просторові зв'язки між мікрофонами та джерелом, а також між кожним мікрофоном і джерелом, зберігають абсолютне положення та орієнтацію в середовищі. Конфігурація джерела мікрофона змінюється випадковим чином. Середнє значення просторового очікування, позначене як $E\{\cdot\}$, потім оцінюється на основі усіх N реалізацій.

Полак спроектував модель в часовому просторі, яка характеризує АІР у формі одного втілення як нестационарний випадковий процес [19].

$$h(t) = b(t)e^{-\bar{\xi}t}, \quad \text{для } t \geq 0 \quad (2.4)$$

де $b(t)$ – стаціонарний випадковий процес Гауса з нульовим математичним сподіванням, що характеризується спектральною густиною $B(f)$, а $\bar{\xi}$ представляє собою постійну середню амплітуду згасання, пов'язану з часом реверберації T_{60} .

$$\bar{\xi} = \frac{3 \ln(10)}{T_{60}} \quad (2.5)$$

Характеристика імпульсу в часовому вимірі може приймати форму лише Гаусової кривої [20], якщо достатня кількість відбиттів перетинається у будь-який момент часу. Таким чином, модель Полака стає ефективною лише після визначеного періоду часу, який відомий як час змішування. Після закінчення цього періоду піку відгуку в повітрі вже не корелюють з окремими відбиттями. З урахуванням того, що щільність відбиття зростає з часом, ця ситуація також

відображається в частотному діапазоні, за винятком того, що розповсюдження відбиття в часовій області може бути визначено лише відносно внутрішніх характеристик середовища. Розповсюдження відбиття в часовій області може бути описане лише в зв'язку із пропускнуою спроможністю джерела або мікрофону [21]. Якщо критерієм є те, що принаймні 10 відбиттів перетинаються протягом характерного часового інтервалу (прийнятого як 24 мс), час змішування визначається за формулою: $t_{mix} = 1000\sqrt{V}$.

Даний вираз було висунуто [22] з метою обчислення часу переходу між ранніми відбиттями та пізньою реверберацією. У дослідженні було продемонстровано, що стохастичні моделі з експоненційним згасанням можуть бути застосовані у геометричній акустиці.

У даному контексті час змішування визначається як інтервал часу, необхідний для того, щоб звукові промені, спочатку прилеглі один до одного, рівномірно розповсюджувалися по всьому просторі середовища. За цей період змішування (при умові, що початок відліку часу встановлено в момент випромінювання акустичного імпульсу джерелом), процес реверберації стає дифузним, означаючи, що щільність звукової енергії та напрям вектору напруженості рівномірно розподілені по всьому об'єму середовища. Характер змішування в середовищі залежить від його геометрії та розсіювальних властивостей меж. При досягненні стадії змішування, щільність відбиттів експоненційно збільшується з часом, а не пропорційно t^2 . Таким чином, значення $1000\sqrt{V}$ може розглядатися як верхня межа для часу змішування в типових середовищах.

2.5 Вибір методу моделювання

На основі вже розглянутих методів моделювання реверберації необхідно визначити основний підхід для подальшого використання у вирішенні завдання усунення ефектів реверберації. Оскільки променеві та хвильові методи, які були розглянуті раніше, характеризуються значною складністю обчислення та

можуть бути застосовані лише в обмеженому частотному діапазоні. Крім того, на їх основі важко розробити ефективний та відносно простий алгоритм фільтрації реверберації. Таким чином, ми будемо далі розглядати алгоритми та засоби усунення ефектів реверберації, які в цілому ґрунтуються на статистичній моделі.

2.6 Штучна генерація реверберації

Для аналізу ефективності алгоритмів усунення ревербераційних ефектів необхідно використовувати еталонний сигнал як без реверберації, так і з нею. Однією з простих структур для виклику реверберації є структура Шредера, побудована на основі всепропускаючого фільтра. На рисунку 2.4 наведено фільтр, який використовує рекурсивну лінію затримки.

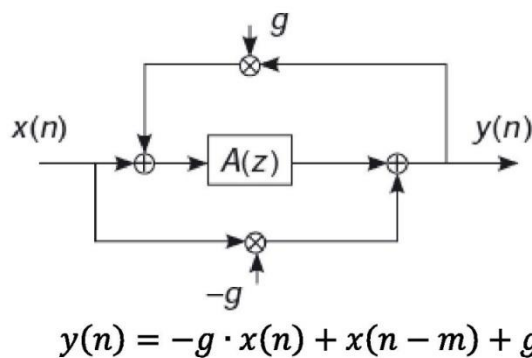


Рис. 2.4 Структура на основі всепропускаючого фільтра: m – довжина затримки у відліках

Структуру фільтра наведено на рисунку 2.4, де $A(z)$ зазвичай замінюється лінією затримки. Цей фільтр дозволяє досягти щільної імпульсної характеристики та рівномірної частотної характеристики. Така конфігурація стала широко використовуваним елементом та застосовується практично в усіх штучних ревербераторах. Припускається, що фільтри зі всією пропускнуою здатністю не призводять до спотворень вхідного сигналу. Тим не менше, це припущення вірне з точки зору сприйняття лише тоді, коли лінія затримки

значно коротша за час адаптації слуху, тобто приблизно 50 мс. Якщо це не виконується, ефекти в часовій області набувають більшої актуальності, і вони суттєво впливають на тембр вхідного сигналу.

Майкл Герзон здійснив узагальнення одноканального та вихідного всепропускаючого фільтра, застосовуючи структури з численними виходами та низькою кількістю входів, де лінія затримки відліків t замінена унітарною мережею порядку N . Прикладами таких тривіальних унітарних мереж є ортогональні матриці та паралельне з'єднання ліній затримки чи всепропускаючих фільтрів. Суть цього узагальнення полягає в підвищенні складності імпульсної характеристики без введення помітного забарвлення частот (див. рис. 2.4).

2.6.1 Структури на базі всепропускаючих та гребінчастих фільтрів

У більшості випадків різні конструкції представляють собою комбінації гребінчастих та повністю пропускних елементарних блоків. Наприклад, розглянемо структуру Мура, яку наведено на рисунку 2.5. Блок (а) реверберації Мура вирішує завдання ранніх відбиттів за допомогою лінії затримки з відводами. Отриманий сигнал передається до блоку (b), який паралельно прямому шляху на одній гілці, на іншій гілці затримується і проходить через дифузійний ревербератор, де відбувається його послаблення. Вихід сигналу ревербера регулюється таким чином, що останні ехо-сигнали виходять з блоку (а), досягають виходу після появи перших ненульових вибірок, що виходять з дифузного ревербера. У вдосконаленій реалізації Мура, ревербератор блоку (b) ефективно реалізований як паралельний набір шести гребінчастих фільтрів, при цьому кожен фільтр першого порядку для фільтрації нижніх частот, а також один фільтр з усією діапазоном частот. Рекомендується встановлювати для останнього фільтра довжину затримки на рівні 6 мс та коефіцієнт фільтра на рівні 0,7. Незважаючи на те, що будь-який фільтр із всепропусканням не вносить забарвлення у АЧХ, його реакція може призвести до звучання з

металевим відтінком або додати непотрібну грубість. Коефіцієнти загасання зворотного зв'язку та фільтрів нижніх частот у гребінчастих фільтрах можуть бути налаштовані для досягнення реалістичного та плавного спаду. Зокрема, ξ_l коефіцієнти ослаблення визначають загальний час загасання серії ехо-сигналів, що генеруються у кожному гребінчастому фільтрі. Якщо бажаний час загасання (зазвичай визначається для рівня ослаблення 60 дБ) позначається T_d , підсилення кожного гребінчастого фільтра повинно бути встановлене в

$$g_i = 10^{-3 \frac{T_d F_S}{m_i}} \quad (2.6)$$

де F_S – частота дискретизації, а m_i – довжина затримки у відліках.

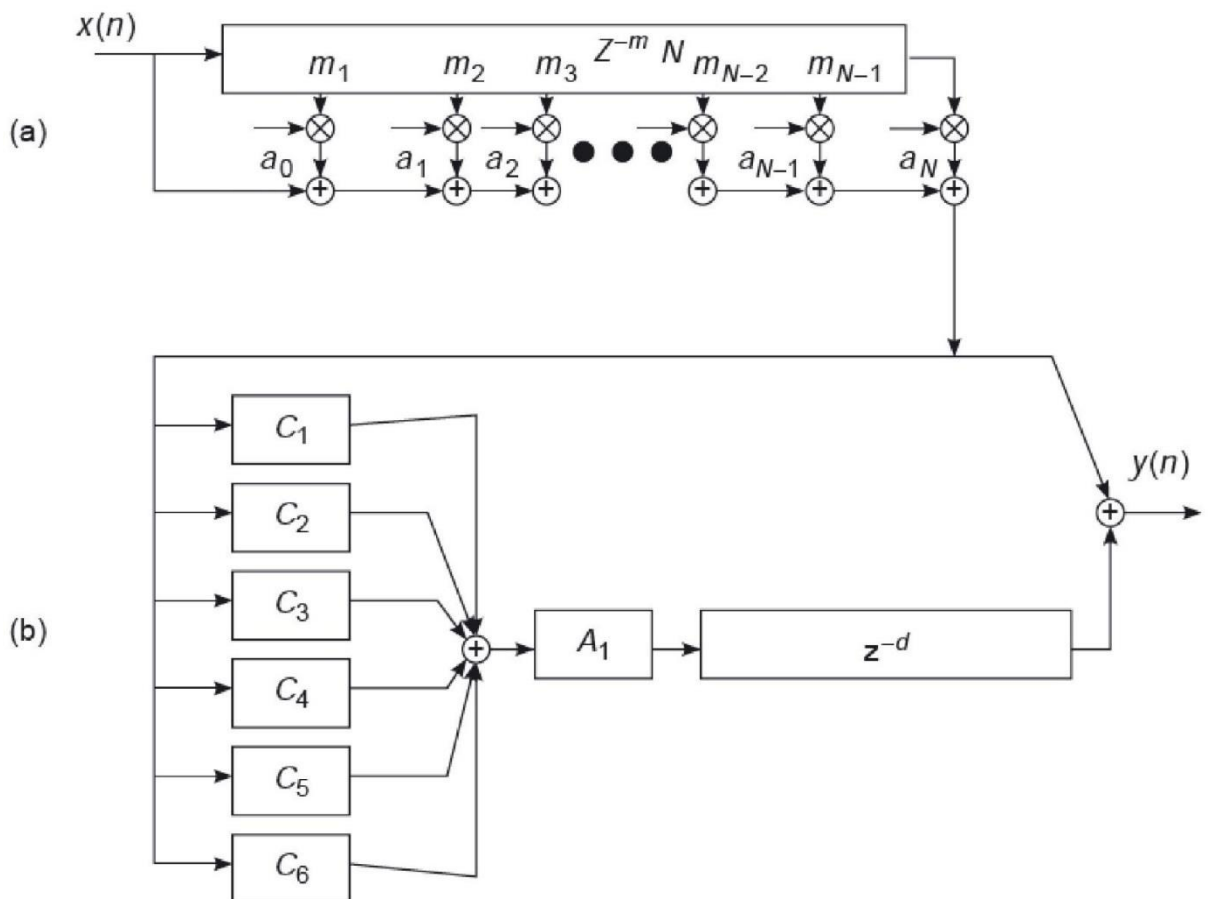


Рис. 2.5 Ревербератор Мура

Окрім того, зниження інтенсивності на високих частотах реалізується за допомогою фільтрів нижніх частот, коефіцієнти яких також можуть бути пов'язані зі зменшенням ефекту реверберації в часовому діапазоні; при цьому конкретна частота може бути визначена шляхом тонкої настройки або експериментів. Наведено приклад коефіцієнтів набору зворотного зв'язку для ефективного зменшення реверберації, а також вказано деякі значення довжин затримки гребневих фільтрів.

$$y(n) = \sum_{j=0}^M [x(j) + \sum_{i=1}^{N-1} a_{i-1} x(j - m_i)] \cdot h_c(n - j) \quad (2.7)$$

де $h_c(n)$ – імпульсна характеристика пізньої реверберації (рис. 2.5), яка охоплює комбіновані гребінчасті та усі пропускаючі фільтри.

Зазвичай їх рекомендується розміщувати в пропорції 1:1,5 від 50 до 80 мс. Шредер висунув ряд теоретичних критеріїв для точнішого вибору довжини затримки [23]; довжина відліків повинна бути взаємно простою, щоб зменшити наслоєння відлунь на імпульсну характеристику та зменшити ефект флаттер-ехо. Цей же критерій можна застосовувати до відстані між кожним відлунням та прямим сигналом до ранніх віддзеркалень.

2.6.2 Мережі зворотніх ліній затримок

Штейнер впровадив структуру штучної реверберації, засновану на системі ліній затримки [24], які утворюють замкнене коло зворотного зв'язку за допомогою матриці, яка наведена на рисунку 2.6.

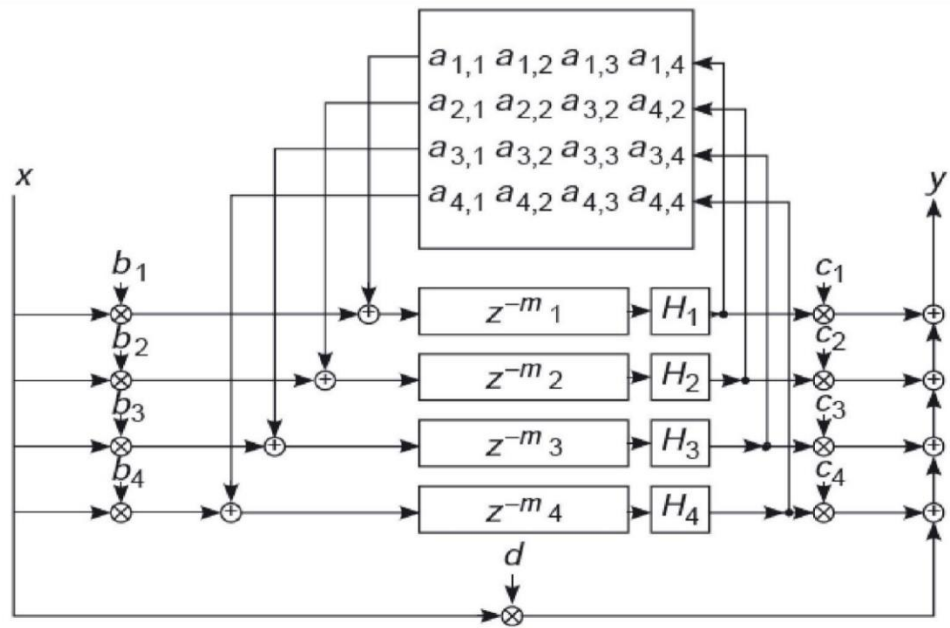


Рис. 2.6 Мережа зворотних ліній затримок 4-го порядку

Згодом такі конструкції стали відомі як мережі зворотних ліній затримки (FDNs). FDN представляють узагальнений вектор рекурсивного гребінчастого фільтра у вигляді

$$y(n) = x(n - m) + g \cdot y(n - m) \quad (2.8)$$

де відлік лінії затримки m був замінений численними лініями затримок різної довжини, а коефіцієнт посилення g був замінений зворотною матрицею G . Stautner і Puckette пропонували матрицю

$$G = g \begin{bmatrix} 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & -1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix} / \sqrt{2} \quad (2.9)$$

За рахунок унікальної специфічної конфігурації, для G достатньо всього одного множника на вихідний канал.

Зв'язок між входом та виходом з матрицею $N \times N$ виражається як

$$\mathbf{y}(\mathbf{n}) = \sum_{i=1}^N \mathbf{c}_i \cdot \mathbf{s}_i(\mathbf{n}) + \mathbf{d} \cdot \mathbf{x}(\mathbf{n}) \quad (2.10)$$

де $\mathbf{s}_i(\mathbf{n})$ представляє вихід i -тої лінії затримки. Крім того, визначається вираз

$$\mathbf{s}_j(\mathbf{n} + \mathbf{m}_j) = \sum_{i=1}^N \mathbf{a}_{ij} \mathbf{s}_i(\mathbf{n}) + \mathbf{b}_j \mathbf{x}(\mathbf{n}) \quad (2.11)$$

де \mathbf{a}_{ij} представляє собою елемент зворотної матриці. Функція передачі системи може бути виражена за допомогою матричного представлення для коефіцієнтів зворотного зв'язку, довжини затримки та векторів \mathbf{q} , \mathbf{b} та \mathbf{c} , які представлені у вигляді N -вимірних векторів:

$$\mathbf{H}(\mathbf{z}) = \frac{Y(\mathbf{z})}{X(\mathbf{z})} = \mathbf{c}^{-T} [\overline{\mathbf{D}}(\mathbf{z}^{-1}) - \overline{\mathbf{A}}]^{-1} \overline{\mathbf{b}} + \mathbf{d} \quad (2.12)$$

де $\overline{\mathbf{D}}(\mathbf{z}^{-1})$ є діагональною матрицею з елементами \mathbf{z}^{-d_i} .

2.6.3 Порівняння отриманих результатів

Як зазначено вище, одним із найбільш поширених алгоритмів є ревербератор Мура (див. рис. 2.7) та мережа зворотних ліній затримок (див. рис. 2.8). Нижче представлені їх імпульсні характеристики.

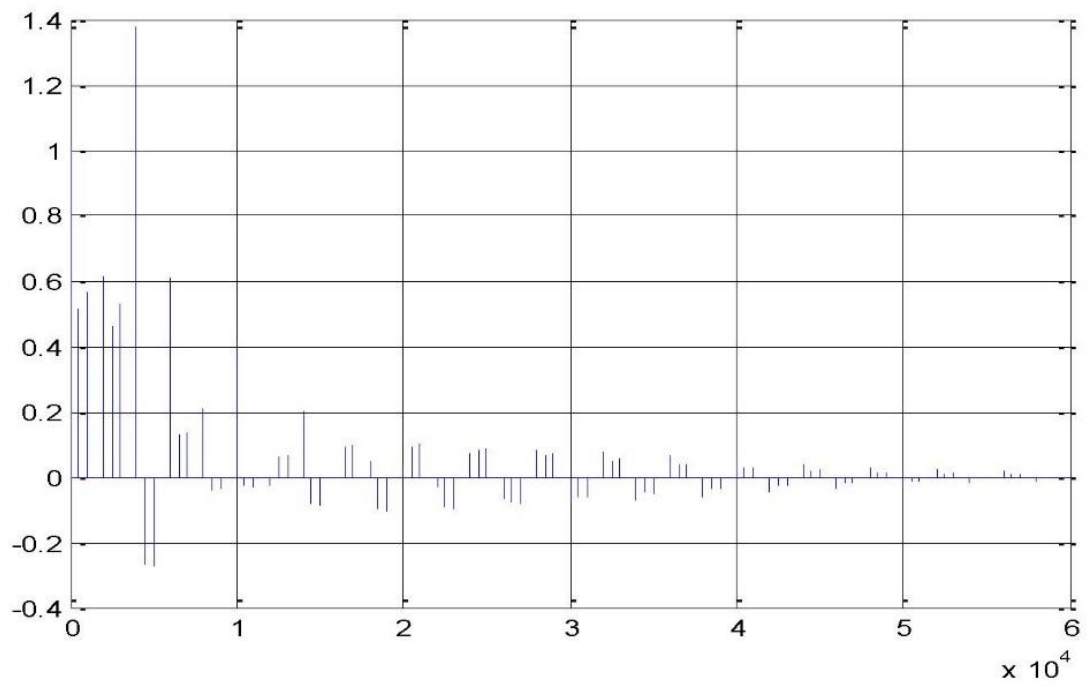


Рис. 2.7 Імпульсна характеристика ревербератора Мура

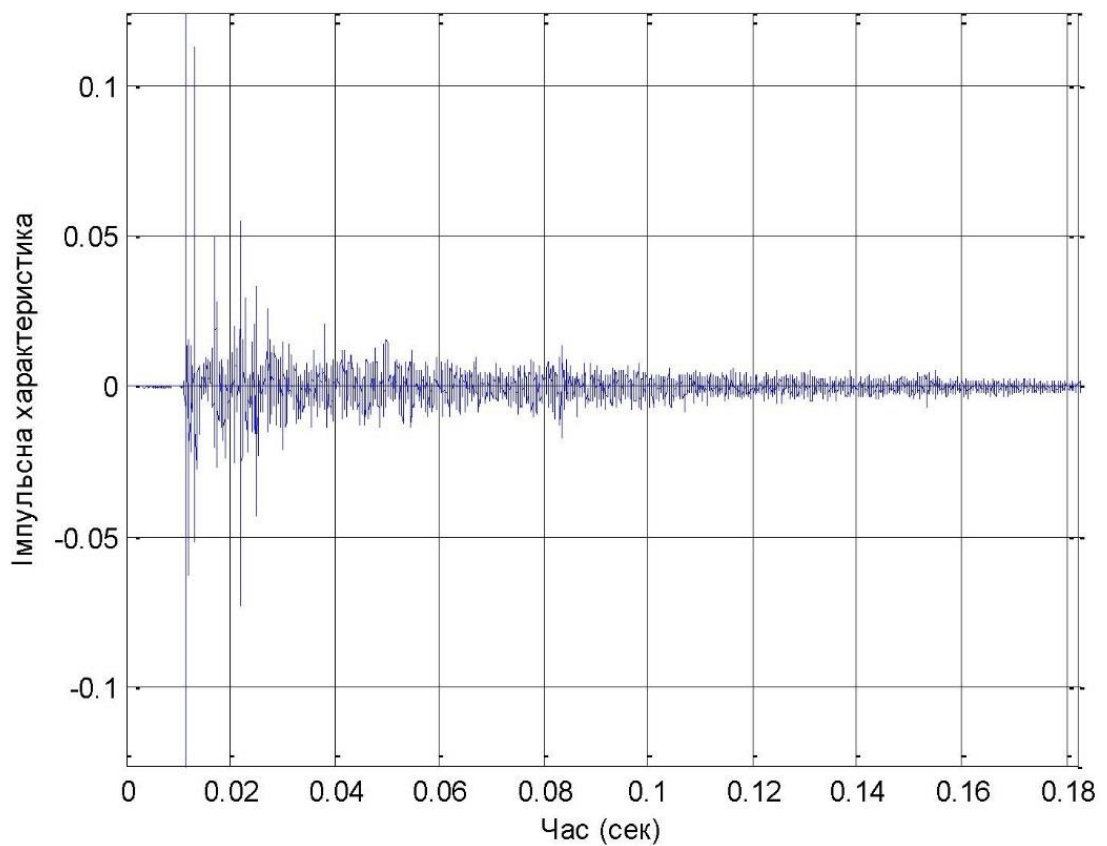


Рис. 2.8 Імпульсна характеристика мережі зворотних ліній затримки

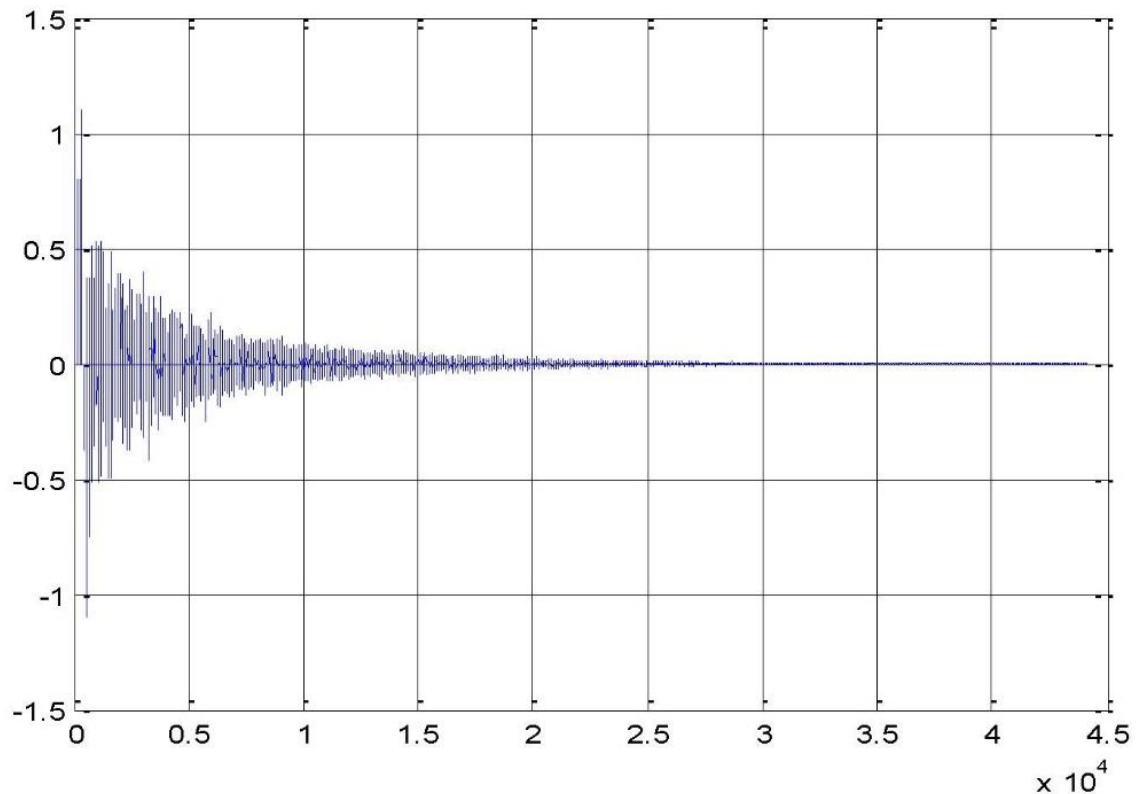


Рис. 2.9 Реальна ІХС

Для порівняння додатково вказано реальну вимірну ІХС (див. рис. 2.9), яка була взята із загальної бази даних для дослідження реверберації.

2.7 Метод підвищення точності розпізнавання звукових сигналів у закритому середовищі

Методи подолання реверберації, які представлені, можуть бути систематизовані за різними критеріями. Такі методи можна класифікувати на одномікрофонні та багатомікрофонні. З метою розгляду одноканальних систем наголос робиться на методах, призначених саме для одномікрофонного подолання реверберації. Більшість алгоритмів, які використовують декілька мікрофонів, не можуть бути застосовані у випадку з одним мікрофоном, оскільки, як правило, вони використовують просторову обробку сигналів. З іншого боку, методи подолання реверберації можна розділити на ті, які

переважно спрямовані на корекцію забарвлення мови, та ті, що зосереджені на ефективному подоланні пізньої реверберації.

В одному із підходів до класифікації методів подолання реверберації [25] визначається типова категоризація, яка не залежить від того, чи проводиться оцінка акустичної імпульсної характеристики, чи ІХС. Цей підхід є результатом розділення на дві основні категорії, відомі як приглушення реверберації та подолання реверберації. Методи першої категорії не враховують оцінку ІХС у той час, як у другій категорії вимагається оцінка ІХС для подальшого подолання реверберації у сигналі.

Особлива увага буде приділена алгоритмам подолання реверберації із використанням одного мікрофона, оскільки це є найбільш розповсюдженим варіантом на мобільних пристроях. Також важливим аспектом є подолання ефекту реверберації для алгоритмів розпізнавання мови.

2.7.1. Приглушення реверберації

Як зазначалося раніше, методи подолання реверберації, що не ґрунтуються на оцінці RIR, поділяються на категорії відповідно до наявності інформації про джерело сигналу чи канал його поширення, а також в залежності від різниці у методах обробки сигналів, задіяних у цьому процесі.

Деякі методи подолання реверберації ґрунтуються на використанні моделювання мовного сигналу шляхом використання базової структури мовного сигналу без еха. Хардвік запропонував модель дуального збудження [26], яка використовується для покращення якості мови. Введення ефекту зміни висоти звуку у цій моделі розширює узагальнену дуальну модель збудження мови. Слід зауважити, що обидві вищезазначені моделі базуються виключно на озвучуванні мовних сегментів.

Був висунутий єдиний ймовірнісний підхід з метою зменшення рівня шуму та подолання реверберації в мовних сигналах. Використовуючи Баєсівське оптимальне оцінювання сигналу для шумоприглушення, а також для

подолання негативного впливу реверберації. Основна концепція даного методу базується в передбаченому тренуванні моделі мови на обширному обсязі даних із без еховою мовою. Ця структура може бути застосована як для одного, так і для кількох мікрофонів, проте подолання реверберації призводить до однаково ефективних результатів [27].

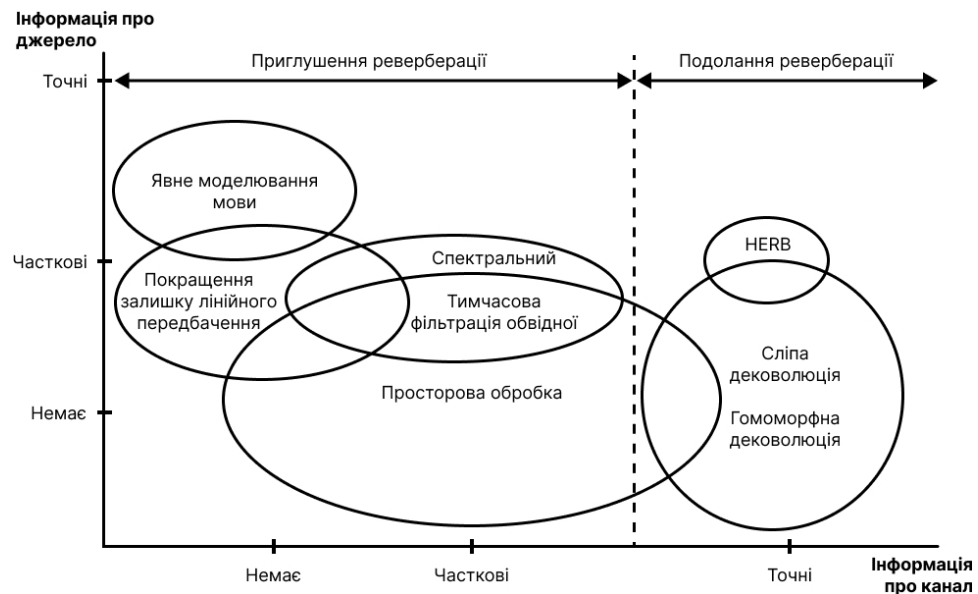


Рис. 2.10 – Класифікація методів зменшення впливу реверберації враховуючи рівень інформації щодо каналу поширення сигналу та джерела

В той час як експериментальні дані свідчать, що оптимальна баєсівська оцінка може перевершити традиційні методи, такі як МСВ, в умовах приглушення шуму, на жаль, ефективність подолання реверберації не може бути оцінена окремо. Крім того, недолік цього підходу полягає в тому, що він значно залежить від фахової підготовки моделі.

2.7.2. Покращення залишкового лінійного передбачення

Моделювання мови у формі згенерованої послідовності, що формується змінним у часі полюсним фільтром, представляє собою широко застосований метод опису мовного сигналу [28]. Моделі стимулювання використовують

послідовність невокалізованої мови, що базується на випадковій послідовності шуму та дзвінких мовних квазіперіодичних імпульсів. Фільтр, використовуваний для формування мовного сигналу, відтворює модель голосового тракту людини. На рисунку 2.11 зображено модель мовної послідовності. Голосовий тракт моделюється полюсним фільтром, коефіцієнти якого оцінюються за допомогою аналізу ЛП записаної мови і отримують назву ЛПС. У цій моделі, ЛР-залишкове представляє собою послідовність генерації, яку можна отримати за допомогою зворотної фільтрації мовного сигналу.

Використання техніки зворотної фільтрації обґрунтовується зауваженням, що в умовах реверберації залишковий звук лінгвальних сегментів містить не лише початкові імпульси, а й кілька інших піків, які виникають через множинне відображення звуку. Однак в цьому підході важливим припущенням є те, що ЛПС не впливає на процес реверберації.

Отже, у цьому варіанті методики подолання реверберації досягається шляхом приглушення піків у послідовності збудження, що виникають внаслідок множинного відбиття відбиттів та синтезу мови за допомогою модифікованого залишкового ЛП та змінного в часі всеполюсного фільтра (ЛП-фільтра) з ЛПС, які обчислені на основі ревербераційної мови подано на рисунку 2.11.

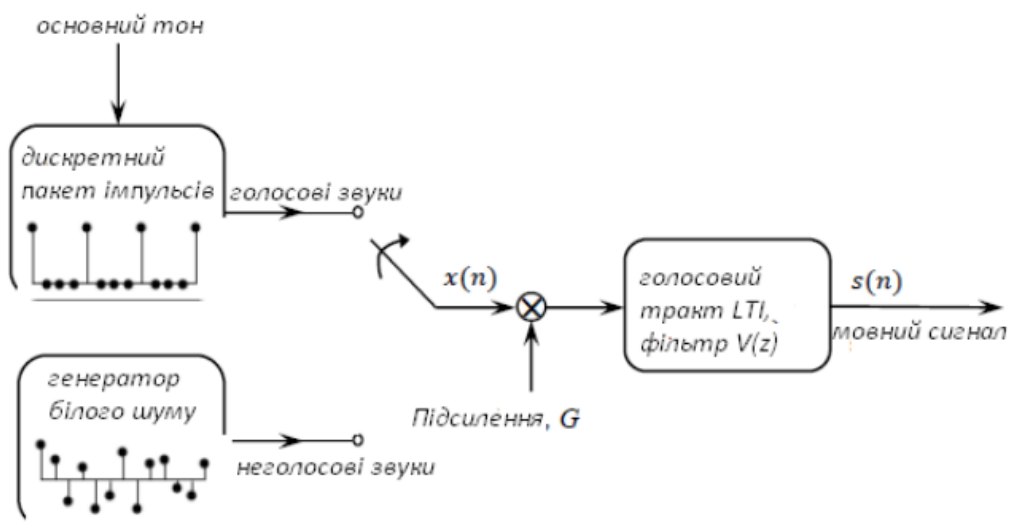


Рис. 2.11 Модель генерації мови з використанням методу явного моделювання

У цьому випадку $x(n)$ відображає зразки ревербераційного сигналу, записаного з M мікрофонів у дискретний час n . Блок аналізу ЛПС виконує функцію методу, яка оцінює полюси фільтра та змінює їх в залежності від часу, як це позначено $\bar{a}(l)$ (де l відображає індекс кадру), і генерує сигнал помилок (різницю), відомий як ЛП-залишковий сигнал.

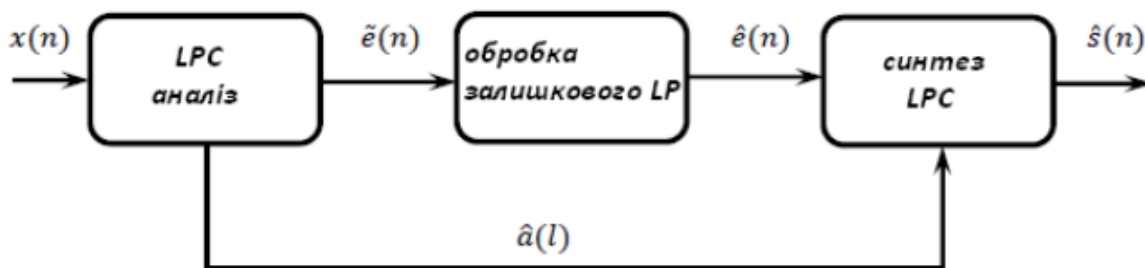


Рис. 2.12 – Загальна організація методу мінімізації реверберації, ґрунтованого на удосконаленні залишкового ЛП

На основі конкретних критеріїв та ознак, що залежать від алгоритму, відбувається зміна залишків LP, і чистий LP-залишковий показник оцінюється. На наступному етапі, покращений мовний сигнал формується за допомогою розрахованих полюсів та оціненого чистого LP-залишку [29].

2.7.3. Спектральне віднімання

Сигнал із реверберацією є результатом згортки чистого сигналу $s(t)$ з каузальною ІХС $h(t)$. Крім того, сигнал з реверберацією можна розглядати як суму ранніх та пізніх компонентів.

$$y(t) = y_E(t) + y_L(t) = \int_{t=0}^{T_L} s(t - \tau)h(\tau)d\tau + \int_{T_L}^{T_R} s(t - \tau)h(\tau)d\tau \quad (2.13)$$

де T_R визначає довжину ІХС $h(t)$, а T_L вказує на інтервал, після якого розпочинається пізня реверберація.

З метою зменшення ефекту пізньої реверберації у спотвореному сигналі, короткочасний спектр пізньої реверберації $Y_L(\mathbf{j}, \mathbf{w})$ віднімається від короткочасного спектру вхідного сигналу з реверберацією $Y(\mathbf{j}, \mathbf{w})$. Спектри $Y_L(\mathbf{j}, \mathbf{w})$ та $Y(\mathbf{j}, \mathbf{w})$ короткочасних сигналів є результатом віконного перетворення Фур'є відповідних сигналів у часовому домені $y_L(t)$ та $y(t)$, де y відображає час, а \mathbf{w} позначає частоту. Очищений спектр сигналу $Y_{SS}(\mathbf{j}, \mathbf{w})$, що відповідає вхідному сигналу без пізньої реверберації, може бути виражений як

$$Y_{SS}(\mathbf{j}, \mathbf{w}) = Y(\mathbf{j}, \mathbf{w}) - Y_L(\mathbf{j}, \mathbf{w}) = \mathbf{G}_{SS}(\mathbf{j}, \mathbf{w})Y(\mathbf{j}, \mathbf{w}). \quad (2.14)$$

Фільтр МСВ $\mathbf{G}_{SS}(\mathbf{j}, \mathbf{w})$ визначається як

$$\mathbf{G}_{SS}(\mathbf{j}, \mathbf{w}) = \frac{Y(\mathbf{j}, \mathbf{w}) - Y_L(\mathbf{j}, \mathbf{w})}{Y(\mathbf{j}, \mathbf{w})} = \mathbf{1} - \frac{Y_L(\mathbf{j}, \mathbf{w})}{Y(\mathbf{j}, \mathbf{w})} \quad (2.15)$$

Отже, для визначення оптимального значення для фільтра МСВ [30] необхідно оцінити короткостроковий спектр пізньої реверберації. Далі будуть розглянуті методи визначення оптимального фільтра МСВ.

2.7.4. Просторова обробка

Одномікрофонні системи використовують лише спектральні та тимчасові характеристики прийнятого сигналу. Реверберація та більшість зовнішніх джерел шуму, зазвичай, також призводять до просторового розподілу. Для додаткового використання цієї особливості необхідно використовувати кілька мікрофонів, а їх виходи слід об'єднати за допомогою відповідного просторового процесора, такого як лінії затримки та сума діаграми напрямленості, або фільтрація та сума діаграми напрямленості чи адаптивний метод формування діаграми напрямленості [31] (див. рис. 2.13).

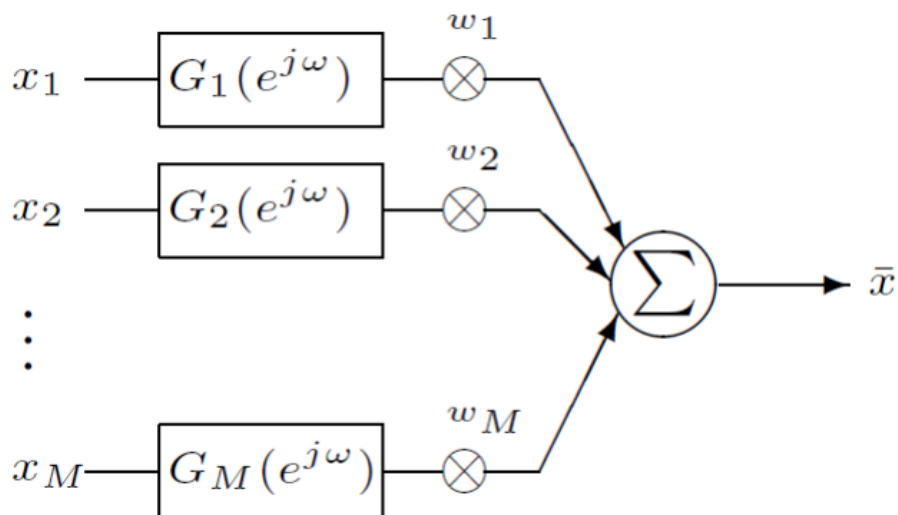


Рис. 2.13 – Загальна організація конструкції діаграми напрямленості

Основний метод мультимікрофонів формується шляхом направлення сформованого променя, де m -тий мікрофон, спрямований в бік джерела, використовує управляючий фільтр $G_m(e^{j\omega})$ та відповідний ваговий коефіцієнт w_m .

Незважаючи на значне покращення якості мови, забезпечене просторовими процесорами, приглушення реверберації обмежено, а приглушення шуму виявляється недостатнім у випадку, коли шумове поле не є когерентним або розсіяним [32]. Додатково до просторового процесора, для досягнення задовільних результатів необхідно використовувати одноканальний фільтр після обробки.

2.7.5. Гомоморфна деконволюція

Даний метод ґрунтується на застосуванні кепстрального аналізу мовного сигналу. Комплексний кепстр має декілька характеристик, що роблять його відмінним вибором для виконання деконволюції. По-перше, сигнали, які в тимчасовій області об'єднуються за допомогою згортки, у комплексному кепстрі об'єднуються адитивно. Внаслідок цього, процес деконволюції зводиться до віднімання у кепстрах. По-друге, комплексний кепстр є мірою

"часової частоти" варіацій в логарифмічному спектрі, і тому сигнал, який повільно змінюється в логарифмічному спектрі, може бути відокремлений від швидкозмінюючих сигналів за допомогою віконної обробки комплексного кепстра (див. рис. 2.14).

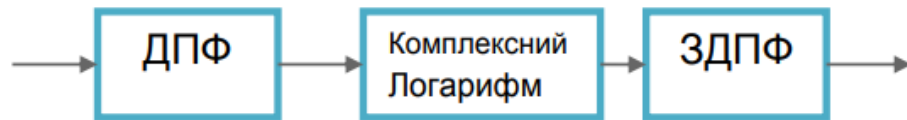


Рис. 2.14 Комплексний кепструм

Мова зазвичай вказує на повільні зміни у логарифмічному спектрі та має складний кепстр, який концентрується близько нуля. Відлуння, що мають затримку від прямого шляху мовлення, можуть бути охарактеризовані за допомогою імпульсної характеристики. У логарифмічному спектрі вони виявляються швидкими "пульсаціями", тоді як у комплексному кепстрі вони складаються з імпульсів, зосереджених неподалік від нуля.

Внаслідок виконання процедур [33] розроблено метод, який дозволяє обчислити комплексний кепстр із сегментів мови з ефектами реверберації. У цьому методі застосовується експоненційно зважена віконна функція до зазначених сегментів, а компоненти косинусного перетворення, що мають відповідні імпульсні характеристики, будуть виключені.

Розглядаючи ехо-сигнали у формі піків у комплексному кепстрі, їхні положення визначаються за допомогою процедури визначення піків і подальше встановлення значень косинусного перетворення у цих точках на нуль. З іншого боку, обчислений кепстр піддається множенню на віконну функцію, призначену для збереження інформації в мовному кепстрі та ефективного вилучення відлуння кепстра. Залишковий кепстр піддаватиметься зворотньому перетворенню в область часу та множенню на зворотню експоненційну віконну функцію (див. рис. 2.15).

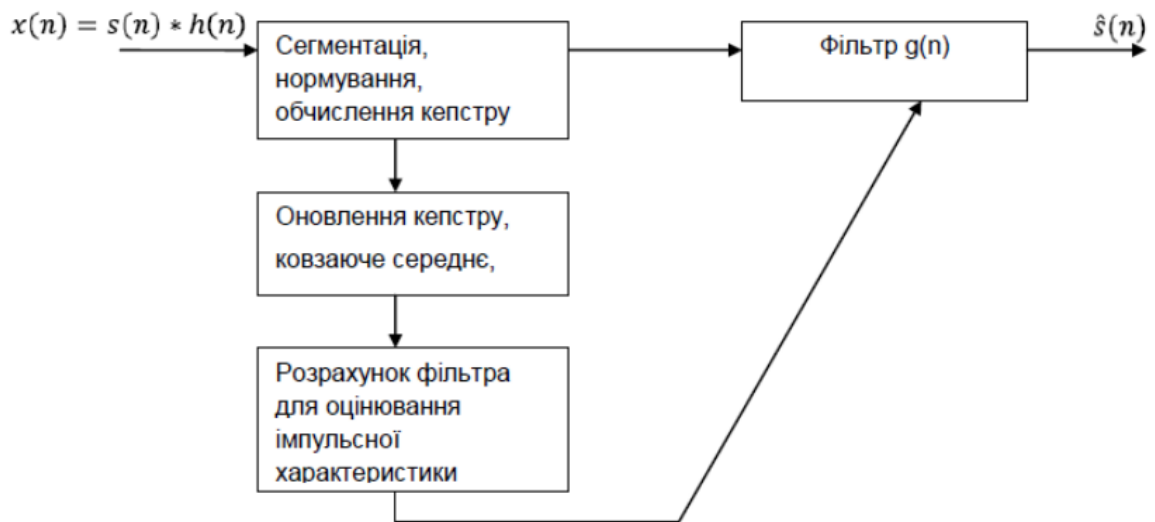


Рис. 2.15 Схема подолання реверберації, заснована на гомоморфній деконволюції

На вищенаведеному рисунку 2.15 зображено блок-схему подолання реверберації на основі гомоморфної деконволюції.

2.7.6. Метод спектрального віднімання для подолання реверберації звукового сигналу

МСВ ґрунтується на використанні статистичного моделювання реверберації в середовищі за допомогою моделі Полака [34]. Реверберований сигнал є результатом згортки чистого сигналу $s(t)$ та каузальної ІХС $h(t)$.

$$y(t) = \int_{-\infty}^{\infty} x(t)h(t - \theta)d\theta \quad (2.16)$$

Із застосуванням моделі Полака

$$h(t) = b(t)e^{-\alpha t}, \text{ для } t \geq 0 \quad (2.17)$$

де $\mathbf{b}(t)$ – це білий гаусівський стаціонарний шум із нульовим середнім значенням. Згортка може бути виражена у вигляді

$$\mathbf{y}(t) = e^{-\alpha t} \int_{-\infty}^t x(t) \mathbf{b}(t-\theta) e^{\alpha \theta} d\theta \quad (2.18)$$

оскільки, якщо $t - \theta$, то $\theta \leq t$, що дозволяє визначити межу інтегралу.

АКФ $r_{xx}(t, t + \tau) = E_x\{x(t)x|t + \tau|\}$ визначається для реверберованого сигналу x у відношенні до часу та зсуву τ :

$$r_{xx}(t, t + \tau) = E \left\{ \int_{-\infty}^t \int_{-\infty}^{t+\tau} s(\theta) s(\theta') h(t - \theta) h(t + \tau - \theta') d\theta d\theta' \right\} \quad (2.19)$$

Застосовуючи статистичну модель Полака, АКФ можна представити у наступному вигляді:

$$r_{xx}(t, t + \tau) = E \left\{ \int_{-\infty}^t \int_{-\infty}^{t+\tau} s(\theta) s(\theta') \mathbf{b}(t - \theta) \mathbf{b}(t + \tau - \theta') e^{-\alpha t} e^{-\alpha t + \tau} e^{-\alpha \theta} e^{-\alpha \theta'} d\theta d\theta' \right\} \quad (2.20)$$

Отже, операція згортки є лінійною та виконується над сигналами $\mathbf{s}(t)$ та $\mathbf{b}(t)$.

Очікуване значення інтегралу дорівнює інтегралу від очікуваного значення, тобто

$$r_{xx}(t, t + \tau) = \int_{-\infty}^t \int_{-\infty}^{t+\tau} E\{s(\theta) s(\theta') \mathbf{b}(t - \theta) \mathbf{b}(t + \tau - \theta')\} e^{-\alpha t} e^{-\alpha t + \tau} e^{-\alpha \theta} e^{-\alpha \theta'} d\theta d\theta' \quad (2.21)$$

У випадку, коли \mathbf{s} та \mathbf{b} розглядаються як незалежні випадкові процеси:

$$r_{xx}(t, t + \tau) = \int_{-\infty}^t \int_{-\infty}^{t+\tau} E\{s(\theta)s(\theta')\}E\{b(t - \theta)b(t + \tau - \theta')\}e^{-\alpha t} e^{-\alpha t + \tau} e^{-\alpha \theta} e^{-\alpha \theta'} d\theta d\theta' \quad (2.22)$$

За використанням теорії, що розглядалась у розділі 2, у статистичній моделі ІХС можна визначити, що

$$r_{xx}(t, t + \tau) = \int_{-\infty}^t \int_{-\infty}^{t+\tau} E\{s(\theta)s(\theta')\}E\{b(t - \theta)b(t + \tau - \theta')\}e^{-\alpha(\theta+\theta'-\tau)} d\theta d\theta' \quad (2.23)$$

$$E\{b(t - \theta)b(t - \tau - \theta')\} = \sigma^2 \delta(\theta - \theta' + \tau) \quad (2.24)$$

де $\delta(\cdot)$ позначає функцію Дірака, а σ^2 відповідає дисперсії $b(t)$.

$$\begin{aligned} r_{xx}(t, t + \tau) &= e^{-2\alpha t} \int_{-\infty}^t \int_{-\infty}^{t+\tau} E\{s(\theta)s(\theta')\}E\{b(t-\theta)b(t+\tau-\theta')\}e^{-\alpha(\theta+\theta'-\tau)} d\theta d\theta' = \\ &= e^{-2\alpha t} \int_{t-\tau}^{t+\tau} E\{s(\theta)s(\theta+\tau)\}\sigma^2 e^{2\alpha\theta} d\theta + e^{-2\alpha t} \int_{-\infty}^{t-\tau} E\{s(\theta)s(\theta+\tau)\}\sigma^2 e^{2\alpha\theta} d\theta \end{aligned} \quad (2.25)$$

Автокореляція в момент t може бути розкладена на дві складові. Перший член залежить від безпосереднього сигналу між часом $t - T$ і t , тоді як другий вплив пов'язаний від сигналу реверберації та відповідає за ранні відбиття. Розглянемо автокореляцію, усереднену просторово в момент часу $t - T$:

$$r_{xx}(t - T, t - T + \tau) = e^{-2\alpha(t-T)} \int_{-\infty}^{t-\tau} E\{s(\theta)s(\theta+\tau)\}\sigma^2 e^{2\alpha\theta} d\theta \quad (2.26)$$

Тепер маємо можливість спостерігати, що автокореляція в момент часу T може бути представлена у вигляді

$$r_{xx}(t, t + \tau) = r_{x_d x_d}(t, t + \tau) + r_{x_T x_T}(t, t + \tau) \quad (2.27)$$

де

$$r_{x_d x_d}(t, t + \tau) = e^{-2\alpha t \int_{t-T}^{t+\tau} E\{s(\theta)s(\theta+\tau)\} \sigma^2} e^{2\alpha\theta} d\theta \quad (2.28)$$

$$r_{x_T x_T}(t, t + \tau) = e^{-2\alpha T} r_{xx}(t - T, t - T + \tau) \quad (2.29)$$

На практиці можна розглядати сигнали як стаціонарні впродовж невеликого часового інтервалу, що є невеликим порівняно із часом реверберації T_r . Це обгрунтовано тим, що експонентний спад виявляється надзвичайно повільним, а мова вважається квазістаціонарною. Позначимо T_s як інтервал часу, протягом якого сигнал мови можна розглядати як стаціонарний, і цей проміжок часу зазвичай становить приблизно 20-40 мс [35]. Ми припускаємо, що $T_s < T < T_r$.

Спектральні характеристики стохастичного сигналу визначаються через виконання перетворення Фур'є АКФ. Іншими словами, розподіл потужності за частотою визначається функцією:

$$\gamma(f) = \int_{-\infty}^{\infty} r(\tau) e^{-2\pi f \tau} d\tau \quad (2.30)$$

і короткострокова спектральна густина приблизно дорівнює:

$$\gamma_{x_T x_T}(t, f) = \gamma_{x_d x_d}(t, f) + \gamma_{x_T x_T}(t, f) \quad (2.31)$$

$$r_{x_T x_T}(t, f) = e^{-2\alpha T} r_{xx}(t - T, f) \quad (2.32)$$

Отже, ми маємо змогу оцінити СЩП вихідного сигналу, використовуючи МСВ СЩП від відповідної характеристики пізньої реверберації.

Метод Лебарта, Бочера та Дебніха. Для отримання амплітуди короткочасного спектру пізньої реверберації за допомогою методу Лемберда, Бочера і Дебніха можна обчислити СЩП пізньої реверберації та взяти квадратний корінь з неї [36]. Користуючись відомістю про те, що СЩП

представляє собою віконне перетворення Фур'є АКФ пізньої реверберації в часовому діапазоні $E[\mathbf{y}_L(t)\mathbf{y}_L(t + \tau)]$.

Значущим є той аспект, що для оцінки СЦП або автокореляції $\mathbf{y}_L(t)$ елементу пізньої реверберації, необхідна більш деталізована інформація стосовно пізньої та ранньої складових реверберації вхідного сигналу $\mathbf{y}(t)$ не є обов'язковою.

Полак представив статистичну модель [37] ІХС, що ґрунтується на послідовності незалежних та рівномірно розподілених випадкових величин $\mathbf{b}(t)$ з нульовим математичним сподіванням та нормальним законом розподілу:

$$\mathbf{h}(t) = \mathbf{b}(t)e^{-\rho t}, \text{ для } t \geq 0 \quad (2.33)$$

де $\bar{\rho}$ – константа середнього затухання, що пов'язана з часом реверберації T_{60} :

$$\bar{\rho} = \frac{3 \ln(10)}{T_{60}} \quad (2.34)$$

За наявності стаціонарного гаусівського шуму $\mathbf{b}(t) \sim N(\mathbf{0}, \sigma^2)$, що ґрунтується на інтегральній характеристиці стану, АКФ вхідного сигналу $\mathbf{y}(t)$ може бути виражена у вигляді

$$E[\mathbf{y}(t), \mathbf{y}(t + \tau)] = \int_{t=0}^t \int_{t=0}^{t+\tau} E[s(\phi)s(\phi')] E[\mathbf{b}(t - \phi)\mathbf{b}(t + \tau - \phi')] e^{\rho(\phi + \phi' - \tau)} d\phi d\phi' = e^{-2\rho t} \int_{t=0}^t E[s(\phi)s(\phi')] \sigma^2 e^{2\rho\phi} d\phi \quad (2.35)$$

Припускаючи, що енергія відбиттів перевищує енергію звуку прямого проходження, можна встановити зв'язок автокореляції вхідного сигналу з автокореляцією відбиттів, здійснених на пізніший час. Шляхом розкладання $\mathbf{y}(t)$ на $\mathbf{y}_E(t)$ та $\mathbf{y}_L(t)$ за визначений часовий інтервал T_L , можна продемонструвати, що пізня реверберація може бути апроксимована як

некорельований шумовий процес. В результаті АКФ сигналу пізньої реверберації матиме такий вигляд:

$$E[y_L(t)y_L(t + \tau)] = e^{-2\rho T_L} E[y(t - T_L)y(t - T_L + \tau)] \quad (2.36)$$

На підставі цих розглядів можна отримати оцінку для СЦП пізньої реверберації $|Y_L(j, \omega)|^2$, що відповідає короткочасному представленню автокореляції.

$$|Y_L(j, \omega)|^2 = e^{-2\rho T_L} |\bar{Y}(j - N_L, \omega)|^2 \quad (2.37)$$

Оскільки це короткостроковий сигнал СЦП, час стає дискретним, розділеним на кадри, що призводить до заміни N_L на T_L . Оцінка СЦП реверберації вхідного сигналу від кадру до кадру не лише залежить від поточного кадру $|Y_L(j, \omega)|^2$, але й від оцінки СЦП реверберації попереднього кадру. Це можна отримати, використовуючи вираз СЦП для реверберації через АКФ, де завжди важливо враховувати довжину часових кадрів. Оцінка СЦП вхідного сигналу може бути рекурсивно отримана, використовуючи середнє значення між поточним та попереднім кадром, і також застосовується згладжування з коефіцієнтом $0 \leq \alpha \leq 1$:

$$|Y_L(j, \omega)|^2 = \alpha |Y(j - N_L, \omega)|^2 + (1 - \alpha) |Y_L(j, \omega)|^2 \quad (2.38)$$

Як зазначено вище, повний короткочасний спектр пізньої реверберації $Y_L(j, \omega)$ можна скласти на основі його амплітуди і фази. З огляду на те, що фаза спектру пізньої реверберації ідентична фазі вхідного реверберованого сигналу, застосування МСВ не призводить до зміни фази. Таким чином, АЧХ фільтра МСВ може бути виражена

$$|G_{SS}(j, \omega)| = \mathbf{1} - \frac{|Y_L(j, \omega)|}{Y(j, \omega)} = \mathbf{1} - \frac{e^{-2\rho T_L} |\bar{Y}(j - N_L, \omega)|}{|Y(j, \omega)|} \quad (2.39)$$

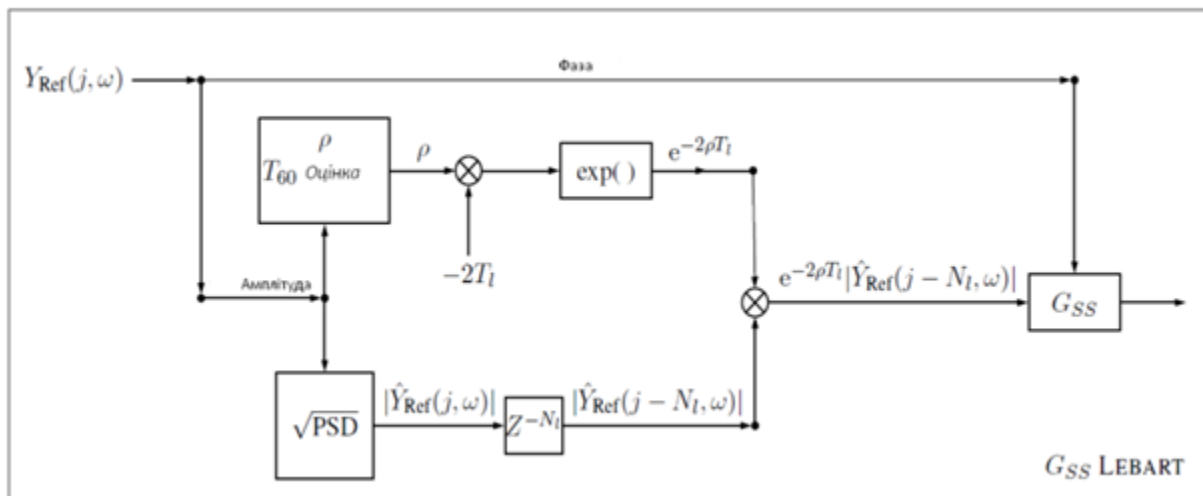


Рис. 2.16 – Фільтр МСВ Лембарта

Метод Ву та Ванга. Цей метод описує короточасний спектр потужності окремої частини сигналу зі затриманою реверберацією як згладжений та зсунутий відповідно до СЩП сигналу з реверберацією. Згладження в часовому домені відображає ефект розмиття запізнених імпульсів у ревербераційному середовищі. Додатково, СЩП запізнених компонентів уявляється як згладжена та зсунута версія СЩП ревербераційного сигналу.

$$|Y_L(j, \omega)|^2 = \gamma r(j - \rho) \cdot |\bar{Y}(j, \omega)|^2 \quad (2.40)$$

Параметр затримки ρ і масштабуючий фактор γ представляють зміщення затримки та визначають ступінь впливу короткострокової пізньої реверберації СЩП. Зсув і згладжування визначаються шляхом складання СЩП сигналу та застосування асиметричної функції згладжування, що адаптується до форми вирівняної імпульсної характеристики реверберації. Розподіл Релеєва задовольняє ці умови і визначається як

$$r(j) = \begin{cases} \frac{j+\alpha}{\alpha^2} e^{-\frac{(j+\alpha)^2}{\alpha^2}}, & \text{якщо } f \geq -\alpha \\ 0, & \text{інше} \end{cases} \quad (2.41)$$

Графічне відображення функції Релея зображено на рисунку 3.8 при значеннях параметрів $\rho = 300$, $\alpha = 200$, $\gamma = 3000$.

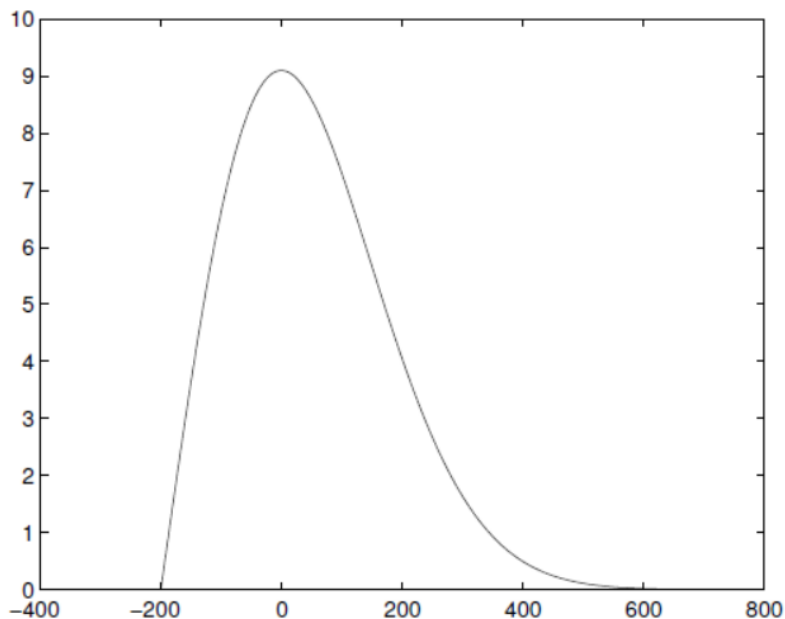


Рис. 2.17 Розподіл Релеєва, який виступає у ролі згладжувальної функції СЦП для ефекту пізньої реверберації

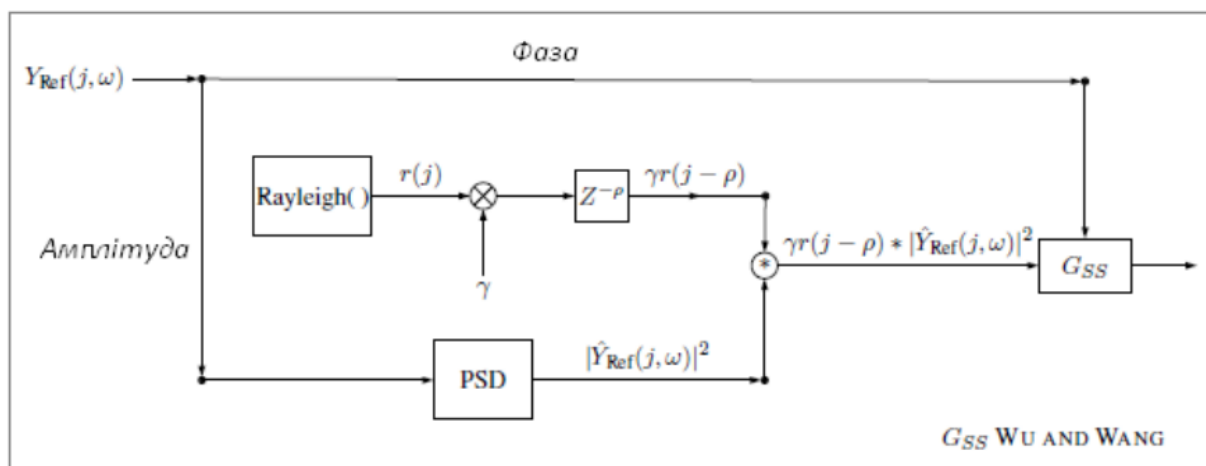


Рис. 2.18 Фільтр МСВ за методом Ву і Ванга.

Спектр сигналу без урахування реверберацій можна визначити шляхом застосування коефіцієнта підсилення, який обчислюється за методом, розробленим Ву і Вангом, з використанням спектру вхідного сигналу із реверберацією.

$$Y_{SS}(j, \omega) = Y(j, \omega) - Y_L(j, \omega) = \left\{ \frac{|\bar{Y}(j, \omega)|^2 - |Y_L(j, \omega)|^2}{|\bar{Y}(j, \omega)|^2} \right\}.$$

$$Y(j, \omega) == G_{SS}(j, \omega)Y(j, \omega) \quad (2.42)$$

Метод Блейкера та Георганті. Не лише використання згортки Релеєвського розподілу зі спектром сигналу, що має реверберацію, призводить до отримання необхідних результатів, але також застосування згортки цієї згладжувальної функції з реверберованим сигналом у часовому домені, яке використовується для досягнення ефекту розмиття пізньої реверберації. Згладжування та зсув у часовому домені для реверберації в даний момент визначається згладженою версією реверберованого сигналу з певних попередніх кадрів (див. рис. 2.13).

$$y_L(j) = yr(j - \rho) \cdot y(j) \quad (2.43)$$

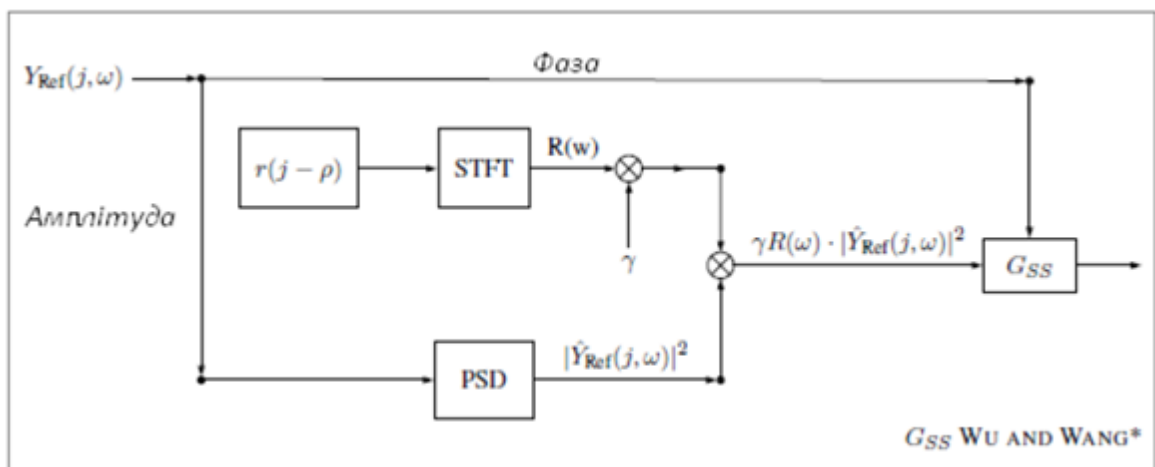


Рисунок 2.19 – Фільтр МСВ згідно з методом Ву та Ванга з використанням модифікації від Блейкера та Георганті.

Згладжування та зсув часового відображення реверберації мовлення можуть бути адаптовані в частотному діапазоні.

$$|Y_L(j, \omega)|^2 = \gamma \cdot STFT\{r(j - \rho)\} \cdot |\bar{Y}(j, \omega)|^2 \quad (2.44)$$

Згортка в області часу еквівалентна перемноженню в області частоти. Множення спектру сигналу реверберації здійснюється застосуванням віконного перетворення Фур'є до затриманого розподілу Релея

$$STFT\{r(j - \rho)\} = R(\omega) \quad (2.45)$$

Метод Фарає та Катаокі. Такий метод визначає підсилення фільтра МСВ, безпосередньо використовуючи оцінку СЦП компоненту пізньої реверберації та вхідного сигналу з реверберацією.

$$G_{SS}(j, \omega) = \frac{|\bar{Y}(j, \omega)|^2 - |Y_L(j, \omega)|^2}{|\bar{Y}(j, \omega)|^2} \quad (2.46)$$

Вони пропонують оцінити короткочасний СЦП пізньої реверберації в кожному кадрі, використовуючи метод сумування відфільтрованих, попередніх версій кадрів ревербераційного сигналу [38]. Спочатку розраховуються вагові коефіцієнти для кожного кадру СЦП, після чого обчислюється спектр потужності попередніх кадрів, які згодом сумуються із відповідними ваговими коефіцієнтами (див. рис. 2.20).

$$|Y_L(j, \omega)|^2 = \sum_{l=1}^K |a_l(j, \omega)|^2 \cdot |\bar{Y}(j - l, \omega)|^2 \quad (2.47)$$

Вагові коефіцієнти дозволяють враховувати вплив спектру сигналу із реверберацією попереднього кадру СЦП на спектр поточного кадру. З чим більша подібність між двома СЦП, тим суттєвіше їхній вплив на спектр попереднього кадру при переході до поточного кадру, зокрема для компоненти пізньої реверберації.

$$a_l(j, \omega) = E \left\{ \frac{|\bar{Y}(l, \omega)| \cdot |\bar{Y}^*(j-l, \omega)|}{|\bar{Y}(j-l, \omega)|^2} \right\} \quad (2.48)$$

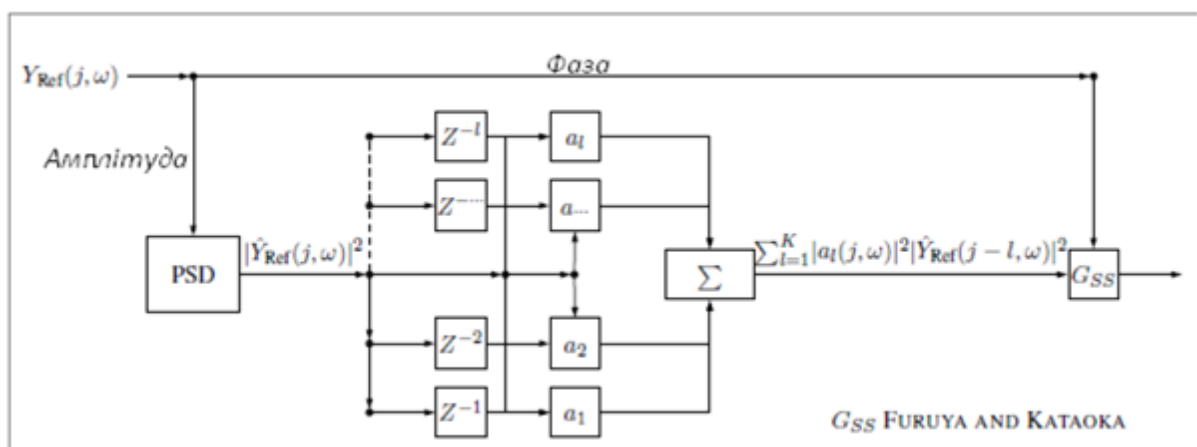


Рис. 2.20 Фільтр МСВ за методом Фарая і Катаоки

Висновок до розділу 2

Отже, в результаті аналізу імпульсних характеристик генераторів встановлено, що вони виявляють схожий характер, але породжують ранні та пізні ревербераційні ефекти у різний спосіб. У Муровій моделі первинна реакція звучить досить дискретно, що спричиняє зернистість звуку під час імпульсних сигналів, тоді як у моделі зворотних ліній затримок ранні відбиття більш розсіяні. З цієї причини обидва алгоритми будуть використовуватися для подальшого аналізу методів приглушення реверберації.

На основі аналізу видатних методів компенсації ревербераційних ефектів для поліпшення якості виявлення акустичних сигналів у ЗС зроблено висновок, що найоптимальнішим для застосування є МСВ.

Здійснено аналіз різних методів до оцінки спектру пізньої реверберації для МСВ. Враховуючи переваги та недоліки вибрано МСВ, який використовує час реверберації для обчислень, підвищуючи точність вимірювань у заданих параметрах середовища. Це застосування дозволяє детальніше проаналізувати ефективність алгоритму, зокрема застосованого методу, оскільки використання генераторів неприродної реверберації дає можливість встановлювати час реверберації і досліджувати результати фільтрації. Отже, у подальших дослідженнях планується використання МСВ з оцінкою спектру пізньої реверберації відповідно до методу Дебніха.

РОЗДІЛ 3

НАКОВО-ДОСЛІДНА ЧАСТИНА

Моделювання було виконано у середовищі MATLAB. Для цього використовувався метод придушення пізньої реверберації, заснований на спектральному відніманні із застосуванням методу Дебніха для оцінки спектру пізньої реверберації. Для створення реверберованого сигналу застосовувались алгоритми генерації реверберації, такі як алгоритм Герцеля та мережа лінійних затримок (див. Розділ 2).

Моделювання виконувалося застосовуючи наступний алгоритм:

- Початковий сигнал передається до генератора штучної реверберації, результатом якого є сигнал із реверберацією (див. рис. 3.1).
- Сигнал із реверберацією вводиться в схему усунення пізньої реверберації, що призводить до отримання відфільтрованого сигналу (див. рис. 3.1).
- Порівнюємо чистий вхідний сигнал та відфільтрований сигнал після застосування алгоритму за допомогою спектрограм та абсолютної різниці.

Важливо відзначити, що оцінити ефективність роботи алгоритму досить складно, оскільки ми фільтруємо лише пізню реверберацію. Наявність ранньої реверберації не дозволяє порівняти вихідний відфільтрований сигнал з чистим початковим сигналом.

3.1. Моделювання гармонічного сигналу

З метою визначення більш зрозумілого виконання алгоритму, проведемо його моделювання на простому сигналі з декількома гармоніками (акордів), із значним періодом реверберації (1,5 с). Як наведено на рисунку 3.1, реверберований сигнал проявляє “хвіст” з пізньою реверберацією, тоді як оцінений сигнал вже має приглушену пізню реверберацію. З цього можна зробити висновок, що алгоритм ефективно працює.

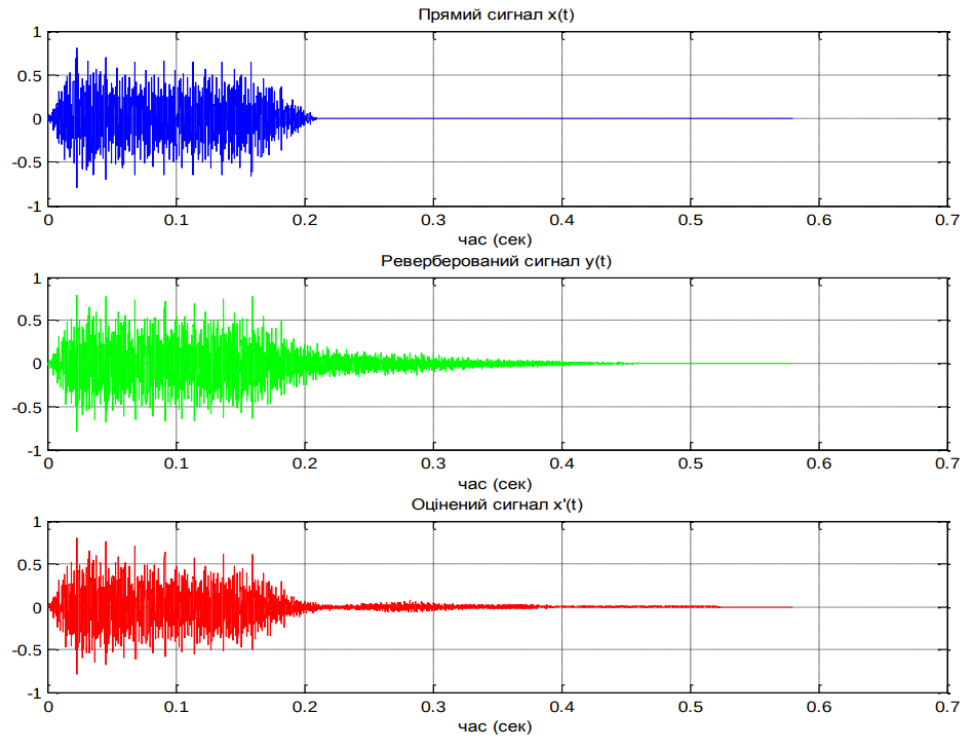


Рис. 3.1 Відображення сигналів у часовому просторі

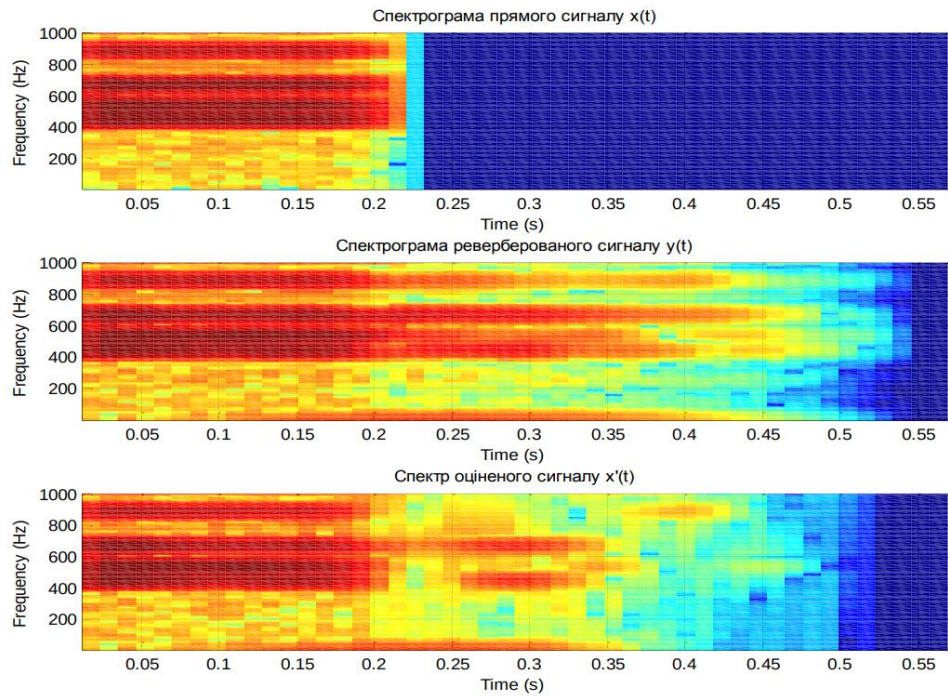


Рис. 3.2 Спектрограма сигналів

3.2 Моделювання з використанням звучання акорду на гітарі

Оскільки алгоритм призначений для обробки звукових сигналів, доцільно моделювати акустичний сигнал. У даному випадку розглянуто акустичний гітарний акорд з тривалістю реверберації (див. рис. 3.3), що дорівнює 1.3 с.

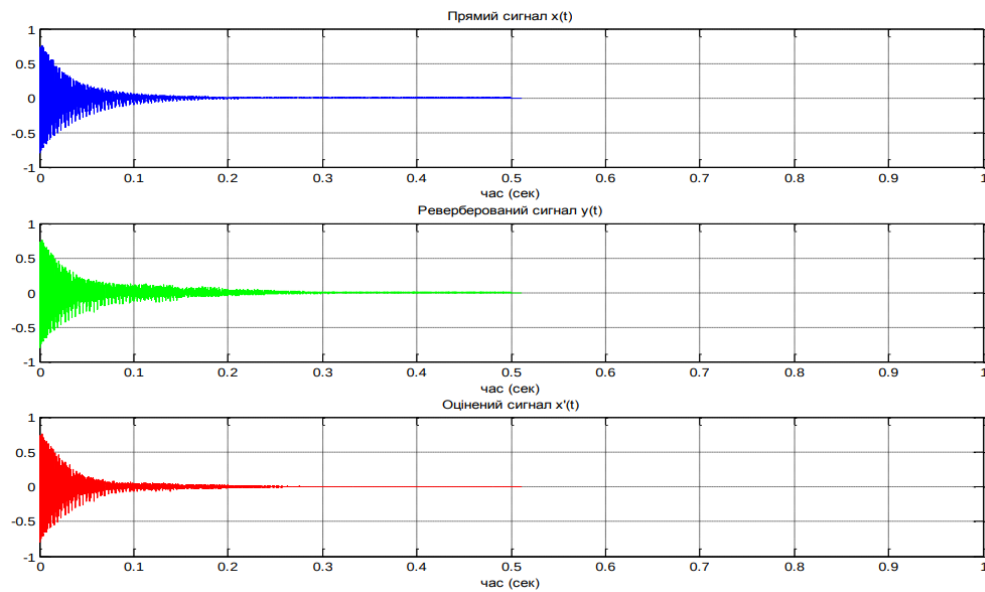


Рис. 3.3 Відображення сигналів у часовому просторі

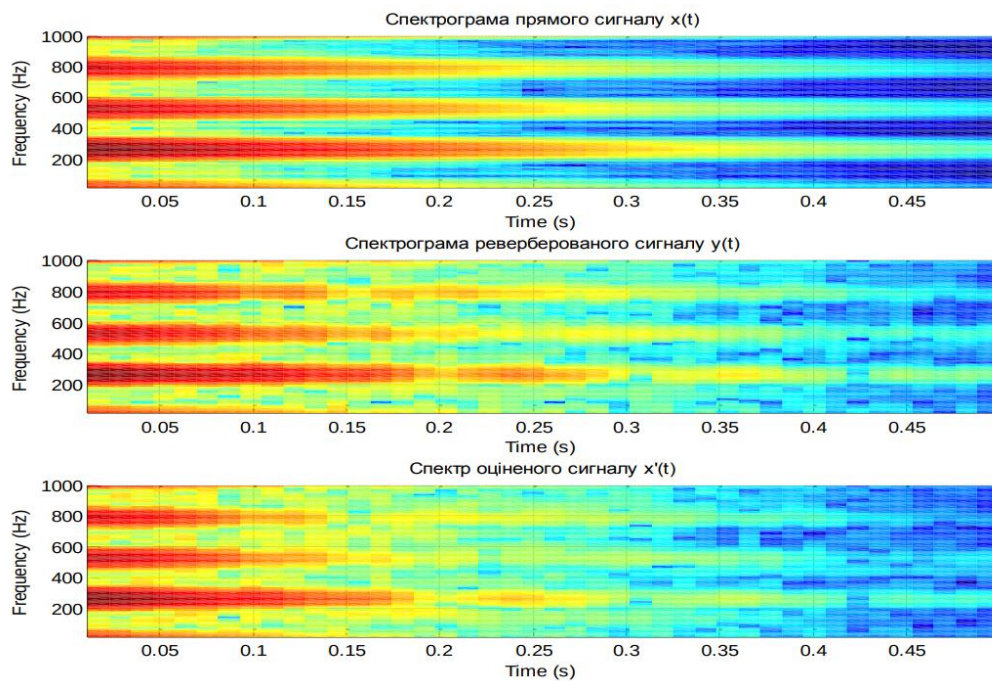


Рис. 3.4 Спектрограма сигналів

3.3 Моделювання мовного сигналу

Оскільки використання моделювання на штучно створених сигналах призначене лише для ілюстрації роботи алгоритму і не може служити перевіркою його якості, розглянемо сценарій для реального мовного сигналу. Для створення штучної реверберації ми також застосуємо згортку сигналу з реальною ІХС (див. рис. 3.5-3.7).

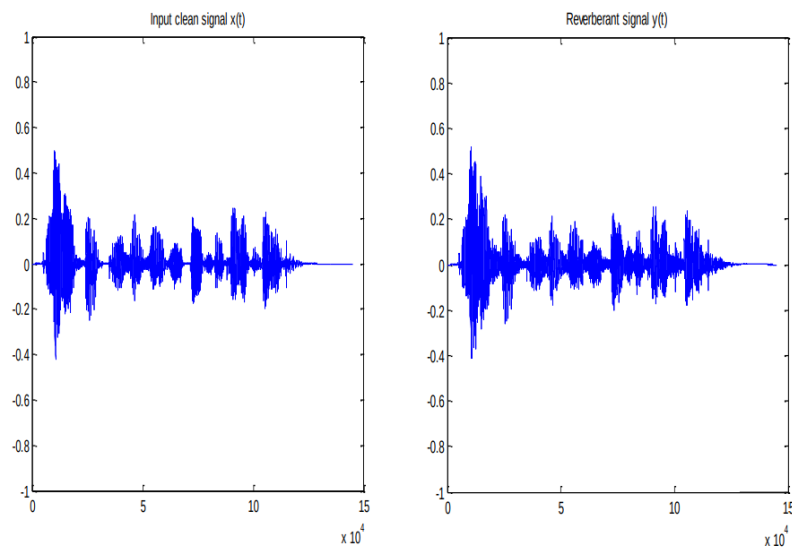


Рис. 3.5 Початковий аудіосигнал (ліворуч) та сигнал із реверберацією (праворуч)

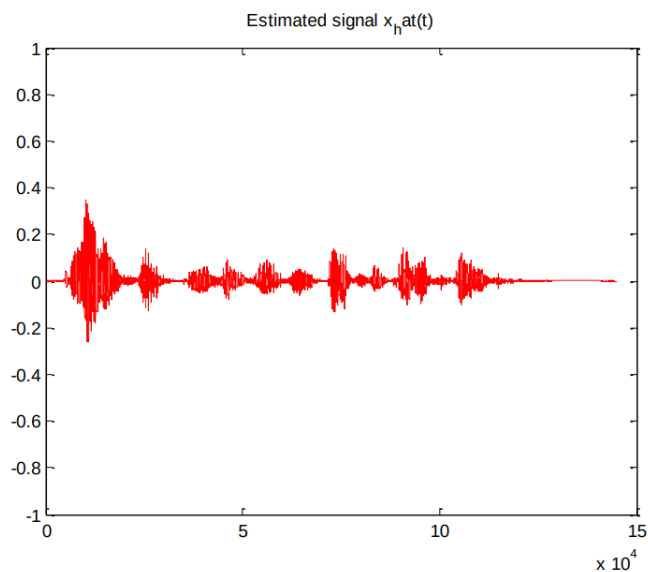


Рис. 3.6 Очищений сигнал

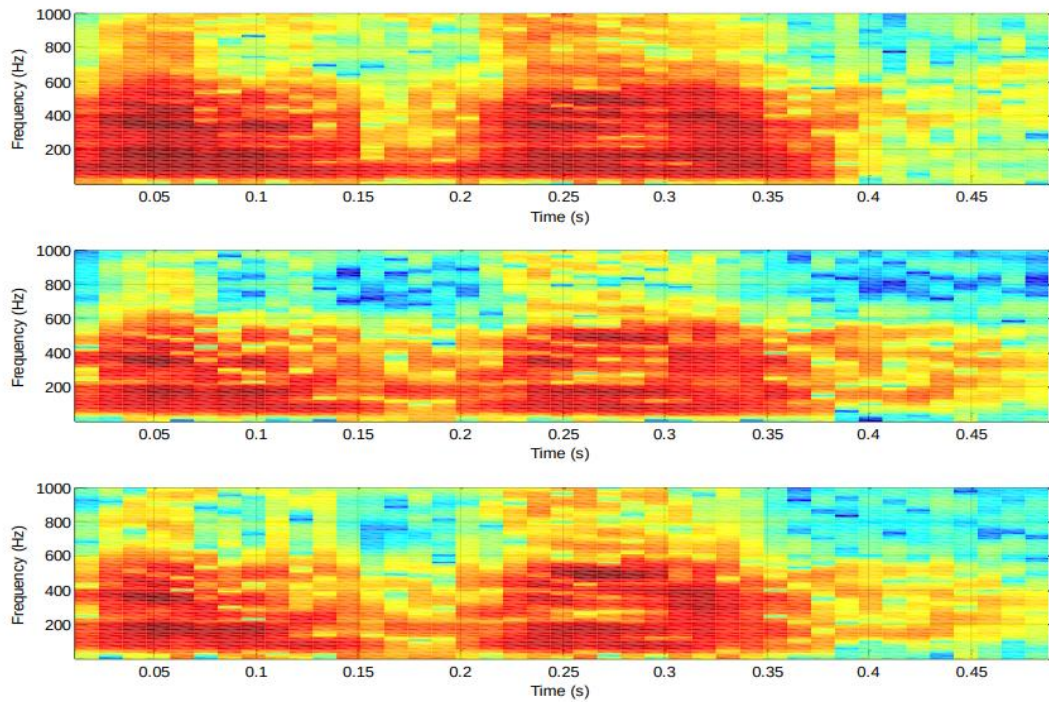


Рис. 3.7 Спектрограма початкового, реверберованого та вихідного сигналів

3.4 Аналіз роботи алгоритму в залежності від вхідних параметрів

Проаналізовано функціонування алгоритму залежно від зміни вхідних параметрів (див. рис. 3.8-3.11). Експерименти здійснювались за допомогою мовного сигналу та генератора штучної реверберації Герцена. Для вимірювання різниці між вхідним сигналом без реверберації та сигналом після застосування алгоритму усунення реверберації використовувалась СКВ, усереднене за тривалістю сигналу. Слід зауважити, що завжди існує певний мінімальний рівень, оскільки у сигналі все ще присутня початкова реверберація.

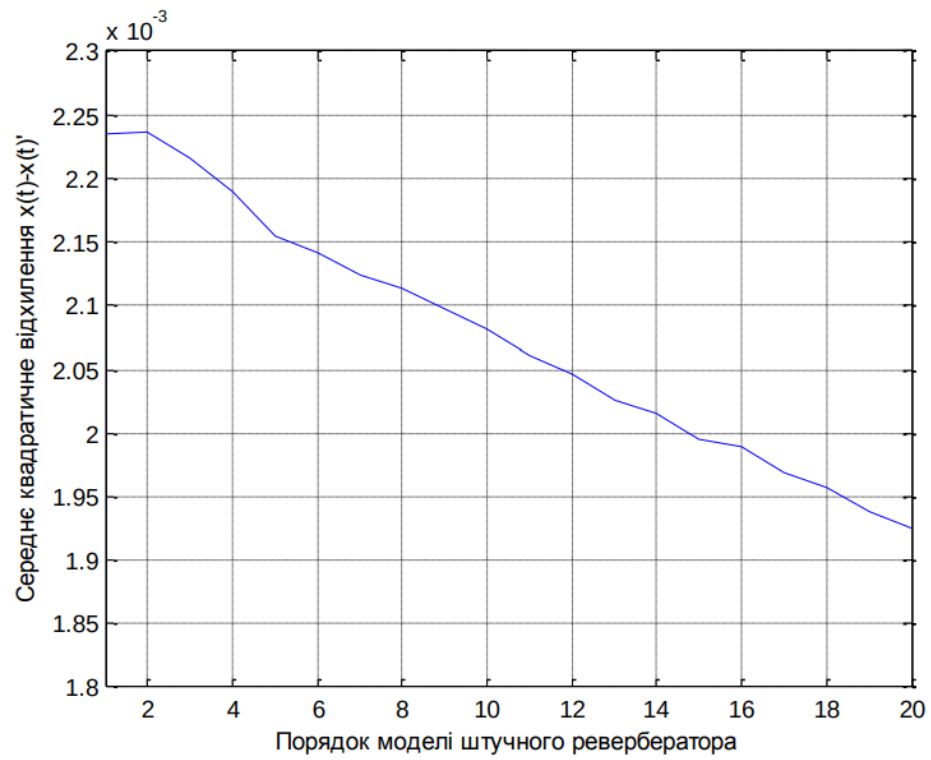


Рис. 3.8 Графік залежності СКВ від порядку моделі штучного ревербератора

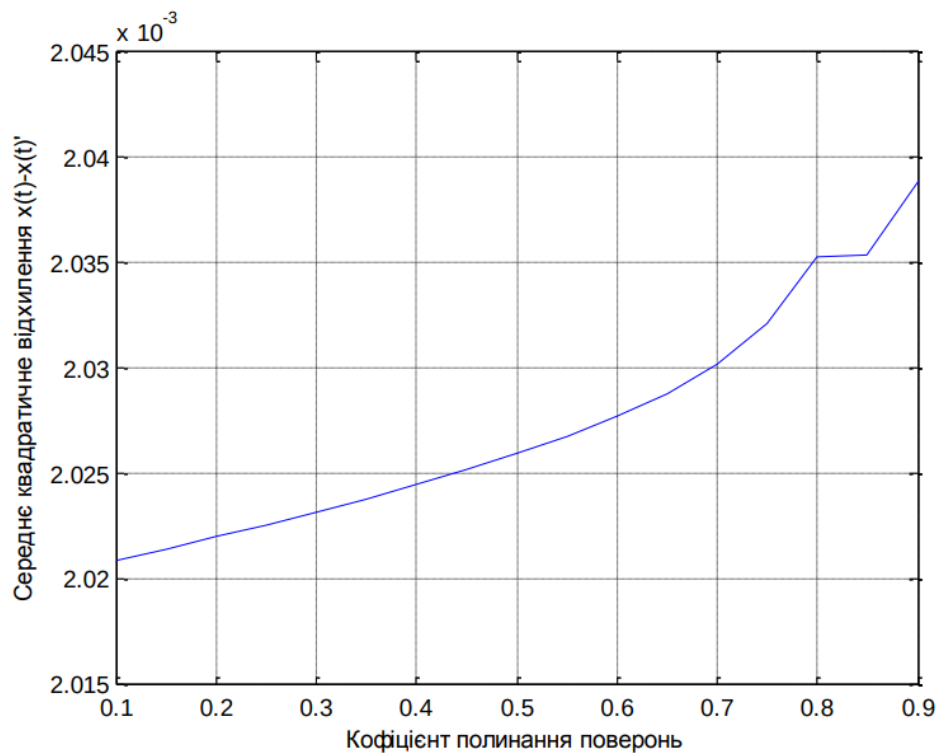


Рис. 3.9 Графік залежності СКВ від коефіцієнта поглиблення поверхонь середовища

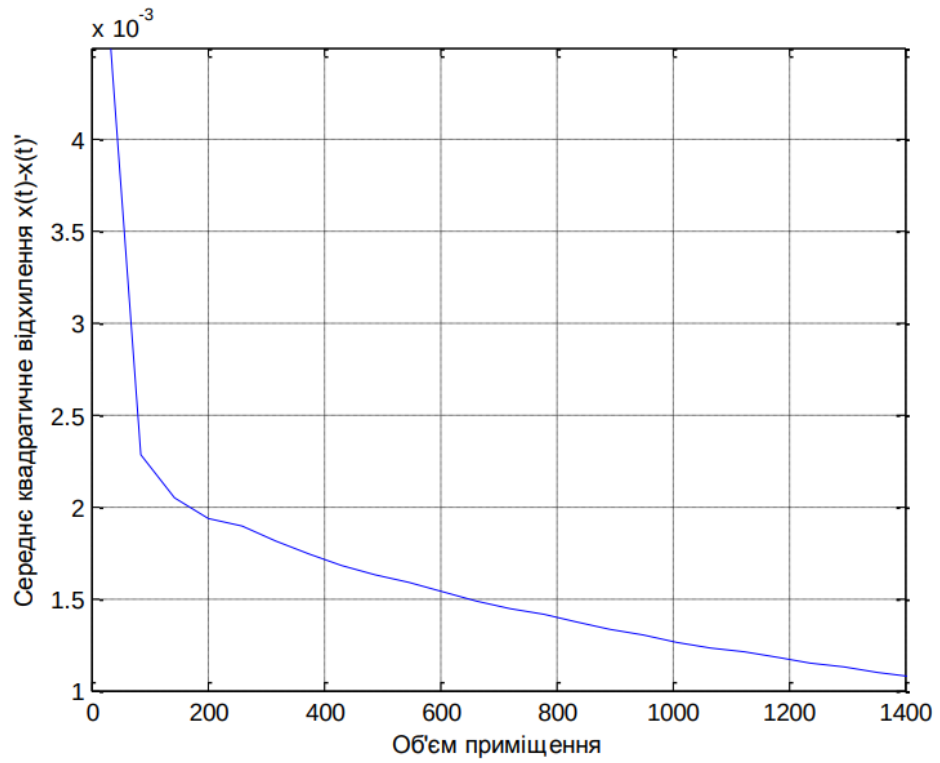


Рис. 3.10 Залежність СКВ від об'єму середовища

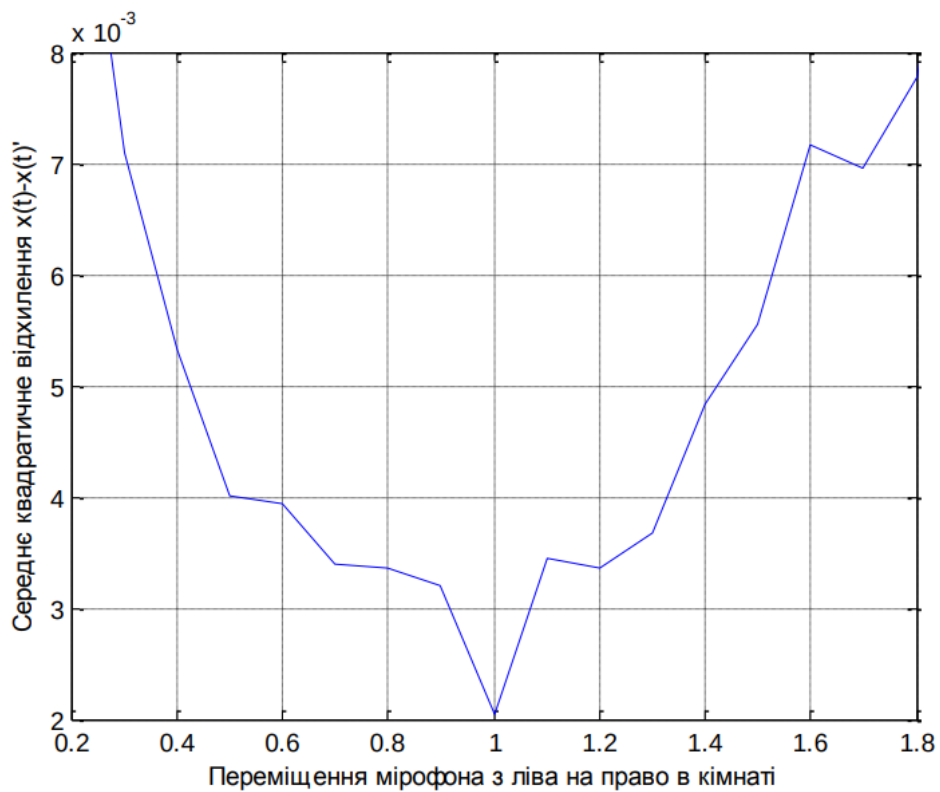


Рис. 3.11 Графік залежності СКВ від розташування мікрофона в середовищі

4. Висновок до розділу 3

У даному розділі виконано процедуру моделювання функціонування алгоритму, використовуючи як синтетичні, так і реальні сигнали. На основі результатів моделювання встановлено, що розроблений алгоритм ефективно усуває пізню реверберацію, залишаючи лише мінімальні залишки у відфільтрованій реалізації сигналу ранньої реверберації. Додатково, проведено аналіз функціонування алгоритму в залежності від зміни вхідних параметрів, які є важливими для реверберації. Це дозволило оцінити ефективність алгоритму в різноманітних умовах.

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

4.1. Охорона праці

Створення оптимальних комфортних умов у виробничих приміщеннях по виготовленню радіотехнічного обладнання є складною задачею, вирішити яку можна наступними заходами та засобами:

- Удосконалення технологічних процесів та устаткування.
- Впровадження нових технологій та обладнання, які не пов'язані з необхідністю проведення робіт в умовах інтенсивного нагріву дасть можливість зменшити виділення тепла у виробничі приміщення.
- Раціональне розміщення технологічного устаткування. Основні джерела теплоти бажано розміщувати безпосередньо під аераційним ліхтарем, біля зовнішніх стін будівлі і в один ряд на такій відстані один від одного, щоб теплові потоки від них не перехрещувались на робочих місцях. Для охолодження гарячих виробів необхідно передбачити окремі приміщення. Найкращим рішенням є розміщення тепловипромінюючого обладнання в ізольованих приміщеннях або на відкритих ділянках.
- Автоматизація та дистанційне управління технологічними процесами. Цей захід дозволяє в багатьох випадках вивести людину із виробничих зон, де діють несприятливі фактори.
- Раціональна вентиляція, опалення та кондиціонування повітря. Вони є найбільш розповсюдженими способами нормалізації мікроклімату у виробничих приміщеннях. Так зване повітряне та водоповітряне душення використовується у боротьбі з перегріванням робітників в гарячих цехах.

Забезпечити нормальні теплові умови в холодний період року в надтогабаритних та полегшених промислових будівлях дуже важко і економічно недоцільно. Найбільш раціональним варіантом в цьому випадку є застосування променистого нагрівання постійних робочих місць та окремих

дільниць. Захист від протягів досягається шляхом щільного закривання вікон, дверей та інших отворів, а також влаштуванням повітряних і повітряно-теплових завіс на дверях і воротах.

Раціоналізація режимів праці та відпочинку досягається скороченням тривалості робочої зміни, введенням додаткових перерв, створенням умов для ефективного відпочинку в приміщеннях з нормальними метеорологічними умовами. Якщо організувати окреме приміщення важко, то в гарячих цехах створюють зони відпочинку – охолоджувальні альтанки, де засобами вентиляції забезпечують нормальні температурні умови.

Для робітників, що працюють на відкритому повітрі зимою, обладнують приміщення для зігрівання, в яких температуру підтримують дещо вищою за комфортну. Застосування теплоізоляції устаткування та захисних екранів В якості теплоізоляційних матеріалів широко використовуються: азбест, азбоцемент, мінеральна вата, склотканина, керамзит, пінопласт.

На виробництві застосовують також захисні екрани для відгородження джерел теплового випромінювання від робочих місць. За принципом захисту щодо дії тепла екрани бувають відбиваючі, поглинаючі, відвідні та комбіновані. Хороший захист від теплового випромінювання здійснюють водяні завіси, що широко використовуються в металургії.

Важливе значення для уникнення перегрівання мають індивідуальні засоби захисту. Спецодяг повинен бути повітряно- та вологопроникним (бавовняним, з льону, грубововняного сукна), мати зручний покрій. Для роботи в екстремальних умовах застосовуються спеціальні костюми з підвищеною теплосвітловіддачею. Для захисту голови від випромінювання застосовують дюралеві, фіброві каски, повстяні капелюхи; для захисту очей – окуляри, темні або з прозорим шаром металу, маски з відкидним екраном. Захист від дії зниженої температури досягається використанням теплового спецодягу, а під час опадів – плащів та гумових чобіт.

Вимоги виробничої санітарії обов'язково враховують при проектуванні виробничих приміщень, процесів та технологічного обладнання. При цьому

зазначається, що в процесі виробництва приладу повинні бути відсутніми або мінімальними (нижче допустимих значень):

- виділення у повітря приміщень, атмосферу та стічні води шкідливих або з неприємним запахом речовин, а також виділення теплоти і вологості в робочих приміщеннях;
- шум, вібрація, ультразвук, електромагнітні хвилі радіочастот, статична електрика та іонізуючі випромінювання.

Відповідно до Закону України "Про охорону праці" на роботах із шкідливими та небезпечними умовами праці, в особливих температурних умовах, в забрудненому середовищі робітникам та службовцям безплатно видаються спецодяг, спецвзуття та інші засоби індивідуального захисту. Перелік робіт та професій, що дають право на одержання ЗІЗ, складається на основі галузевих норм адміністрацією підприємства та погоджується із місцевими органами Державної служби України з питань праці. Порядок видачі, зберігання та використання спеціального одягу, взуття та інших засобів індивідуального захисту визначається НПАОП 0.00-4.01-08 "Положення про порядок забезпечення працівників спеціальним одягом, спеціальним взуттям та іншими засобами індивідуального захисту".

4.2. Безпека в надзвичайних ситуаціях

Порядок дій у разі виникнення надзвичайних ситуацій.

Хімічна небезпека

Хімічне забруднення – розповсюдження небезпечних хімічних речовин у навколишньому середовищі в концентраціях чи кількостях, що протягом певного часу створюють загрозу життю та здоров'ю людей і/або негативно впливають на навколишнє природне середовище [].

Аварії (катастрофи) на підприємствах, транспорті 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.

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.

Дії у випадку 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 залиште зону хімічного забруднення.

4. Якщо засобів індивідуального захисту немає і вийти із району аварії неможливо, залишайтеся у приміщенні і негайно та надійно герметизуйте приміщення. Зменште 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 у герметичну валізу та підготуйтеся до евакуації.

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

8. Залишаючи приміщення 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 місце 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

Радіоактивне забруднення – забруднення навколишнього середовища радіоактивними речовинами, вміст яких перевищує рівень, встановлений згідно зі стандартами, нормами і правилами радіаційної безпеки [].

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 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, 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

Пожежа – неконтрольований процес знищування або пошкодження вогнем майна, під час якого виникають чинники, небезпечні для істот та навколишнього природного середовища [].

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.

4.3 Висновки до розділу 4

У підрозділі з охорони праці проаналізовано оптимальні комфортні умови у виробничих приміщеннях по виготовленню радіотехнічного обладнання.

У підрозділі з безпеки в надзвичайних ситуаціях наведено та описано порядок дій при виникненні хімічної небезпеки, радіаційної небезпеки, пожежі, повені та паводків.

ВИСНОВКИ

У кваліфікаційній роботі освітнього рівня «Магістр» вирішено важливу наукову проблему, а саме розроблено підхід для вдосконалення ефективності виявлення акустичних сигналів у ЗС.

Отримані результати включають:

– Аналіз фізичних характеристик та параметрів акустичних сигналів. На основі цього аналізу визначено основні вимоги до впровадження ефективного методу покращення якості виявлення аудіосигналів у ЗС. Однією із цих вимог є ліквідація ревербераційних ефектів, зокрема приглушення кінцевого звукового післязвучання.

– Вдосконалено метод та алгоритм для створення фальшивої реверберації з визначеним часом реверберації. Це було виконано для проведення дослідження процесу виявлення аудіосигналів у ЗС.

– Вдосконалено підхід покращення якості виявлення аудіосигналів у ЗС, використовуючи МСВ, модель Полака та метод Дебніха. Проаналізовано його ефективність, що призвело до підвищення точності розпізнавання акустичних сигналів у ЗС шляхом пригнічення пізньої реверберації.

– Спроектовано ПЗ для проведення досліджень щодо процесу виявлення аудіосигналів у ЗС.

ПЕРЕЛІК ДЖЕРЕЛ

- 1 F.Everest. The Master Handbook of Acoustics. McGraw-Hill, 1994.
- 2 Encyclopaedia Britannica Online. Acoustics. 2007.
- 3 P. A. Naylor and N. D. Gaubitch, *Speech Dereverberation*. Berlin, Germany: Springer, 2010.
- 4 E.A.P Habets, “Single- and multi-microphone speech dereverberation using spectral enhancement,” Ph.D. dissertation, T. U. Eindhoven, 2007.
- 5 Kuttruff, H.: room acoustics, 4 edn. Taylor & Francis (2000).
- 6 Savioja, L., Rinne, T.J., Takala, T.: Simulation of room acoustics with a 3-D finite difference mesh. In: Proc. Int. Computer Music Conf., pp. 463–466. Denmark (1994).
- 7 Kleiner, M., Dalenb.ack, B., Svensson, P.: Auralization – an overview. J. Acoust. Soc. Am. 41(11), 861–875 (1993).
- 8 Wen, J.Y.C., Gaubitch, N.D., Habets, E.A.P., Myatt, T., Naylor, P.A.: Evaluation of speech dereverberation algorithms using the MARDY database. In: Proc. Int. Workshop Acoust. Echo Noise Control (IWAENC). Paris, France (2006).
- 9 Kuttruff, K.H., Schroeder, M.R.: On frequency response curves in rooms. comparison of experimental, theoretical, and monte carlo results for the average frequency spacing between maxima. J. Acoust. Soc. Am. 34(1), 76 – 80 (1962).
- 10 Savioja, L.: Modeling techniques for virtual acoustics. Doctoral dissertation, Helsinki University of Technology, Espoo, Finland (1999).
- 11 Kleiner, M., Dalenb.ack, B., Svensson, P.: Auralization – an overview. J. Acoust. Soc. Am. 41(11), 861–875 (1993).
- 12 Savioja, L., Backman, J., Jarvinen, A., Takala, T.: Waveguide mesh method for low-frequency simulation of room acoustics. In: Proc. of the 15th Int. Congr. Acoust. (ICA‘95), vol. 2, pp. 1–4. Trondheim, Norway (1995).
- 13 Krokstad, A., Strom, S., Sorsdal, S.: Calculating the acoustical room response by the use of a ray tracing technique. J. Sound Vib. 8, 118–125 (1968).

- 14 Kulowski, A.: Error investigation for the ray tracing technique. *Appl. Acoust.* 15(4), 263–274 (1982).
- 15 Allen, J.B., Berkley, D.A.: Image method for efficiently simulating small-room acoustics. *J. Acoust. Soc. Am.* 65(4), 943–950 (1979).
- 16 Peterson, P.M.: Simulating the response of multiple microphones to a single acoustic source in a reverberant room. *J. Acoust. Soc. Am.* 80(5), 1527–1529 (1986).
- 17 Nelson, P.A., Elliott, S.J.: *Active control of sound*. Academic, London (1993).
- 18 Polack, J.D.: *La transmission de l'énergie sonore dans les salles*. Thèse de doctorat d'état, Université du Maine, Le Mans (1988).
- 19 Gustafsson, T., Rao, B.D., Trivedi, M.: Source localization in reverberant environments: modeling and statistical analysis. *IEEE Trans. Speech Audio Process.* 11(6), 791803 (2003).
- 20 Reichardt, W., Lehmann, U.: Raumeindruck als oberbegriff von rumlichkeit und halligkeit, erluterungen des raumeindrucks masses. *Acustica* 40, 174–183 (1978).
- 21 Polack, J.D.: Playing billiards in the concert hall: the mathematical foundations of geometrical room acoustics. *Appl. Acoust.* 38(2), 235–244 (1993).
- 22 Pietrzyk, A.: Computer modeling of the sound field in small rooms. In: *Proc. of the 15th AES Int. Conf. on Audio, Acoustics and Small Spaces*, vol. 2, pp. 24–31. Copenhagen, Denmark (1998).
- 23 Gardner, W.G. 1998. Chapter 3. Reverberation Algorithms, in Kahrs, M. And Brandenburg, K. Editors. *Applications of Digital Signal Processing to Audio and Acoustics*. Kluwer Academic Publishers.
- 24 Schroeder, M. R., Logan, B. F. 1961. Colorless Artificial Reverberation. *J. Audio Engineering Society*. Vol. 9, No. 3.
- 25 E.A.P Habets, “Single- and multi-microphone speech dereverberation using spectral enhancement,” Ph.D. dissertation, T. U. Eindhoven, 2007.

- 26 J. Hardwick, C.D. Yoo, and J.S. Lim, "Speech enhancement using the dual excitation speech model," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 1993, pp. 367–370.
- 27 C.D. Yoo, "Speech enhancement based on the generalized dual excitation model with adaptive analysis window," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 1995, vol. 1, pp. 832–835.
- 28 H. Attias and L. Deng, "Speech Denoising and Dereverberation Using Probabilistic Models," *Advances in Neural Inform. Process. Syst.*, vol. 13, pp. 758–764, 2001.
- 29 J.R. Deller, J.G. Proakis, and J.H.L. Hansen, *Discrete-Time Processing of Speech Signals*. New York: MacMillan, 1993.
- 30 M. Tonelli, M.G. Jafari, and M.E. Davies, "A multi-channel Maximum Likelihood approach to de-reverberation," in *Proc. European Signal Process. Conf. (EUSIPCO)*, Florence, Italy, Sept. 2006.
- 31 B. Yegnanarayana, "Speech enhancement using excitation source information," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2002, vol. 1, pp. 541–544.
- 32 H. Wang and F. Itakura, "An implementation of multi-microphone dereverberation approach as a preprocessor to the word recognition system," *J. Acoust. Soc. Jap.*, vol. 13, no. 5, pp. 285–293, 1992.
- 33 A. Oppenheim, R. Schafer, T. Stockham, "Nonlinear filtering of multiplied and convolved signals", *IEEE Trans. Audio and Electroacoustics*, Vol. AU-16, No. 3, 1968.
- 34 M. Wu and D. L. Wang, "A two-stage algorithm for one microphone reverberant speech enhancement," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 14, no. 3, pp. 774–784, May 2006.
- 35 P. Krishnamoorthy and S. R. M. Prasanna, "Reverberant speech enhancement by temporal and spectral processing," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 17, no. 2, pp. 137–148, Feb. 2009.

36 S. Mosayyebpour, M. Esmaili, T. A. Gulliver, “Single-Microphone Early and Late Reverberation Suppression in Noisy Speech,” *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 21, no. 2, pp. 322–335, Feb. 2013.

37 E. A. P. Habets, N. Gaubitch, and P. A. Naylor, “Temporal selective dereverberation of noisy speech using one microphone,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2008, pp. 4577–4580.

38 J. S. Erkelens and R. Heusdens, “Single-microphone late-reverberation suppression in noisy speech by exploiting long-term correlation in the DFT domain,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2009, pp. 3997–4000.

39 N.D. Gaubitch, P.A. Naylor, and D. Ward, “On the use of linear prediction for dereverberation of speech,” in *Proc. of the Int. Workshop Acoust. Echo Noise Control (IWAENC)*, Kyoto, Japan, 2003, pp. 99–102.

40 N.D. Gaubitch, D.B. Ward, and P.A. Naylor, “Statistical analysis of the autoregressive modeling of reverberant speech,” *J. Acoust. Soc. Am.*, vol. 120, pp. 4031–4039, Dec. 2006.

41 D. Gesbert and P. Duhamel, “Robust blind identification and equalization based on multi-step predictors,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 1997, vol. 26, no. 5, pp. 3621–3624.

42 Khvostivska L., Khvostivskyi M., Dunets V., Dediv I. Matematychnе, alhorytmichne ta prohramne zabezpechennia synfaznoho vyivlennia radiosyhnaliv v elektronnykh komunikatsiinykh merezhakh iz zavadamy [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, pp. 48-57 [in Ukrainian].

43 Хвостівський М.О., Дунець В.Л., Дедів І.Ю. Методичні рекомендації з оформлення кваліфікаційних робіт магістра за спеціальністю 172 Телекомунікації та радіотехніка. Тернопіль: ТНТУ імені Івана Пулюя, 2020. 21 с.

ДОДАТКИ

Лістинг програми для алгоритму подолання реверберації

Лістинг А.1 – tawfAlgorithm.m

```
function [x_hat,G] = tawfAlgorithm(a, Rt, T, Y)
tawfConstants;
% Reverb model parameters N =
round(T*fs/frameLen); if N==0
N=1;
end
delta = 3 * log(10) / Rt;
% Dereverberation algorithm Pyy =
tawfEstimatePSD(Y, a);
Prr = tawfEstimateReverbPower(1*exp(-2*delta*T),
Pyy, N);

% Calculate gain coefficients and threshold (Eqn
7.) G = sqrt((Pyy - Prr)./Pyy);
G(G < thr) = thr; G(isnan(G)) = thr;
% for i = 1:frameLen
%           G(i,:) = smooth(G(i,:))';
% end
% Apply gain to STFT to scale out reverb power W
= (G .* Y);
% Reconstruct Dereverberated Signal
x_hat = tawfInverseSTFT(W, winFunc); % IFFT &
Overlap Add end
```

Лістинг А.2 – tawfEstimatePSD.m

```
function [estPSD] = tawfEstimatePSD(signalSTFT,
weight)
%tawfEstimatePSD Estimation of the PSD of a
supplied signalSTFT
%           Estimates the PSD of the input STFT
using Eqn (8) from Shi & Ma.
%           Uses a running average the weights
previous frame powers using the
%           scaling factor specified by 'weight'
estPSD = zeros(size(signalSTFT));
%scf = (1/2/pi/size(signalSTFT,1)); estPSD(:,1)
= abs(signalSTFT(:,1)).^2; for ii =
2:size(estPSD,2)
estPSD(:,ii) = weight * estPSD(:,ii-1) + (1-
weight) * abs(signalSTFT(:,ii)).^2; end

end
```

Лістинг А.3 – tawfEstimateReverbPower.m

```
function [reverbPower] =
tawfEstimateReverbPower(reverbModel, signalPSD,
N)
%UNTITLED4 Summary of this function goes here
%           Detailed explanation goes here

reverbPower = zeros(size(signalPSD));
reverbPower(:,N+1:end) = reverbModel *
signalPSD(:,1:end-N);

end
```

Лістинг А.4 – tawfSTFT.m

```
function [Y] = tawfSTFT(y, frameLen, overlapLen,
window)
%UNTITLED6 Summary of this function goes here
%           Buffer data, apply window and compute
the fft

yb = buffer(y, frameLen, overlapLen, 'nodelay');
winFunc = repmat(window, 1, size(yb, 2));
yb = yb .* winFunc;
Y = fft(yb);

end
```

Лістинг А.5 – tawfInverseSTFT.m

```
function [x_hat] = tawfInverseSTFT(W, winFunc)
%UNTITLED3 Summary of this function goes here
%           Detailed explanation goes here

xb_hat = ifft(W);
invWinFunc =
repmat(1./winFunc, 1, size(xb_hat, 2)); xb_hat =
xb_hat .* invWinFunc;

m = size(xb_hat, 1); n = size(xb_hat, 2);

%% Overlap-add helpers - assumes 50% overlap

% TODO: odd number m upperMat = xb_hat(1:m/2, :);

lowerMat = xb_hat(m/2+1:end, :);

leftCol = [upperMat(:); zeros(m/2, 1)];
leftCol(1:m/2) = 2 * leftCol(1:m/2); rightCol =
[zeros(m/2, 1); lowerMat(:)];
rightCol(end-m/2:end) = 2 * rightCol(end-
```



```
m/2:end); x_hat = mean([leftCol,rightCol],2);
end
```

Лістинг А.6 – tawfGetMSE.m

```
function mse = tawfGetMSE(x,x_hat) mse =
mean((x-x_hat).^2);
end
```

Лістинг А.7 – tawfGenerateRIR.m

```
function [h] = tawfGenerateRIR(fs, mic, n, r,
rm, src)
%RIR          Room Impulse Response.
%          [h] = RIR(FS, MIC, N, R, RM, SRC)
performs a room impulse
%          response calculation by means of
the mirror image method.
%
%          FS = sample rate.
%          MIC = row vector giving the x,y,z
coordinates of
%          the microphone.
%          N = The program will account for
(2*N+1)3 virtual sources
%          R = reflection coefficient for the
walls, in general -1<R<1.

%          RM = row vector giving the
dimensions of the room.
%          SRC = row vector giving the x,y,z
coordinates of
%          the sound source.
%          EXAMPLE:
%
%          >>fs=44100;
%          >>mic=[19 18 1.6];
%          >>n=12;
%          >>r=0.3;
%          >>rm=[20 19 21];
%          >>src=[5 2 1];
%          >>h=rir(fs, mic, n, r, rm, src);

nn=-n:1:n;          % Index for the
sequence rms=nn+0.5-0.5*(-1).^nn; % Part of
equations 2,3,& 4 srcs=(-1).^(nn); % part of
equations 2,3,& 4 xi=srcs*src(1)+rms*rm(1)-
mic(1);          % Equation 2
yj=srcs*src(2)+rms*rm(2)-mic(2); % Equation
3 zk=srcs*src(3)+rms*rm(3)-mic(3); % Equation
4
```

```
[i,j,k]=meshgrid(xi,yj,zk);          % convert
vectors to 3D matrices d=sqrt(i.^2+j.^2+k.^2);
                                          % Equation 5
time=round(fs*d/343)+1;              % Similar to
Equation 6

[e,f,g]=meshgrid(nn, nn, nn);        % convert
vectors to 3D matrices
c=r.^(abs(e)+abs(f)+abs(g));          % Equation 9
e=c./d;                               % Equivalent to
Equation 10

h=full(sparse(time(:),1,e(:)));      % Equivalent
to equation 11 h=h/max(abs(h)); % Scale output
```

Лістинг програми для генерації штучної реверберації

Лістинг Б.1 – delaynetwork.m

```

%fs=44100;
gain=0.97;
% Create an impulse
x = zeros(1,1*fs); x(1) = 1;
[x,fs]=wavread('msspl.wav'); y = zeros(1,fs);
b = [1 1 1 1];
c = [0.8 0.8 0.8 0.8];
% Feedback matrix a(1,:)= [0 1 1 0];
a(2,:)= [-1 0 0 -1];
a(3,:)= [1 0 0 -1];
a(4,:)= [0 1 -1 0];
a2=a*(1/sqrt(2)) * gain;
% Delay lines, use prime numbers m=[149 211 263 293]';
z1=zeros(1,max(max(m)));
z2=zeros(1,max(max(m)));
z3=zeros(1,max(max(m)));
z4=zeros(1,max(max(m))); for n = 1:length(y)
tmp = [z1(m(1)) z2(m(2)) z3(m(3)) z4(m(4))];
y(n) = x(n) + c(1)*z1(m(1)) + c(2)*z2(m(2)) ...
+ c(3)*z3(m(3)) + c(4)*z4(m(4));
z1 = [(x(n)*b(1) + tmp*a2(1,:))' z1(1:length(z1)-1)];
z2 = [(x(n)*b(2) + tmp*a2(2,:))' z2(1:length(z2)-1)];
z3 = [(x(n)*b(3) + tmp*a2(3,:))' z3(1:length(z3)-1)];
z4 = [(x(n)*b(4) + tmp*a2(4,:))' z4(1:length(z4)-1)]; end
plot(y) grid on; figure(2);
y = spectrogram(y, 16); imagesc(abs(y))

```

Лістинг Б.2 – Schroeder1.m

```

%A proposed natural sounding reveberator a1=0.6;
a2=0.4; a3=0.2; a4=0.1; a5=0.8; a6=0.8; a7=0.8; R1=2000; R2=4000;

```

```

R3=1000; R4=500; R5=2000;
R6=2000;
d = zeros(1,1*60000); d(1) = 1;
%[d,r]=wavread('mssp1.wav');
num1=[0,zeros(1,R1-1),1];
den1=[1,zeros(1,R1-1),-a1];
d1=filter(num1,den1,d); num2=[0,zeros(1,R2-1),1];
den2=[1,zeros(1,R2-1),-a2];
d2=filter(num2,den2,d); num3=[0,zeros(1,R3-1),1];
den3=[1,zeros(1,R3-1),-a3];
d3=filter(num3,den3,d); num4=[0,zeros(1,R4-1),1];
den4=[1,zeros(1,R4-1),-a4];
d4=filter(num4,den4,d); dIIR=d1+d2+d3+d4;
num5=[a5,zeros(1,R5-1),1];
den5=[1,zeros(1,R5-1),a5];
dALL1=filter(num5,den5,dIIR); num5=[a6,zeros(1,R6-1),1];
den5=[1,zeros(1,R6-1),a6];
dALL2=filter(num5,den5,dALL1); dTOTAL=d+a7*dALL2;
%soundsc(dTOTAL,r); x=d;
[outcomb1,b1,a1] = JIPComb(x,0.6,0.7,0.1);
[outcomb2,b2,a2] = JIPComb(x,0.6,0.6,0.2);
[outcomb3,b3,a3] = JIPComb(x,0.6,0.5,0.3);
[outcomb4,b4,a4] = JIPComb(x,0.6,0.4,0.4);
[outcomb5,b5,a5] = JIPComb(x,0.6,0.3,0.5);
[outcomb6,b6,a6] = JIPComb(x,0.6,0.2,0.6);
% sum the ouptut of the 4 comb filters
apinput = outcomb1 + outcomb2 + outcomb3 + outcomb4 + outcomb5 +
outcomb6;
%find the combined filter coefficients of the the comb filters
[b,a]=parallelcoefficients(b1,a1,b2,a2);
[b,a]=parallelcoefficients(b,a,b3,a3);
[b,a]=parallelcoefficients(b,a,b4,a4);
[b,a]=parallelcoefficients(b,a,b5,a5);
[b,a]=parallelcoefficients(b,a,b6,a6);
% send the output of the comb filters to the allpass filter

```

```
[y,b7,a7] = allpass(apinput,ag,ad);  
%find the combined filter coefficients of the the comb filters in  
series with the allpass filters  
[b,a]=seriescoefficients(b,a,b7,a7);  
% add the scaled direct signal  $y = y + k*x$ ;  
% normalize the output signal  $y = y/\max(y)$ ;  
plot(y)  
grid on
```