

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації
(повна назва)

Кафедра медіаінженерії та інформаційних радіоелектронних систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти другий (магістерський)

Методи подавлення квазістаціонарних шумів у звукозаписі
(тема)

Виконав:
студент 2 курсу, групи МІМ-20-1
Медінцев Р.В.
(прізвище, ініціали)

Спеціальність 172 Телекомунікації та
радіотехніка
(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Медіаінженерія
(повна назва освітньої програми)

Керівник проф. Тихонов В.А.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Карташов В.М.
(прізвище, ініціали)

2021 р.

Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації

Кафедра медіаінженерії та інформаційних радіоелектронних систем

Рівень вищої освіти другий (магістерський)

Спеціальність 172 Телекомунікації та радіотехніка

(код і повна назва)

Тип програми освітньо-професійна

(освітньо-професійна або освітньо-наукова)

Освітня програма Медіаінженерія

(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____

(підпис)

«_____» _____ 20__ р.

ЗАВДАННЯ НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Медінцеву Руслану Володимировичу

(прізвище, ім'я, по батькові)

1. Тема роботи Методи подавлення квазістаціонарних шумів у звукозаписі.

затверджена наказом по університету від " 08 " 11 2021 р. № 1676 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 08.12.2021 р.

3. Вихідні дані до роботи Провести дослідження базових методів шумоподавлення. Методом натурного експерименту дослідити багатосмуговий шумоподавлювач. Отримати числові характеристики поліпшення відношення сигнал-шум. Провести теоретичний аналіз методу спектрального віднімання. Дослідити шумоподавлення на основі спектрального віднімання у звукових редакторах. Отримати спектрограми зашумленого та обробленого сигналів. Провести аналіз результатів досліджень.

4. Перелік питань, що потрібно опрацювати в роботі _____

Вступ

1. Аналіз характеристик і сприйняття, причин зашумлення і алгоритмів обробки звукових сигналів.

2. Дослідження базових алгоритмів шумоподавлення у звукових сигналах.

3. Дослідження алгоритмів шумоподавлення з високим спектральним розрізненням.

4. Дослідження алгоритмів шумопогнікання на основі спектрального віднімання.

Висновки

Перелік посилань

Додатки

5. Перелік графічного матеріалу із зазначенням обов'язкових креслеників, схем, плакатів, комп'ютерних ілюстрацій _____

1. Проблемна ситуація (1 аркуш).

2. Постановка задачі(1 аркуш).

3. Пороговий шумоподавлювач (1 аркуш).

4. Дослідження порогового шумоподавлювача (1 аркуш).

5. Динамічний шумоподавлювач (1 аркуш).

6. Багатосмуговий гейт (1 аркуш).

7. Метод спектрального віднімання (1 аркуш).

8. Дослідження спектрального віднімання (1 аркуш).

9. Алгоритм мультирозрізнення (1 аркуш).

10. Зведена таблиця результатів (1 аркуш).

11. Висновки (1 аркуш).

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналітичний огляд літератури	02.11.21–10.11.21	
2	Теоретичне обґрунтування методу	11.11.21–17.11.21	
3	Дослідження базових методів	18.11.21–24.11.21	
4	Дослідження спектральних алгоритмів	25.11.21–1.12.21	
5	Дослідження спектрального віднімання	02.12.21–5.12.21	
6	Обробка результатів досліджень	25.11.21–5.12.21	
7	Графічна частина роботи	25.11.21–5.12.21	
8	Перевірка керівником	5.12.21-6.12.21	
9	Перевірка на академічний плагіат	7.12.21	
10	Перевірка завідувачем кафедри, рецензування	7.12.21–8.12.21	

Дата видачі завдання _____ 08.11.2021 р. _____

Студент _____

(підпис)

Керівник роботи (проекту) _____

(підпис)

проф. Тихонов В.А.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи: 65 сторінок, 41 рисунок, 9 таблиць, 32 джерела.

ГЕЙТ, ЗВУКОВИЙ СИГНАЛ, БАГАТОСМУГОВЕ
ШУМОПОДАВЛЕННЯ, СПЕКТРАЛЬНЕ ВІДНІМАННЯ, ШУМ,
ШУМОПОДАВЛЕННЯ

Об'єкт дослідження – процес подавлення широкосмугового шуму у звукових сигналах.

Мета дослідження – визначення числових і якісних характеристик різних алгоритмів шумоподавлення.

Розглянуто основні види шумів у звукових трактах і їх вплив на якісні показники звукозаписів. Проведено аналіз і дослідження базових методів шумоподавлення. Методом натурального експерименту досліджено багатосмуговий граничний шумоподавлювач із високим спектральним розрізненням. Отримано числові характеристики поліпшення відношення сигнал-шум для різних умов. Проведено аналіз методу спектрального віднімання. Досліджено методи шумоподавлення на основі спектрального віднімання у звукових редакторах. Отримано спектрограми зашумленого та обробленого сигналів, проведено аналіз результатів досліджень.

РЕФЕРАТ

Объяснительная записка к квалификационной работе: 65 страниц, 41 рисунок, 9 таблиц, 32 источника.

ГЕЙТ, ЗВУКОВОЙ СИГНАЛ, МНОГОПОЛСНОЕ
ШУМОПОДАВЛЕНИЕ, СПЕКТРАЛЬНОЕ ВЫЧИТАНИЕ, ШУМ,
ШУМОПОДАВЛЕНИЕ

Объект исследования – процесс подавления широкополосного шума в звуковых сигналах.

Цель исследования – определение числовых и качественных характеристик разных алгоритмов шумоподавления.

Рассмотрены основные виды шумов в звуковых трактах и их влияние на качественные показатели звукозаписей. Проведен анализ и исследование базовых методов шумоподавления. Методом натурального эксперимента исследован многополосный предельный шумоподавитель с высоким спектральным разрешением. Получены числовые характеристики улучшения отношения сигнал-шум для разных условий. Проведен анализ метода спектрального вычитания. Исследованы методы шумоподавления на основе спектрального вычитания в звуковых редакторах. Получены спектрограммы зашумленного и обработанного сигналов, проведен анализ результатов исследований.

ABSTRACT

Explanatory note to the qualification work: 65 pages, 41 figures, 9 tables, 32 sources.

GATE, SOUND SIGNAL, MULTI-BAND NOISE SUPPRESSION, SPECTRAL SUBTRACTION, NOISE, NOISE SUPPRESSION

The object of study – the process of suppressing broadband noise in audio signals.

The purpose of the study is to determine the numerical and qualitative characteristics of different noise reduction algorithms.

The main types of noise in sound paths and their influence on qualitative indicators of sound recordings are considered. The analysis and research of basic methods of noise suppression are carried out. A multi-band boundary noise suppressor with high spectral resolution was investigated by the method of field experiment. Numerical characteristics of signal-to-noise improvement for different conditions are obtained. The analysis of the spectral subtraction method is carried out. Noise suppression methods based on spectral subtraction in sound editors have been studied. Spectrograms of noisy and processed signals were obtained, analysis of research results was performed.

ЗМІСТ

Перелік умовних позначень, символів, одиниць, скорочень і термінів.....	9
Вступ.....	10
1 АНАЛІЗ ХАРАКТЕРИСТИК І СПРИЙНЯТТЯ, ПРИЧИН ЗАШУМЛЕННЯ І АЛГОРИТМІВ ОБРОБКИ ЗВУКОВИХ СИГНАЛІВ.....	12
1.1 Аудіосигнали та їх характеристики.....	12
1.2 Сприйняття звукових сигналів.....	15
1.3 Цифрове представлення звукових сигналів.....	19
1.4 Алгоритми роботи лінійних цифрових фільтрів.....	23
1.5 Шумові спотворення та їх причини.....	26
1.6 Висновки по розділу 1.....	27
2 ДОСЛІДЖЕННЯ БАЗОВИХ АЛГОРИТМІВ ШУМОПОДАВЛЕННЯ У ЗВУКОВИХ СИГНАЛАХ.....	29
2.1 Вимоги до шумів, що подавляються.....	29
2.2 Пороговий шумоподавлювач (гейт).....	29
2.3 Динамічний шумоподавлювач.....	32
2.4 Динамічний обмежувач шуму.....	34
2.5 Багатосмуговий гейт.....	36
2.6 Висновки по розділу 2.....	37
3 ДОСЛІДЖЕННЯ АЛГОРИТМІВ ШУМОПОДАВЛЕННЯ З ВИСОКИМ СПЕКТРАЛЬНИМ РОЗРІЗНЕННЯМ.....	38
3.1 Постановка задачі.....	38
3.2 Шумоподавлення при фіксованому спектрі шуму.....	39
3.3 Алгоритм рекурсивного згладжування за часом.....	41
3.4 Алгоритм багатомасштабних перетворень.....	43
3.5 Висновки по розділу 3.....	45
4 ДОСЛІДЖЕННЯ АЛГОРИТМІВ ШУМОПОГНИКАННЯ НА ОСНОВІ СПЕКТРАЛЬНОГО ВІДНІМАННЯ.....	46

4.1 Алгоритм спектрального віднімання.....	46
4.2 Дослідження алгоритму спектрального віднімання.....	49
4.3 Дослідження алгоритмів шумоподавлення у звукових редакторах.....	54
4.4 Висновки по розділу 4.....	57
Висновки.....	58
Перелік посилань.....	62
Додаток А. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ ШУМОПОДАВЛЕННЯ У РЕДАКТОРАХ ЗВУКУ.....	66
Додаток Б. Графічний матеріал.....	73
Додаток В. Відомість кваліфікаційної роботи.....	85

ПЕРЕЛІК УМОВИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ,
СКОРОЧЕНЬ І ТЕРМІНІВ

АІМ – амплітудно-імпульсна модуляція;
АЦП – аналого-цифровий перетворювач;
АЧХ – амплітудно-частотна характеристика;
ШПФ – швидке перетворення Фур'є;
ДІКМ – диференціально-імпульсна кодова модуляція;
ДПФ – дискретне перетворення Фур'є;
ДР – детектор рівня;
ІКМ – імпульсно-кодова модуляція;
ЗС – звуковий сигнал;
СГМ – спектральна густина потужності;
ФВЧ – фільтр верхніх частот;
ФНЧ – фільтр нижніх частот;
ЦАП – цифро-аналоговий перетворювач;
Attack – час спрацьовування шумоподавлювача;
CCIR – МККР, Міжнародний консультативний комітет з радіо;
DNL – динамічний обмежувач шуму;
DNR – динамічний шумоподавлювач;
FFT – швидке перетворення Фур'є;
Gate (гейт) – пороговий шумоподавлювач;
Release – час відновлення;
SNR – відношення сигнал-шум;
Soft knee – "м'який" поріг спрацьовування шумоподавлювача;
SPL (Sound Pressure Level) – рівень звукового тиску;
STFT – короткочасне перетворення Фур'є;
Threshold – поріг спрацьовування шумоподавлювача;
VCA – підсилювач, керований напругою.

ВСТУП

Сучасна звукотехніка є однією з областей масової технологічної діяльності, в якій засобами електроніки здійснюється обробка, накопичення та розповсюдження в електричній формі сигналів звукового діапазону частот. Вона спрямована на задоволення потреб людини у знаннях, культурі, освіті.

Завдяки повсюдному поширенню звукотехнічних пристроїв у поєднанні із засобами масової аудіовізуальної інформації та комунікації формується змістовна частина навколишнього штучного акустичного середовища, що оточує людину. Таке середовище надає, як правило, позитивний раціональний та емоційний вплив на людей.

Світовий випуск різноманітних виробів аудіотехніки сягає сотень мільйонів одиниць на рік. До того ж масове застосування, однакове використання та загальні принципи функціонування звукотехнічних пристроїв означають, що науково-технічні, навчально-методичні та довідкові відомості про звукотехніку з однаковим успіхом можуть бути як створені, так і використані у будь-якій країні світу.

Цифрова обробка звукових сигналів грає дедалі важливішу роль із поширенням мультимедійних можливостей персональних комп'ютерів. Зародження основних методів цифрової обробки відбувалося у середині 20 століття, коли з'являлися перші комп'ютери. У ті роки великі зусилля під час створення алгоритмів цифрової обробки витрачалися на їх ефективну реалізацію. Зараз, коли потужності навіть персональних комп'ютерів зросли настільки, що дозволяють у реальному часі проводити складну обробку сигналів, на передній план виходить питання якості алгоритмів, а не їх швидкодії.

Часто користувачі готові пожертвувати швидкістю обчислення задля досягнення найкращих результатів.

Алгоритми цифрової обробки переходять із розряду вузькоспеціальних у розряд повсюдно використовуваних. Раніше вони застосовувалися в таких

завданнях, як професійний звукозапис та обробка звуку, радіолокація. Тепер же алгоритми ЦОС все активніше використовуються в повсякденному людино-машинному інтерфейсі, який стає мультимедійним: це і чіпи цифрових фотокамер, що обробляють зображення, і мобільні телефони, що кодують і обробляють звук, і персональні комп'ютери, що відіграють роль домашнього центру розваг за рахунок широких можливостей обробки звуку, зображень та відео.

Зменшення широкосмугового шуму в аудіосигналі – одне з завдань, з якими звукорежисери стикалися з ранніх часів. Вже на початку 40-х років у радіоприймачах широко використовувалися найпростіші гейти для подавлення перешкод при відсутності сигналу.

У кваліфікаційній роботі досліджуються сучасні методи шумоподавлення широкосмугових шумів у звуковому сигналі.

Розглядаються адитивні стаціонарні шуми. Багато рівних і однорідних у часі шумів задовольняють цим умовам (наприклад, шум магнітної стрічки, гул наведень електромережі, шипіння мікрофонного підсилювача, шум вентиляції в студії звукозапису і т.д.). Такі шуми, як клацання, випадання, нелінійні спотворення, скрипи і шарудіння, не є стаціонарними і не розглядаються у цій роботі.

1 АНАЛІЗ ХАРАКТЕРИСТИК І СПРИЙНЯТТЯ, ПРИЧИН ЗАШУМЛЕННЯ І АЛГОРИТМІВ ОБРОБКИ ЗВУКОВИХ СИГНАЛІВ

1.1 Аудіосигнали та їх характеристики

Звуком називаються малі механічні коливання у твердих, рідких чи газоподібних середовищах, частоти яких перебувають у діапазоні 20–20000 Гц. Частоти нижче 20 Гц називаються інфразвуковими, а вище 20000 Гц ультразвуковими.

Тиск, щільність звукової енергії, силу звуку зручно визначати у рівнях, що виражаються в децибелах.

$$N_p = 20 \lg\left(\frac{p}{p_0}\right), \quad N_I = 20 \lg\left(\frac{I}{I_0}\right). \quad (1.1)$$

N_p – рівень звукового тиску (SPL – Sound Pressure Level);

N_I – рівень потужності звуку (PWL – Power Level).

За опорну величину 0 дБ SPL приймається звуковий тиск $p_0 = 2 \cdot 10^{-5}$ Па, тобто звук, який відповідає порогу слухового відчуття. Рівень потужності акустичного сигналу, що відповідає цьому порогу, вимірюють в дБ PWL по відношенню до опорної величини $I_0 = 10^{-12}$ Вт/м².

При використанні одиниці дБ рівень громових гуркотів оцінюється приблизно в 120 дБ, шум літака або музика на рок-фестивалі відповідає рівню 110 дБ, шум потягу – 100 дБ, звуки галасливої вулиці – 80 дБ. Розмова в кімнаті відповідає рівню звуку приблизно 50...60 дБ, а шепіт – 20...30 дБ.

Аудіосигнал – електричний сигнал, що повторює за формою відповідні йому акустичні коливання.

У зв'язку з більшими діапазонами зміни аудіосигналів в аудіотехніці оперують не безпосередніми значеннями, а рівнями [1]. За рівень параметра

приймають величину пропорційну логарифму відносного значення цього параметра. Рівні прийнято вимірювати децибелах ([дБ] чи [dB]). Значення рівня енергетичних параметрів:

$$N_{\text{Э}} = 10 \cdot \lg \left(\frac{K_{\text{Э}}}{K_{0\text{Э}}} \right) \text{ [дБ]}. \quad (1.2)$$

Рівень лінійних параметрів визначається виразом:

$$N_{\text{Л}} = 20 \cdot \lg \left(\frac{K_{\text{Л}}}{K_{0\text{Л}}} \right) \text{ [дБ]}. \quad (1.3)$$

Якщо K_0 нормовано, то рівень є абсолютним, інакше – відносним. В останньому випадку вказується, відносно якої величини розрахований рівень. Значення K_0 відповідає опорному (еталонному) значенню рівня. Відносний рівень пов'язаний з абсолютною рівністю:

$$N_{\text{ОТН}}(K) = N_{\text{АБС}}(K) - N_{\text{АБС}}(K_0), \quad (1.4)$$

де $N_{\text{АБС}}(K_0)$ – абсолютний рівень для заданої величини K_0 ;

$N_{\text{АБС}}(K)$ – абсолютний рівень для заданої величини K .

Електричні рівні. За умовне (нормоване) значення електричного рівня прийнято величину потужності $P_0 = 10^{-3}$ Вт, що виділяється в навантаженні $R_0 = 600$ Ом. При цьому величина напруги становить:

$$U_0 = \sqrt{P_0 \cdot R_0} = \sqrt{0,001 \cdot 600} = \sqrt{0,6} = 0,775 \text{ (В)} = 775 \text{ (мВ)}. \quad (1.5)$$

Абсолютний рівень електричної потужності:

$$N_P = 10 \cdot \lg \left(\frac{P}{P_0} \right) = 10 \cdot \lg \left(\frac{P_{[Bm]}}{10^{-3}_{[Bm]}} \right) = 10 \cdot \lg P_{[mBm]}. \quad (1.6)$$

Відповідно абсолютний рівень за напругою:

$$N_U = 20 \cdot \lg \left(\frac{U}{U_0} \right) = 20 \cdot \lg \left(\frac{U_{[B]}}{0,775_{[B]}} \right) = 20 \cdot \lg \left(\frac{U_{[mB]}}{775} \right). \quad (1.7)$$

Рівні аудіосигналу за потужністю і напругою збігаються, якщо вимірювання проводиться на навантаженні опором 600 Ом. Якщо опір навантаження R_H відрізняється від 600 Ом, то рівень потужності відрізняється від рівня по нарузі на величину, що визначається виразом [1]

$$N_P - N_U = 10 \cdot \lg \left(\frac{600}{R_H} \right) \approx 27,8 - 10 \cdot \lg R_H. \quad (1.8)$$

Динамічний діапазон аудіосигналу – це різниця між його квазімінімальним та квазімаксимальним рівнями:

$$DD = L_{\max} - L_{\min}. \quad (1.9)$$

Квазімаксимальний рівень аудіосигналу (L_{\max}) – рівень, для якого відносна тривалість існування рівнів не нижча за нього дорівнює 2% для музичних програм (1% для мовних).

Квазімінімальний рівень аудіосигналу (L_{\min}) – рівень, для якого відносна тривалість існування рівнів не нижча за нього становить 98% для музичних програм (99% для мовних).

Приклади динамічних діапазонів реальних аудіосигналів наведено у таблиці 1.1.

Таблиця 1.1 – Динамічний діапазон реальних аудіосигналів

Вид сигналу	Промова диктора	Телефонна розмова	Невеликі ансамблі	Симфонічний оркестр
Динамічний діапазон, дБ	25...35	35...45	45...55	65...110

1.2 Сприйняття звукових сигналів

Сприйняттям звуку називається обумовлений зовнішніми впливами на органи слуху пізнавальний процес, у якому предмети та явища відбиваються у свідомості як образи.

Частотне сприйняття.

Людина відчуває зміни звукового тиску, що відбуваються із частотою 20...20000 Гц. Звуки з частотою нижче 40 Гц рідко зустрічаються в музиці і не існують у розмовній мові [1]. Сприймаються такі низькі звуки здебільшого суб'єктивними гармоніками, що виникають через нелінійності слуху.

На дуже високих частотах чітке музичне сприйняття зникає і виникає якість невизначене звукове відчуття, що залежить від особливості слухача, його віку. Експерименти показали, що люди, які ледь сприймають сигнали вище 12 кГц, дуже легко розпізнають у музичній передачі недостатність верхніх частот.

У частотному діапазоні чутних звуків людина запам'ятовує лише кілька сотень градацій частоти, причому число цих градацій різко зменшується із зменшенням інтенсивності звуку й у середньому не більше 100...150. Проте людина може опосередковано розрізнити зміну частоти до 0,3% на середніх частотах, наприклад, за умови зіставлення двох тонів, безпосередньо наступних один за одним або по биттям частот двох тонів.

Амплітудне сприйняття.

Область чутних людиною звуків обмежується за інтенсивністю порогами: знизу – порогом чутності і згори – порогом больових відчуттів.

Поріг чутності - це мінімальний рівень сигналу, що визначає межу переходу з чутного стану в нечутне і назад [1]. На середніх частотах $p_0 = 2 \cdot 10^{-5}$ Па, $I_0 = 10^{-12}$ Вт/м². У бік нижчих та найвищих звукових частот чутливість слуху падає (рис. 1.1), величина порога зростає.

У міжнародному стандарті ISO/R-226 прийняті стандартні значення порогів чутності (табл. 1.2) для слухачів віком 18...30 років.

Таблиця 1.2 – Стандартні значення порогів чутності (ISO/R-226)

Частота, Гц	100	200	400	800	1000	2000	3150	5000	8000	12500
Рівень, дБ	25,1	13,8	7,2	4,4	4,2	1,0	-3,6	-1,1	15,3	11,6

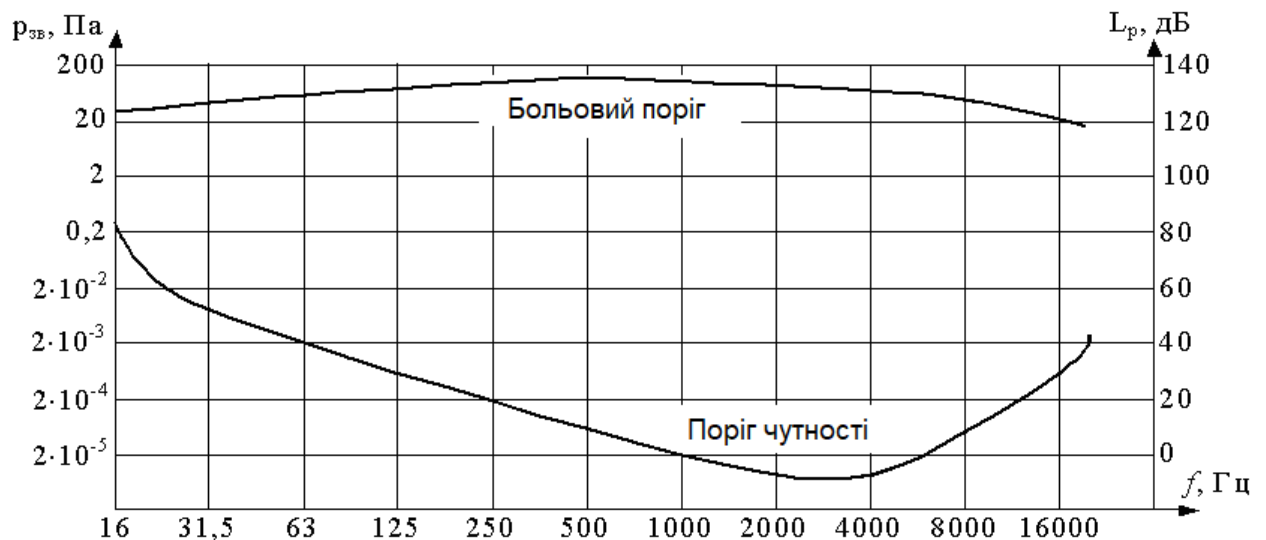


Рисунок 1.1 – Частотна залежність порогів сприйняття звуку

При плавному збільшенні звукового тиску слухове відчуття наростає стрибками. Елементарний стрибок зміни слухового відчуття ΔI називається порогом зміни гучності. Значення порога помітності залежить від поточної інтенсивності I_0 та частоти f (рис. 1.2).

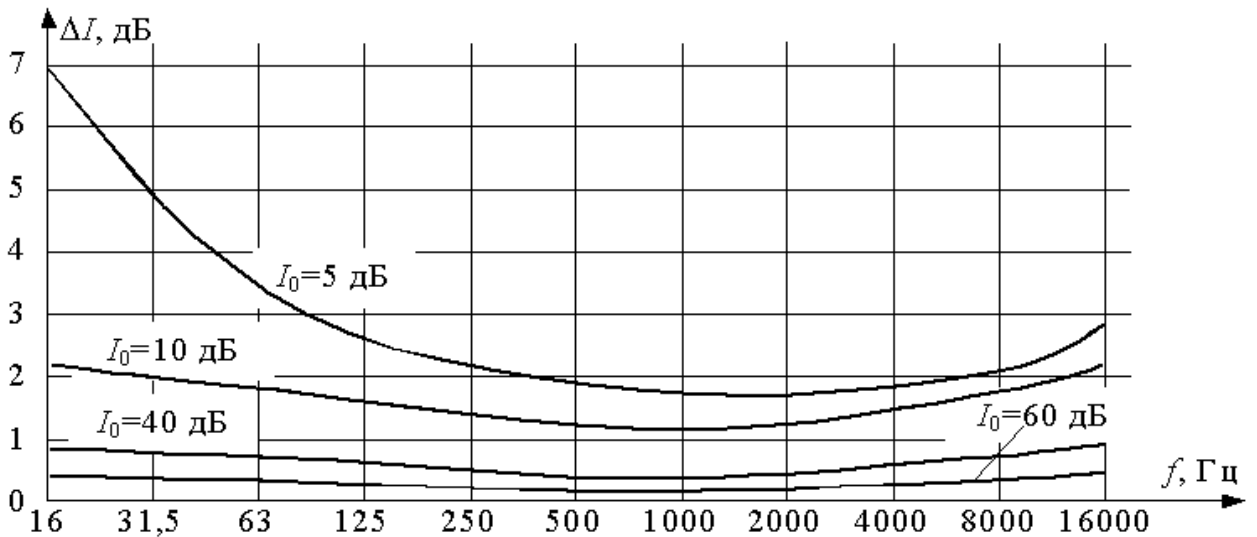


Рисунок 1.2 — Залежність порога помітності зміни інтенсивності від частоти тону

На високих і середніх рівнях відчуття рівень порогу помітності становить 0,2...0,6 дБ, але в низьких рівнях сягає кількох децибел. Число стрибків зміни слухового відчуття на середніх частотах дорівнює приблизно 250, на низьких і високих частотах воно зменшується, і в середньому по частотному діапазону становить близько 150.

Інтенсивність больового порога становить близько 10 Вт/м², і, таким чином, динамічний діапазон слуху від порога чутності до больового порога дорівнює 130 дБ сигналу з частотою 1000 Гц.

Тимчасове сприйняття звуку.

Слуховий апарат людини, як будь-яка коливальна система, інерційний. При зникненні звуку слухове відчуття зникає не відразу, а поступово, зменшуючись до нуля (рис. 1.3).

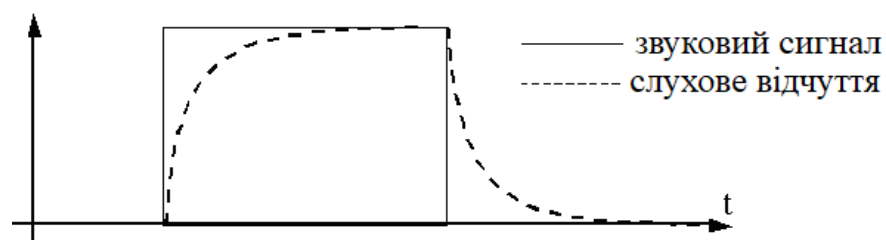


Рисунок 1.3 — Наростання та спад слухових відчуттів

Час, протягом якого відчуття рівня гучності зменшується на 8...10 фон, називається постійної часу слуху. Вона становить 50 мілісекунд.

Відчуття гучності звукового імпульсу залежить не тільки від його рівня, а й від тривалості дії на вухо. Короткочасний звук, що триває 10–12 мс, сприймається вухом тихіше, ніж звук такого ж рівня, що діє на слух протягом 150–400 мс.

Маскування звуку.

Суб'єктивна оцінка гучності залежить від умов, за яких звук сприймається слухачем. У реальних умовах акустичний сигнал не існує в абсолютній тиші. Одночасно з ним на слух впливають сторонні шуми, що ускладнюють звукове сприйняття і маскують основний сигнал. Тобто поріг чутності прийому слабкого сигналу за умов шуму і перешкод зростає. Ефект маскування чистого синусоїдального тону стороннім шумом оцінюється величиною, що вказує на скільки децибел підвищується поріг чутності сигналу, що маскується над порогом його сприйняття в тиші (1.12).

$$M = L_{ПТ} - L_{ПШ}, \quad (1.10)$$

де $L_{ПТ}$ – рівень порога чутності у тиші;

$L_{ПШ}$ – рівень порога чутності при шумі.

Тон будь-якої частоти маскується нижчими тонами значно ефективніше, ніж вищими (несиметричність кривої а) на рис. 1.4).

Якщо одночасно існують два складні звукові сигнали, що складаються з певних спектрів звукових частот, то виникає ефект взаємного маскування. Причому ефект маскування буде сильнішим, якщо основна енергія обох сигналів лежить в одній області діапазону звукових частот.

При маскуванні у часовій області розрізняють явища післямаскування (зміна порога чутності після сигналу високого рівня) та передмаскування (зміна порога чутності перед надходженням сигналу максимального рівня).

Більш слабкий сигнал стає нечутним за 5-20 мс до включення сигналу маскування і стає чутним через 50-200 мс після його включення.

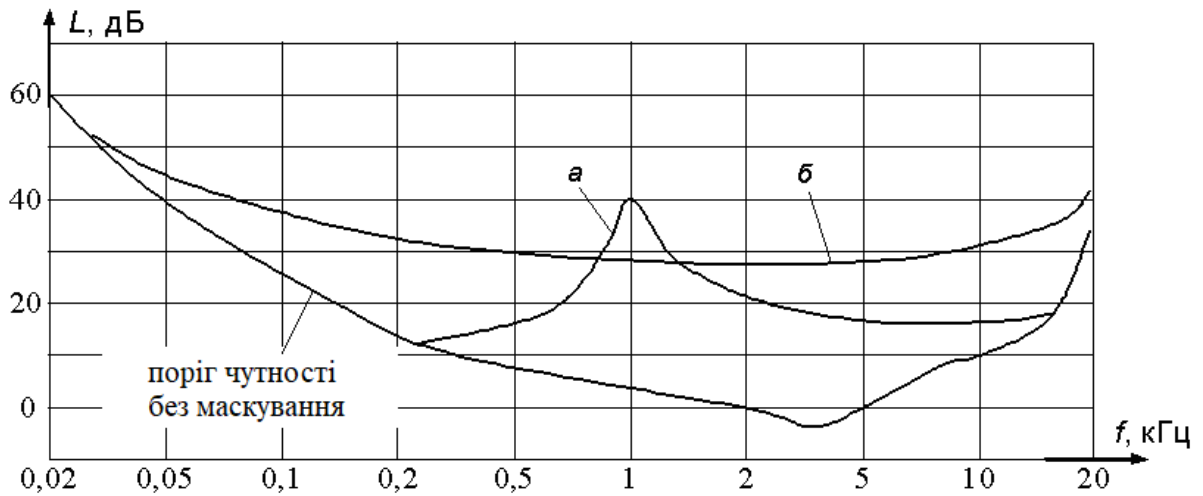


Рисунок 1.4 – Криві пороги чутності при маскуванні:

а – чистим тоном частотою 1 кГц; б – білим шумом

Дані ефекти використовуються як розробки алгоритмів стиснення аудіоданих, і при побудові алгоритмів шумоочищення звуку.

1.3 Цифрове представлення звукових сигналів

Серед алгоритмів цифрового представлення звукових сигналів (ЗС) найвідомішою є імпульсно-кодова модуляція (ІКМ). Процедура перетворення аналогового сигналу на цифровий складається з трьох операцій: дискретизації за часом, квантування отриманої сукупності відліків та заміни квантованих значень сигналу послідовністю чисел (кодування).

Дискретизація.

Значення частоти дискретизації обмежує роздільну здатність аналого-цифрового перетворення в часі і, отже, найвищу частоту на вході АЦП. Процедура дискретизації з однією і тією ж частотою f_d представлена на рис.1.5 одночасно для двох тональних сигналів низької ($F1$ – суцільна лінія) і високої ($F2$ – штрихова лінія) частот. В обох випадках після дискретизації маємо ідентичні часові послідовності значень відліків. Отже, сигнали цих

двох частот $F1$ та $F2$ не можна розрізнити і після зворотного перетворення правильно відновити (реконструювати).

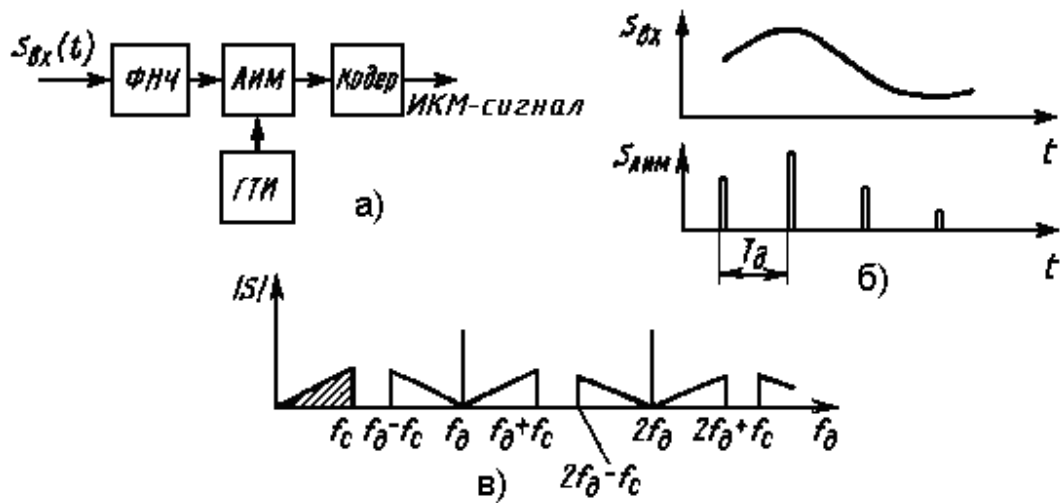


Рисунок 1.5 – Отримання сигналу з ІКМ

Відповідно до теореми відліків В.А. Котельникова неспотворена передача безперервного (аналогового) сигналу зі смугою частот $0 \dots F_{max}$ дискретною послідовністю його відліків можлива тільки в тому випадку, якщо частота f_d пов'язана з максимальною частотою F_{max} вихідного сигналу співвідношенням [2,3]

$$f_d \geq 2 F_{max} . \quad (1.11)$$

Отже, якщо потрібно передати синусоїдальне коливання з частотою 20 кГц, то необхідна частота дискретизації повинна бути більше 40 кГц, лише в цьому випадку можливе точне відновлення безперервного сигналу.

Усі сигнали, частота яких $F > f_d/2$, при відновленні інтерпретуються неправильно та трансформуються у сигнали нижчої частоти (рис. 1.6).

При передачі ЗС чи його обробці крок дискретизації T_d не повинен бути однаковим у різних трактах каналу чи різних стадіях обробки.

З урахуванням цих міркувань при АЦП сигналів ЗВ у цифрових трактах первинного та вторинного розподілу програм для ЗС з $F_{max} = 15$ кГц прийнято значення $f_d = 32$ кГц, що відповідає вчетверному значенню f_d для

сигналу в телефонному каналі. У трактах формування програм при $F_{max} = 20$ кГц прийнято значення $f\delta = 48$ кГц, лазерних програвачах $f\delta = 44,1$ кГц [2,3].

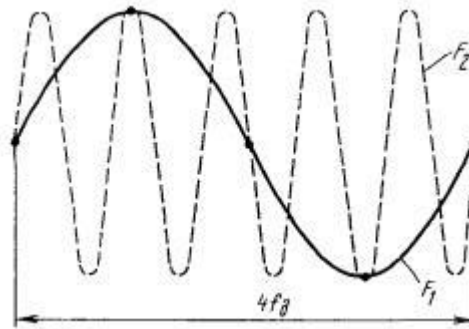


Рисунок 1.6 – Невірне відновлення аналогового сигналу

Боротьба з похибками цифрового перетворення через можливу нестабільність частоти дискретизації зводиться до обмеження смуги частот ЗС фільтрами зі смугою пропускання від 0 до F_{max} на вході АЦП та виході ЦАП та до виконання вимоги теореми В.А. Котельникова. Висуваються жорсткі вимоги до крутості спаду частотних показників ФНЧ поза смуги пропускання і стабільності частоти тактових генераторів. Виконання цих вимог сьогодні не викликає труднощів.

Квантування.

При квантуванні безперервному безлічі миттєвих значень відліків аналогового сигналу ставлять у відповідність кінцеве безліч значень – рівнів квантування. Інакше кажучи, кожне значення відліку замінюється найближчим до нього дозволеним значенням.

Відстань між сусідніми дозволеними рівнями квантування називають кроком квантування. Процедуру квантування можна розглядати як результат проходження вхідного сигналу через пристрій з амплітудною характеристикою ступінчастої форми (рис.1.7), яка називається шкалою квантування.

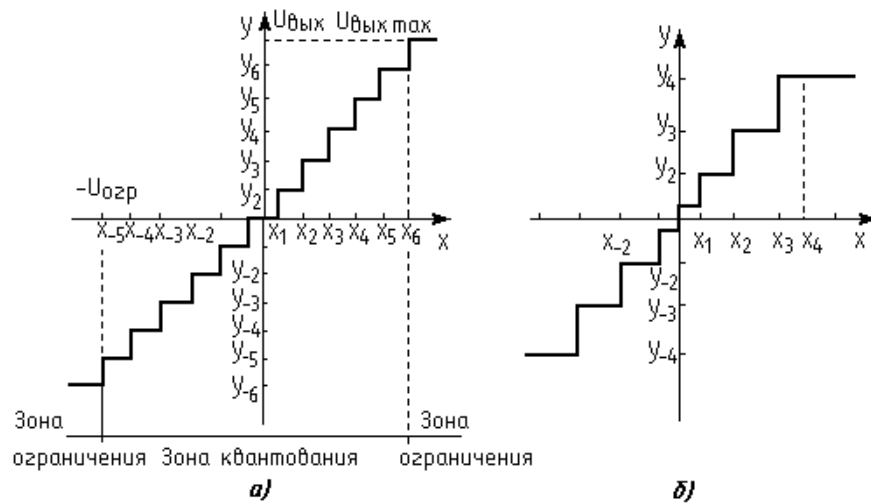


Рисунок 1.7 – Шкали квантування
(а – рівномірна, б – нерівномірна)

Одним з найважливіших показників цифрових систем передачі аналогових сигналів є величина відношення потужності сигналу P_c і потужності шуму квантування $P_{шкв}$ на виході ЦАП.

Для довільної шкали квантування з кроком Δ [2,3]:

$$P_{шкв} = \Delta^2 / 12. \quad (1.12)$$

З (1.12) випливає важливий висновок: при рівномірному квантуванні потужність шуму квантування визначається виключно кроком квантування і залежить від величини сигналу. Тому при зменшенні рівня сигналу відношення потужності сигналу до потужності шуму квантування знижується.

Спектр шуму квантування рівномірний у смузі частот $0 \dots f_d/2$. Шум квантування з'являється лише за наявності сигналу. При відсутності сигналу на вході АЦП, на виході ЦАП шум буде також присутній. Наявність теплового шуму вхідних аналогових блоків АЦП, нестабільність напруги живлення, перехідні перешкоди від сусідніх каналів, дрейф постійної складової в підсилювачах постійного струму та дія інших факторів призводять до того, що найнижчий перший рівень квантування досягається навіть за відсутності ЗС на вході АЦП (рис. 1.8).

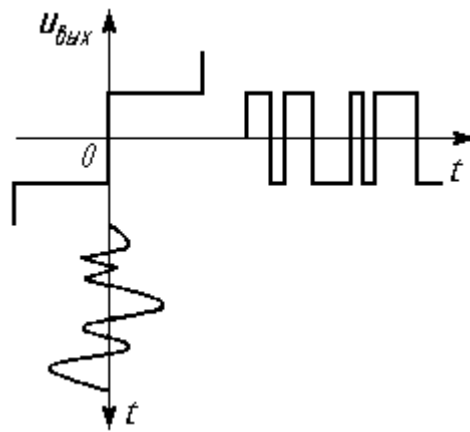


Рисунок 1.8 – Перетворення вхідних шумів на квантоване коливання в АЦП

При m -розрядному кодуванні та пік-факторі сигналу k [2,3]

$$P_c/P_{шкв} = 6m - 20 \lg k + 4,8. \quad (1.13)$$

З (1.13) видно, що з рівномірному квантуванні у випадку збільшення числа розрядів m на одиницю відношення С/Ш підвищується на 6 дБ. У той же час, перехід, наприклад, від 14-розрядного кодового слова до 15-розрядного означає збільшення швидкості передачі цифрового ІКМ-сигналу лише на 7%.

1.4 Алгоритми роботи лінійних цифрових фільтрів

Лінійний цифровий фільтр (ЛЦФ) – пристрій, у якому вихідні відліки сигналу представлені у вигляді лінійної комбінації попередніх відліків вхідного та вихідного сигналів та поточного відліку вхідного сигналу.

Основне різницеве рівняння для ЛЦФ має вигляд [6]:

$$y(nT) = \sum_{i=0}^M a_i \cdot x(nT - iT) + \sum_{j=1}^N b_j \cdot y(nT - jT), \quad n \geq 0. \quad (1.14)$$

де T – період дискретизації;

n – поточний номер відліку сигналу ($n = 0, 1, \dots$);

j – починається з 1, щоб рівняння було приведеним.

Перша сума визначає нерекурсивну частину фільтра, друга – рекурсивну частину фільтра.

Якщо a_i і $b_j \equiv \text{const}$, то ЛЦФ із постійними параметрами, інакше – ЛЦФ із змінними параметрами.

Більше з чисел M та N є порядком фільтра.

Якщо $b_j \equiv 0$, то отримуємо нерекурсивні фільтри або фільтри з кінцевою імпульсною характеристикою, так звані КІХ-фільтри.

Фільтри без зворотного зв'язку, тобто. $b_j \equiv 0$ називають однорідними, трансверсальними або фільтрами з багатовідвідною лінією затримки [2,6].

Якщо $b_j \neq 0$, отримуємо клас рекурсивних фільтрів.

Різницеve рівняння (1.14) точно визначає алгоритми функціонування не цифрових, а дискретних фільтрів, тобто. не враховуються ефекти квантування. У той самий час це різницеve рівняння визначає всі типи лінійних дискретних фільтрів (смугові, режекторні, ФНЧ, ФВЧ та інших.).

Форми реалізації (структурні схеми) ЛЦФ можуть бути різними [6]: пряма (рис. 1.9), канонічна, послідовна (рис. 1.10), паралельна (рис. 1.11) та інші.

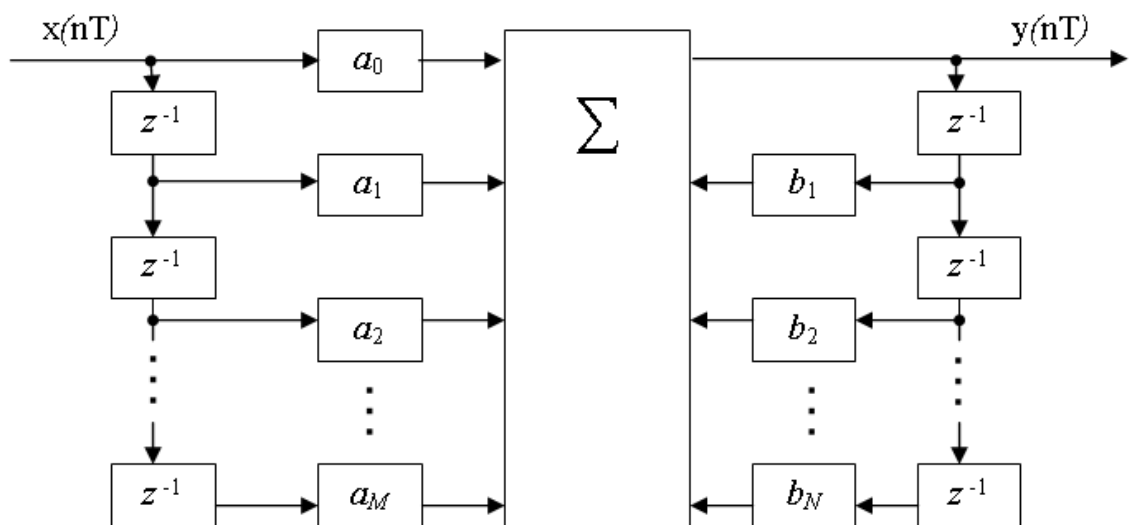


Рисунок 1.9 – Основна (пряма) форма реалізації ЛЦФ

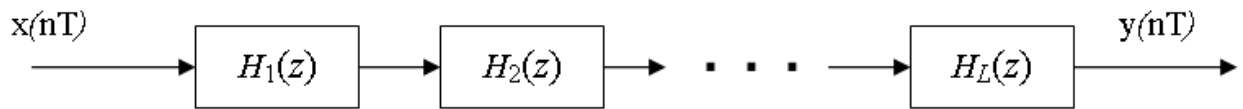


Рисунок 1.10 – Послідовна форма реалізації ЛЦФ

Для послідовного фільтра (рис. 1.10):

$$H_{\Sigma}(z) = \prod_{l=1}^L H_l(z). \quad (1.15)$$

Як правило, $H_l(z)$ це ЦФ 1-го або 2-го порядків.

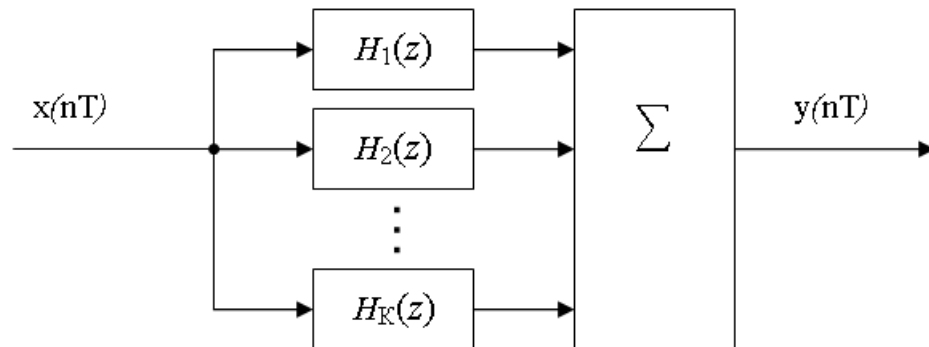


Рисунок 1.11 – Паралельна форма реалізації ЛЦФ

Для паралельної форми реалізації ЛЦФ рис. 1.11,

$$H(z) = \sum_{k=1}^K H_k(z). \quad (1.15)$$

Існують інші форми реалізації ЛЦФ (наприклад, сходова). Для дискретних фільтрів, тобто. коли відсутня операція квантування сигналів за рівнем, вихідний сигнал залежить від форми реалізації ЛЦФ.

До характеристик ЛЦФ із постійними параметрами відносяться [6]:

- системна (передавальна) функція у Z-формі;
- імпульсна та перехідна характеристики;

– комплексно-частотна, амплітудно-частотна та фазочастотна характеристики;

– груповий час запізнення);

– точнісні характеристики (ТХ).

Для цифрових фільтрів, коли вхідні та проміжні сигнали представлені квантованими за рівнем дискретними послідовностями, різні форми реалізації призводять до різних вихідних сигналів.

Таким чином, похибки фільтрації цифрових даних залежать як від розрядності уявлення вхідного та вихідного масивів чисел, так і від послідовності виконання операцій над ними.

1.5 Шумові спотворення та їх причини

За відсутності сигналу на вході підсилювача з його виході діє деяка невелика напруга. Ця напруга обумовлена переважно його власними перешкодами, серед яких розрізняють фон, наведення, мікрофонний ефект, теплові шуми резисторів і пасивних елементів, шуми підсилювальних елементів.

Фон зазвичай з'являється внаслідок недостатньої фільтрації пульсуючої напруги джерела живлення, що працює від мережі змінного струму. Гармонічні складові фону кратні частоті мережі живлення.

Наведення утворюються через паразитні електричні, магнітні, гальванічні або електромагнітні зв'язки ланцюгів підсилювача з джерелами перешкод.

Мікрофонний ефект є результатом перетворення механічних коливань елементів підсилювача в електричні, що проходять на вихід підсилювача. Спектр цих коливань займає діапазон 0,1...10000 Гц. Він помітно проявляється в інтегральних підсилювачах з великим коефіцієнтом посилення, виконаних на одній підкладці. Щоб усунути його,

використовують раціональну конструкцію елементів підсилювача, більш надійне їх кріплення, демпфування, застосовують пристрої, що амортизують.

Теплові шуми обумовлені тепловим безладним (випадковим) рухом в обсязі провідника (або напівпровідника) вільних носіїв зарядів. В результаті на кінцях провідника, який має деякий опір, діє випадкова, флуктуаційна ЕРС, звана ЕРС шуму $E_{ш}$. Її спектр є суцільним та практично рівномірним у діапазоні частот від нуля до сотень мегагерц. Шум із подібним спектром називають «білим».

Фон, наведення та мікрофонний ефект у підсилювачі можуть бути зменшені до будь-яких заданих значень. Теплові шуми і шуми підсилювальних елементів принципово непереборні і зазвичай вдається лише мінімізувати частку шумів, створюваних підсилювальними елементами.

Шумові властивості підсилювачів звукової частоти оцінюють ставленням сигнал/шум. Під цією величиною розуміють виражене децибелах відношення вихідної напруги сигналу при номінальній вихідній потужності підсилювача $P_{ном}$ до сумарної напруги шумів на виході. У якісних підсилювачах звукової частоти відношення "сигнал/шум" досягає 60...110 дБ.

1.6 Висновки по розділу 1

Цифрова обробка звукових сигналів відіграє важливішу роль із поширенням мультимедійних можливостей сучасних інформаційних технологій. Зародження основних методів цифрової обробки відбувалося у середині 20 століття, коли з'явилися перші комп'ютери. У ті роки великі зусилля під час створення алгоритмів цифрової обробки витрачалися на їх ефективну реалізацію. Зараз, коли потужності навіть персональних комп'ютерів зросли настільки, що дозволяють у реальному часі проводити складну обробку сигналів, на передній план виходить питання якості алгоритмів, а не їх швидкодії.

Як розробники, так і користувачі готові пожертвувати швидкістю обчислення для досягнення якісніших результатів.

Алгоритми цифрової обробки сигналів переходять із розряду вузькоспрямованих у розряд дуже поширених. Раніше вони використовувалися в таких задачах, як професійний звукозапис та обробка звуку, радіолокація. Сьогодні алгоритми цифрової обробки все активніше використовуються в повсякденному інтерфейсі, що відіграє роль домашнього центру розваг за рахунок широких можливостей обробки звуку, зображень та відео.

Зменшення широкосмугового шуму в звуковому сигналі – одне з завдань, з якими звукорежисери стикалися з ранніх часів. У кваліфікаційній роботі досліджуються сучасні методи подавлення широкосмугового шуму у аудіосигналі.

Розглядаються адитивні стаціонарні шуми. Багато рівномірних і однорідних у часі шумів задовольняють цим умовам (наприклад, шум магнітної стрічки, гул наведень електромережі, шипіння мікрофонного підсилювача, шум вентиляції в студії звукозапису і т.д.). Такі шуми, як клацання, випадіння, нелінійні спотворення, скрипи і шорохи, які не є стаціонарними, не розглядаються у даній кваліфікаційній роботі.

Дана кваліфікаційна робота виконана на кафедрі МІРЕС ХНУРЕ. На кафедрі проводяться дослідження в таких наукових областях, як виявлення та розпізнавання БПЛА за результатами акустичного спостереження [14-17], створення систем зондування атмосфери за допомогою акустичних хвиль [18-21]. Цілий ряд студентських доповідей [22-25] і атестаційних робіт магістрів минулих років [26-29] присвячені дослідженню систем озвучення приміщень. Отже, дослідження в даній роботі відповідають традиційному напрямку наукових робіт колективу кафедри МІРЕС.

2 ДОСЛІДЖЕННЯ БАЗОВИХ АЛГОРИТМІВ ШУМОПОДАВЛЕННЯ У ЗВУКОВИХ СИГНАЛАХ

2.1 Вимоги до шумів, що подавляються

У цій роботі розглядаються адитивні стаціонарні шуми. Стаціонарність означає, що такі властивості шуму, як потужність та спектр, не змінюються у часі. Адитивність означає, що шум накладається на чистий сигнал і не вносить до нього інших видів спотворень – модуляції, нелінійні спотворення. Багато рівних і однорідних у часі шумів задовольняють цим умовам. Наприклад, шум магнітної стрічки, шум наведень електромережі, шипіння мікрофонного підсилювача, шум вентиляції в студії звукозапису - всі ці шуми приблизно стаціонарні та адитивні.

З іншого боку, такі шуми, як клацання, випадання, нелінійні спотворення, скрипи та шарудіння не є стаціонарними і не розглядаються в магістерській атестаційній роботі. Для них існують свої методи реставрації.

Розглянуті шуми не обов'язково мають бути широкосмуговими. Наприклад, свист наведення від рядкового генератора ЕЛТ-монітора (тон із частотою близько 16 кГц) або гул фону від електромережі (тон 50 Гц з гармоніками) теж є стаціонарними та адитивними перешкодами і допускають подавлення розглянутими алгоритмами.

2.2 Пороговий шумоподавлювач (гейт)

Назва походить від англійського слова "Gate" - клапан, ворота. Робота гейта [7,8] заснована на принципі автоматичного зменшення посилення у тракті відтворення в паузах, коли шуми проявляються найбільше. Для визначення паузи використовується відмінність рівнів сигналу та шуму. Структурна схема гейту показано на рис.2.1.

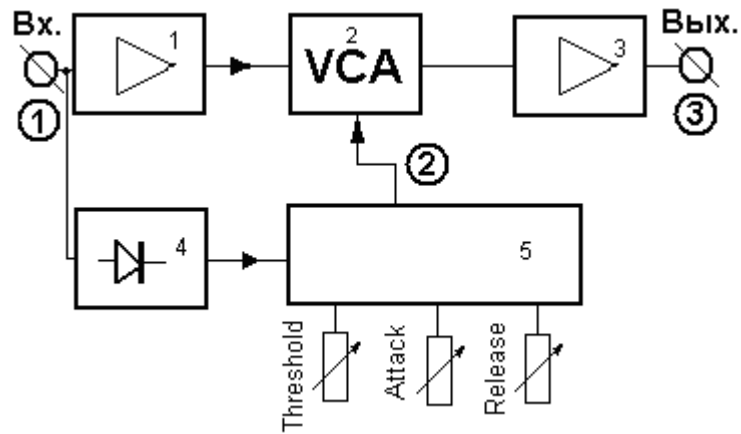


Рисунок 2.1 – Структурна схема порогового шумоподавлювача (1,3 – підсилювач, 2 – керований підсилювач, 4 – детектор рівня, 5 – пристрій управління)

Поріг спрацьовування (Threshold) встановлюють вручну таким чином, щоб зменшення шуму не супроводжувалося помітним зниженням рівня слабого сигналу. Деякі реалізації порогових шумоподавлювачів повністю відключають аудіосигнал від виходу звукового тракту, поки рівень сигналу знову не перевищить встановлений поріг спрацьовування (рис.2.2)

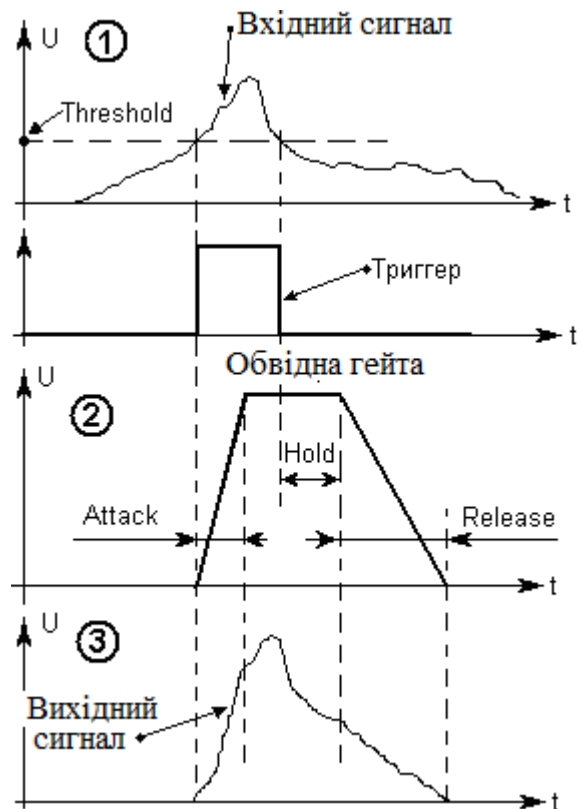


Рисунок 2.2 – Осцилограми у схемі порогового шумоподавлювача

При надто високому порозі тихі сигнали, які можна порівняти за рівнем із шумом, будуть пригнічуватися гейтом. При надто низькому порозі випадкові сплески шуму періодично відкриватимуть гейт і прориватимуться в результуючий сигнал.

Для зменшення ефектів різкого відкриття та закриття гейту сплесками шуму існують схеми уповільнення роботи гейту у часі. Такі параметри, як час атаки і час відновлення (attack time, release time), задають швидкість реакції гейту зміни рівня вхідного сигналу і допомагають згладити процеси відкриття і закриття.

У складніших пристроях гейти можуть набувати рис експандерів – схожих пристроїв динамічної обробки, що зменшують рівень слабких сигналів. Ступінь придушення шуму може залежати від рівня вхідного сигналу, а передавальна характеристика може мати "м'який поріг" (soft knee, рис.2.3).

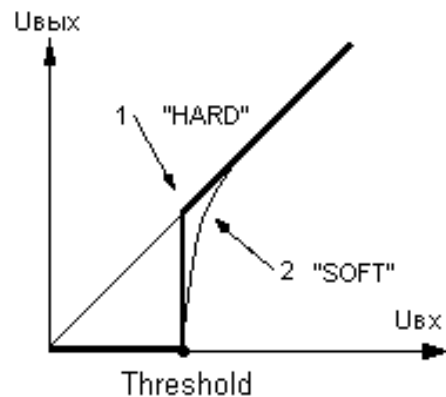


Рисунок 2.3 – Амплітудна характеристика гейту

Проведено дослідження роботи гейту у програмі Sound Forge. Як вихідна шумна фонограма використана запис мови на тлі гулу і широкосмугового шуму. На диску цьому запису відповідає файл S+N.wav. Сигналограма вихідного запису показано на рис. 2.4 а. На ній помітна шумова доріжка у паузах між словами.

Вихідний запис оброблений у редакторі Sound Forge за допомогою фільтра Noise gate з параметрами: поріг -20 дБ, час атаки 3 мс, час відновлення 100 мс. На рис. 2.4 б показана сигналограма після гейта.

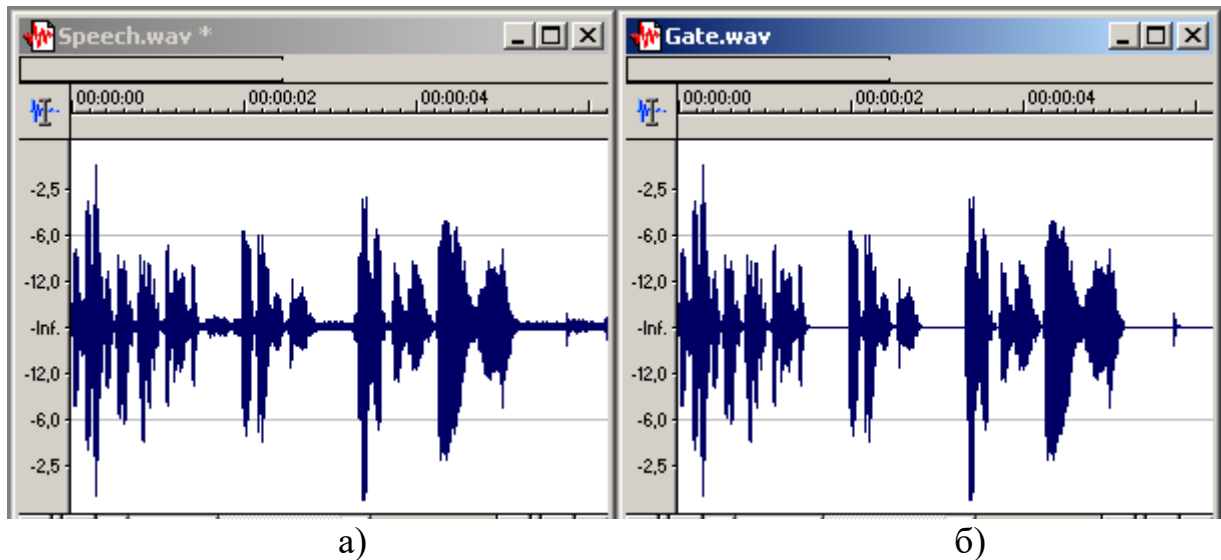


Рисунок 2.4 – Сигналограми вихідного зашумленого запису (а) та обробленої Noise gate (б)

З рис.2.4, а також із запису (файл NoiseGate.wav на диску додатку), можна зробити висновок, що Noise gate повністю пригнічує шум у паузах. До недоліків можна віднести ефект модуляції шуму сигналами низької амплітуди та подавлення корисних сигналів низької амплітуди. Фактично гейт придушив шум лише у паузах, а під час звучання програми його дії немає. Це істотний недолік, особливо при малому відношенні сигнал/шум.

Основною перевагою гейту є простота. Пороговий шумоподавлювач дозволяє зменшити сумарний рівень шуму в каналі відтворення на 1...2 дБ.

2.3 Динамічний шумоподавлювач

Принцип роботи системи динамічного придушення шуму [7,8] – Dynamic Noise Reducer (DNR), заснований на тому, що через особливості слухового сприйняття зменшення рівня звукового сигналу призводить до суб'єктивного звуження смуги чутних частот (рис. 1.1). Таким чином,

відповідне звуження смуги пропускання тракту зменшує вплив перешкод, частоти яких лежать поза спектром чутного сигналу.

Система DNR (рис. 2.5) автоматично звужує смугу пропускання звукового тракту (за рахунок обмеження ВЧ, де шуми особливо помітні) при малих рівнях вхідного сигналу та розширює – при більших.

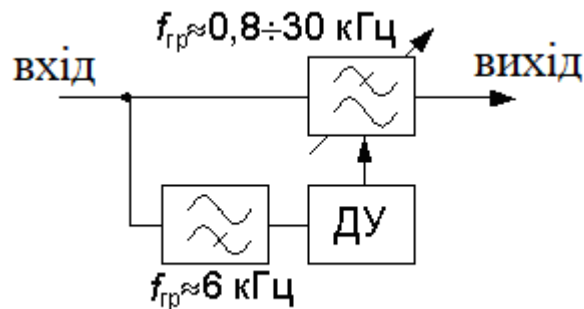


Рисунок 2.5 – Структурна схема динамічного шумоподавлювача (DNR)

З цією метою канал передачі аудіосигналу (як правило, на виході підсилювача відтворення) включений ФНЧ з керованою граничною частотою в межах від 800 Гц до 30 кГц (рис. 2.5). Керуючий сигнал виділяється з корисного сигналу за допомогою ФВЧ ($f_{гр} = 6$ кГц) і формується детектором рівня (ДК).

У паузі та за невеликих рівнях сигналу смуга пропускання звукового тракту встановлюється мінімальною (показано суцільною лінією на рис. 2.6).

При збільшенні гучності (рівня сигналу) смуга пропускання розширюється (показано стрілкою на рис. 2.6), але одночасно збільшується маскування шумів корисним сигналом і придушення шумів стає не обов'язковим.

У середньому система динамічного придушення шуму дає вигоду по відношенню «сигнал/шум» $10 \div 14$ дБ.

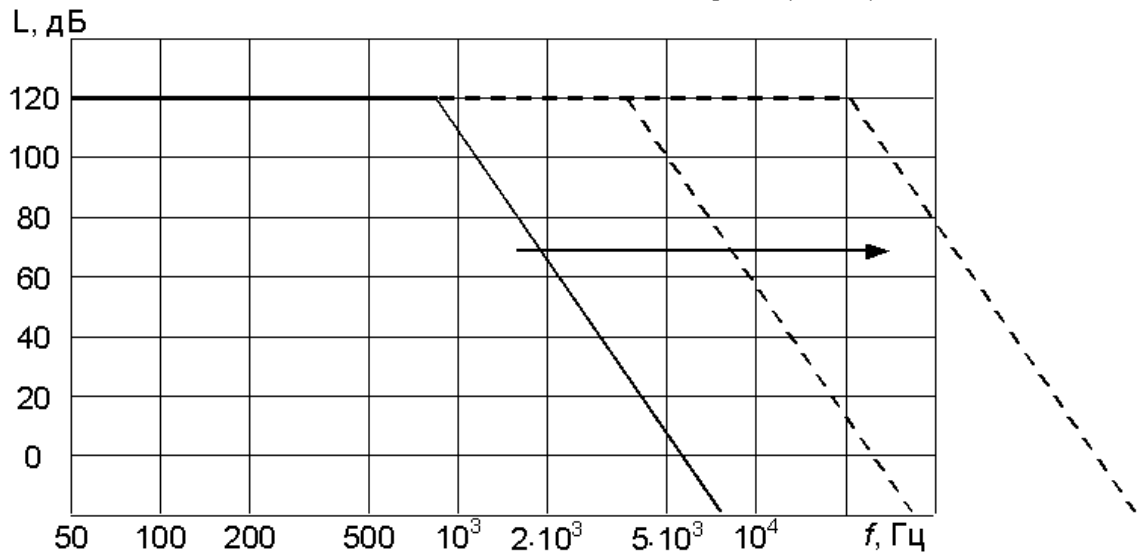


Рисунок 2.6 – АЧХ ФНЧ з керованою частотою зрізу

2.4 Динамічний обмежувач шуму

Система динамічного обмеження шуму (Dynamic Noise Limiter –DNL) також застосовується у випадках, коли найбільше заважає шум вищих частот, у паузах або тихому відтворенні запису [7,8].

Структурна схема динамічного обмежувача шуму зображено на рис. 2.7.

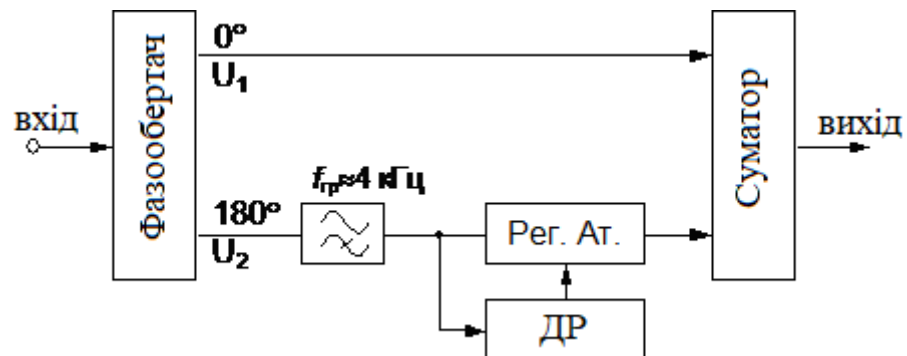


Рисунок 2.7 – Структурна схема динамічного обмежувача шуму (DNL)

Аудіосигнал надходить на фазообертач, на виходах якого сигнали зсунуті на 180 градусів. Один з них (U_1) надходить на суматор безпосередньо, а другий (U_2) – через ланцюг, до якого входять ФВЧ з фіксованою граничною частотою (зазвичай $f_{гр} \approx 4$ кГц) та система АРП із

затримкою, що складається з регульованого атенюатора та детектора рівня (ДР). Коефіцієнт передачі регульованого атенюатора обернено пропорційний керуючій напрузі, що формується детектором рівня. Пороговий рівень сигналу U_2 , при якому починає змінюватися коефіцієнт передачі атенюатора, обирається приблизно на 38 дБ менше від номінального.

Поки рівень вхідного аудіосигналу вище порогового значення коефіцієнт передачі атенюатора близький до нуля і корисний сигнал (U_1) надходить на вихід системи без змін. Коли рівень сигналу опускається нижче порогового значення, коефіцієнт передачі атенюатора підвищується і складові спектру з частотами вище 4 кГц віднімаються у суматорі і, таким чином, компенсуються.

Амплітудно-частотні характеристики динамічного обмежувача шуму за різних рівнів вхідного сигналу наведено на рис. 2.8.

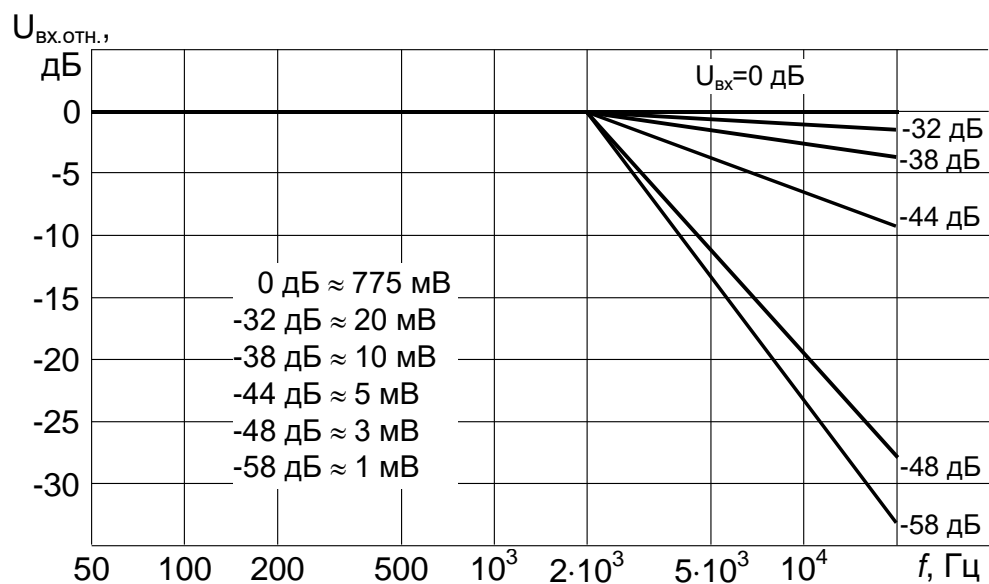


Рисунок 2.8 – АЧХ динамічного обмежувача шуму при різних рівнях вхідного сигналу

Система DNL дає вигравш щодо «сигнал/шум» на частоті 6 кГц приблизно 5 дБ, але в частоті 10 кГц – до 20 дБ. DNL використовується також у пристроях, які призначені для реставрації старих фонограм.

2.5 Багатосмуговий гейт

Багатосмуговий гейт (Multiband gate) можна вважати комбінацією порогового шумоподавлювача Noise gate та динамічних систем DNR та DNL [7,8]. У багатосмуговому гейті ступінь подавлення у кожному з кількох частотних смуг залежить від потужності сигналу у цій смузі (а в деяких випадках – і від розподілу потужностей сигналу у сусідніх полосах). Завдяки наявності декількох частотних смуг можна вибірково подавлювати тільки ті смуги, які в даний момент містять шум. Структурна схема трисмугового гейту показана на рис. 2.9.

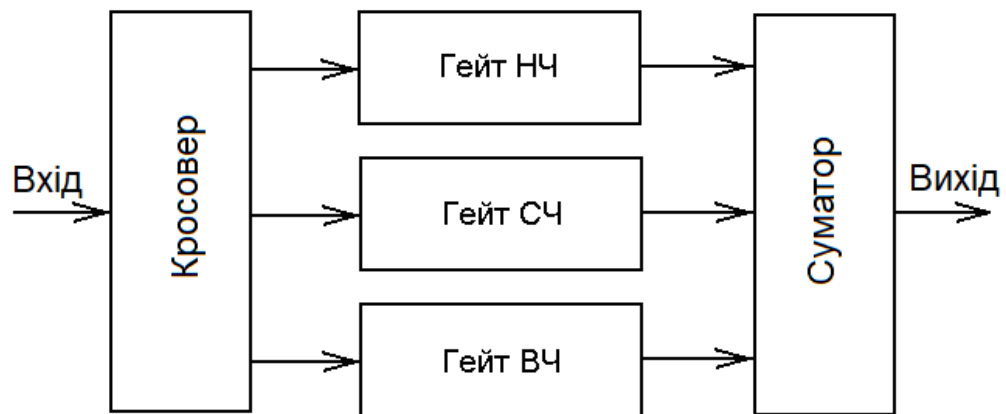


Рисунок 2.9 – Структурна схема трисмугового гейту

Проведено дослідження роботи трисмугового гейту Multiband NoiseGate NX3, випущеного фірмою 123creative.com у форматі VST-плагіну. Плагін було підключено до звукового редактора Cubase.

Частотні смуги мали такі межі: 0...700 Гц, 0,7...3,2 кГц, 3,2 кГц і вище. У всіх смугах поріг -20 дБ, час атаки 3 мс, час відновлення 100 мс. На рис. 2.10 а показана сигналограма після Noise gate, а на рис.2.10 б - після Multiband NoiseGate NX3.

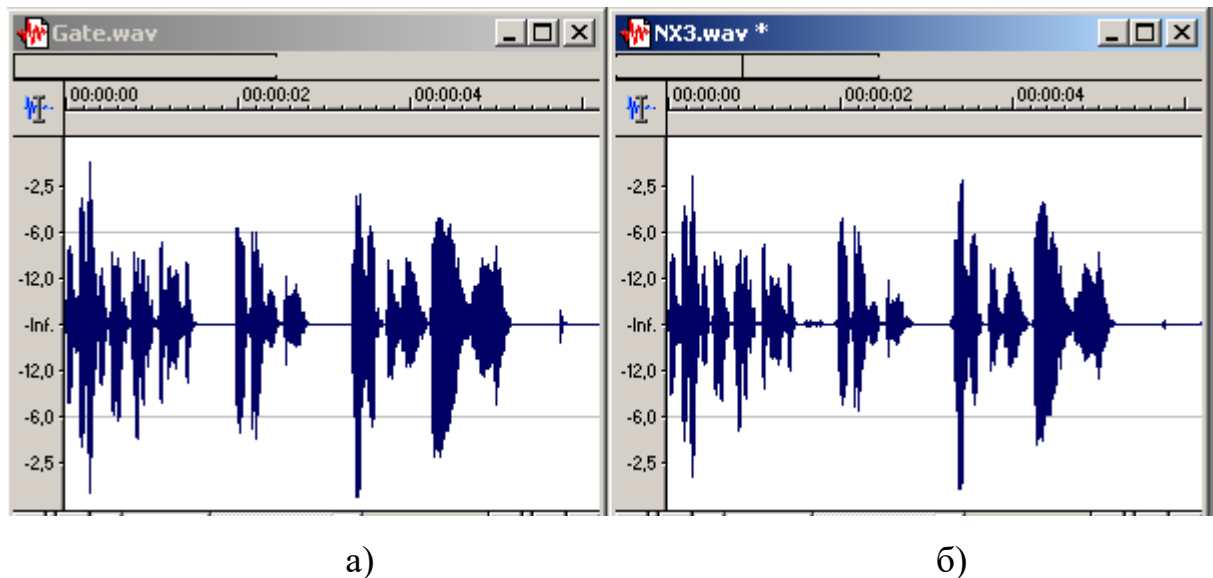


Рисунок 2.10 – Сигналограми записів оброблених Noise gate (а)
та Multiband NoiseGate NX3 (б)

З рис.2.10, а також із запису (файл MultibandGate.wav на диску програми), можна зробити висновки:

- у паузах широкосмугові гейти відкриваються від слабких звуків;
- тембр звучання приглушений у порівнянні з вихідним записом;
- помітність роботи трисмугового гейту на слух та погіршення якості менше, ніж у односмугового гейту.

2.6 Висновки по розділу 2

Таким чином, основні недоліки роботи трисмугового гейту пов'язані з малим числом смуг. Можливості аналогової електроніки обмежує кількість смуг значенням 4...6. Зняти це обмеження вдається з переходом на цифрову обробку звуку.

Алгоритми шумоподавлення з великою спектральною роздільною здатністю досліджуються в розділі 3.

3 ДОСЛІДЖЕННЯ АЛГОРИТМІВ ШУМОПОДАВЛЕННЯ З ВИСОКИМ СПЕКТРАЛЬНИМ РОЗРІЗНЕННЯМ

3.1 Постановка задачі

Розглянемо таку задачу. Є запис звукових сигналів (мова, музика тощо) на певному носії. Внаслідок недосконалості пристроїв запису-відтворення та/або носія інформації, сигнал, що відтворюється, не є точною копією записаного. Якщо позначити вихідний сигнал через $x(t)$, а адитивний шум у пристрої запису-відтворення через $n(t)$, можна записати:

$$s(t) = f(x(t)) + n(t). \quad (3.1)$$

де $f(x)$ – передатна функція системи запису-відтворення, як правило, дуже близька до лінійної, тобто $f(x(t)) \approx k \cdot x(t)$, де $k = const$.

Таким чином, одним із основних факторів, що викликають недосконалість запису, є адитивний шум. Його джерелом є хаотичний рух носіїв заряду в елементах аналогових кіл систем запису/відтворення. Існує ряд алгоритмів, спрямованих на підвищення відношення сигнал/шум [7,8]. Серед них слід зазначити застосування смугових фільтрів, у тому числі адаптивних та нелінійних фільтрів. Вони розглянуті у розділі 2.

Ринок систем цифрового шумоподавлення насичений і не перестає розвиватися. Проте лише деякі компанії мають справді якісні продукти і розробляють принципово нові ефективні алгоритми. Серед успішних розробок необхідно відзначити TC Electronic BackDrop, Waves Z-Noise, Algorithmix NoiseFree, iZotope RX, SoundSoap 2, розробки CEDAR. Також є безліч алгоритмів для зображень, що дозволяють досягти високої якості усунення шуму при адаптації їх до аудіосигналів (наприклад, Non-Local Means). Існують і інші способи усунення шуму, наприклад, з допомогою компандера з високим спектральним дозволом. Такі системи ґрунтуються не

тільки на алгоритмі спектрального віднімання зліпка шуму, але також на ефекті маскування і привносять у результуючий сигнал небажані модуляційні спотворення [8].

Систем з адаптивним вибором діапазону шуму досить мало, і вони спираються на прості принципи відокремлення перешкод від сигналу. Зазвичай класифікація проводиться по енергії: якщо вона опускається нижче деякого порогового значення, то фрагмент позначається як зашумлений. Подібним чином влаштована одна з американських запатентованих систем: якщо енергія в середині частотного діапазону перевищує деяку константу, то низькочастотний фільтр відключається [9]. Це дозволяє позбутися, наприклад, шуму електромережі в паузах між мовними сигналами. Розглянемо підходи, які найбільш наближені до тематики, що досліджується. Даний розділ для зручності розбито на дві частини:

- шумоподавлення при відомому спектрі шуму;
- шумоподавлення з адаптацією до шуму, що змінюється.

3.2 Шумоподавлення при фіксованому спектрі шуму

В основі цього пункту є припущення, що в середньому спектр шуму не змінюється в часі. Його слід задати один раз (наприклад, за допомогою усереднення вручну зазначеного фрагмента, що містить лише перешкоди), після чого з ним проводиться спектральне віднімання для всього аудіозапису. Алгоритм спектрального віднімання не завжди виявляється досить ефективним і для його удосконалення можливе використання багатомасштабних перетворень та різних алгоритмів згладжування спектра, що розглядаються нижче.

Алгоритм спектрального віднімання широко використовується і описаний у джерелах [1, 4, 5]. Адитивні стаціонарні шуми породжуються акустичними джерелами, звукозаписною апаратурою тощо.

Стационарність означає, що властивості шуму (потужність, спектральний склад) не змінюються у часі. Адитивність означає, що шум додається до "чистого" сигналу $y[t]$ і не залежить від нього. Запишемо вираз (3.1) у вигляді

$$x[t] = y[t] + \text{noise}[t], \quad (3.2)$$

де t – час.

Для подавлення адитивних стаціонарних шумів використовується алгоритм спектрального віднімання. Він складається з наступних стадій:

1. Розкладання сигналу за допомогою короткочасного (віконного) перетворення Фур'є (STFT), що компактно локалізує енергію сигналу [2].
2. Створення зліпка шуму, що віднімається (noise footprint). Зліпок шуму отримуємо усередненням за часом амплітуд спектру, взятих із заздалегідь підготовленої ділянки шуму, що не містить корисного сигналу.

$$\text{footprint}[f] = (\sum \text{noise}[f, t]) / K, \quad (3.3)$$

де $\text{noise}[f, t]$ – спектр шуму,

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

t – порядковий номер поточного STFT вікна;

K – кількість вікон на ділянці із шумом.

3. "Віднімання" (в узагальненому сенсі) амплітудного спектра шуму з амплітудного спектра сигналу.

4. Зворотне перетворення STFT – синтез результуючого сигналу. Віднімання амплітудних спектрів здійснюється за формулою:

$$Y[f, t] = \max\{ X[f, t] - k \cdot W[f, t], 0 \}, \quad (3.4)$$

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

$Y[f, t]$ – амплітудний спектр результуючого очищеного сигналу;

k – коефіцієнт подавлення.

Фазовий спектр очищеного сигналу береться рівним фазовому спектру зашумленого сигналу. Результат роботи алгоритму МСВ показано на рис. 3.1.

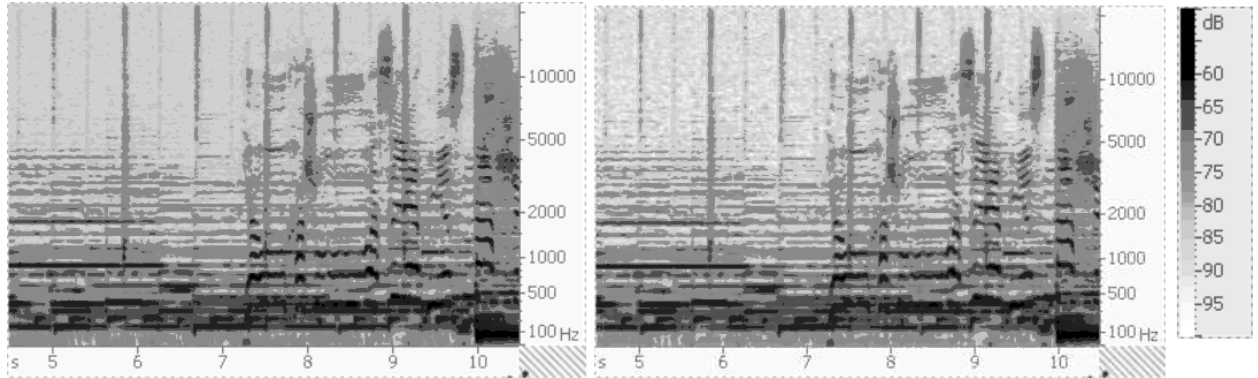


Рисунок 3.1 – Спектрограми зашумленого сигналу (ліворуч) та очищеного сигналу (праворуч)

3.3 Алгоритм рекурсивного згладжування за часом

Одна з проблем алгоритму спектрального віднімання – "музичний шум". Він утворюється внаслідок того, що коефіцієнти STFT шумових сигналів статистично випадкові. Це призводить до їхнього нерівномірного подавлення. В результаті, очищений сигнал містить короточасні та обмежені по частоті сплески енергії, які на слух сприймаються як "дзвіночки" або "вода, що ллється".

Алгоритми боротьби з «музичним шумом»:

- завищення порога (недолік – втрачаємо більше сигналу);
- обмеження $G(f, t)$ знизу (щоб музичний шум маскувався природним шумом);
- збільшення часу спрацьовування гейтів (при надто сильному збільшенні виходить шумове ехо та розмиваються атаки у сигналі);
- згладжування $G(f, t)$ за часом та/або за частотою.

На рис.3.2 показані отримані спектрограми сигналів для зазначених алгоритмів.

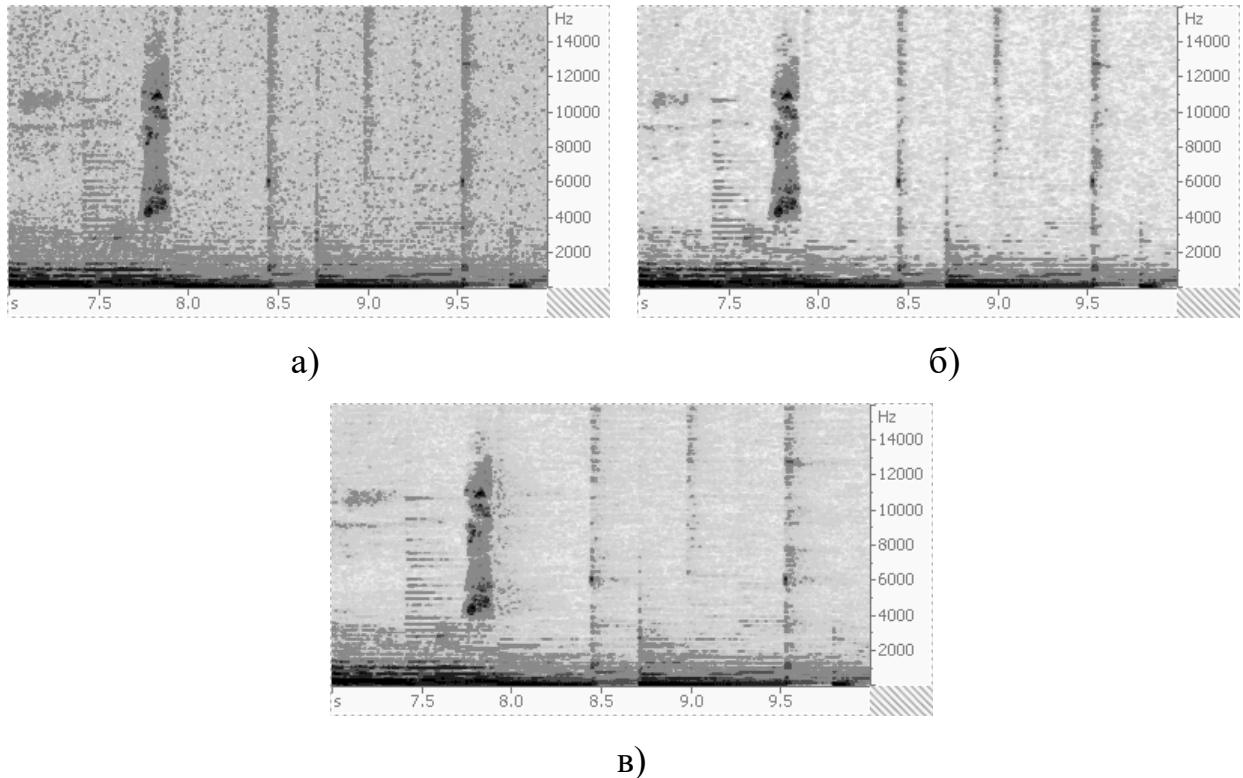


Рисунок 3.2 – Спектрограми сигналів: а) зашумлений сигнал, б) просте спектральне віднімання, в) згладжування за часом

Для подавлення артефакту як «музичного шуму» найчастіше застосовується алгоритм рекурсивного згладжування першого порядку за часом [1]:

$$S[f,t] = S[f,t-1] + \alpha \cdot (X[f,t] - S[f,t-1]), \quad (3.5)$$

де $S[f,t]$ – згладжений спектр сигналу;

$X[f,t]$ – f -та амплітуда спектра сигналу в момент часу t ;

α – коефіцієнт усереднення;

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

t – час.

Далі з усередненого спектра $S[f,t]$ віднімається зліпок шуму і проводиться зворотне перетворення Фур'є (iSTFT) з урахуванням перекриття вікон.

3.4 Алгоритм багатомасштабних перетворень

При використанні банків з постійним розрізненням неможливо отримати водночас якісне розрізнення як за частотою, так і за часом. Фільтри з фіксованою роздільною здатністю призводять до небажаного розмиття сигналу в частотній та часовій областях (рис. 3.3).

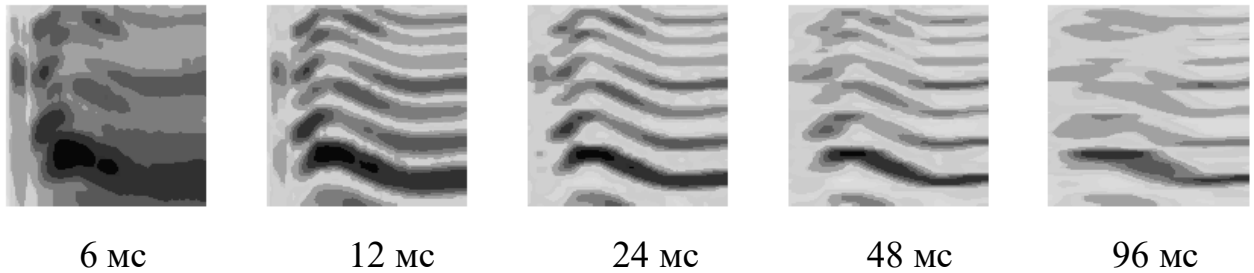


Рисунок 3.3 – Розмиття спектру та вибір кращого розрізнення

Банки фільтрів з адаптивною частотно-часовою роздільною здатністю, зокрема багатомасштабні перетворення (multi-resolution FFT = MR FFT), допомагають вирішити цю проблему [3, 12].

В основі MR FFT лежать:

- обчислення STFT, що взяті від одного сигналу, з різними розрізненнями (розмірами вікна);
- оцінка ступеня розрідженості відповідних областей спектра;
- Мікшування спектрограм з урахуванням їх розрідженості.

Спочатку виконується незалежна один від одного робота ефектів (фільтрів) з фіксованими розрізненнями (розмірами FFT):

STFT => Застосування ефекту => inverse STFT.

Результуючий сигнал кожного ефекту передається в мікшер. Запускається аналізатор, який розраховує коефіцієнти мікшування, спираючись на ступеня розрідженості областей спектрограм (рис. 3.4).

Ступінь розрідженості (розмиття) для кожної області Ω обчислюється як відношення середнього арифметичного модулів коефіцієнтів $a_i \in \Omega$ до їх середнього квадратичного.

$$S_{\Omega} = \frac{\sum_{i=1}^N a_i}{\sqrt{N} \sqrt{\sum_{i=1}^N a_i^2}}$$

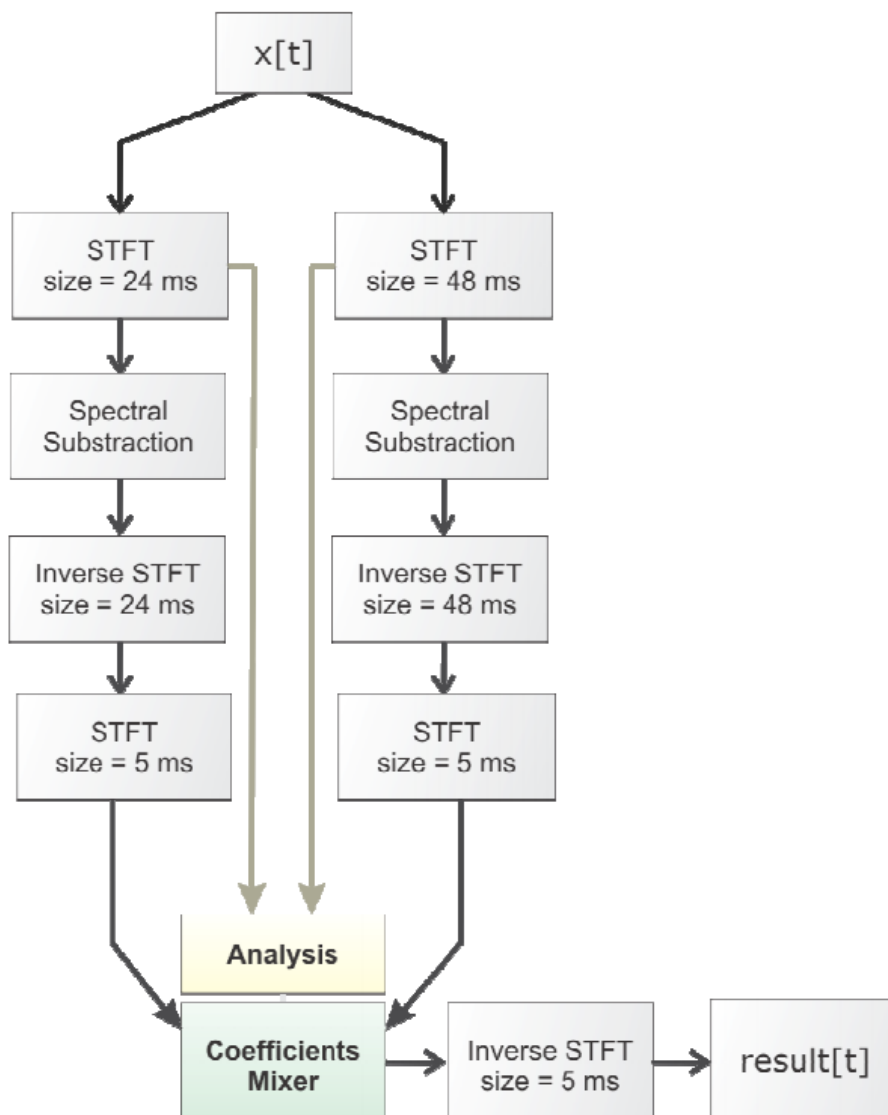


Рисунок 3.4 – Схема алгоритму багатомасштабних перетворень

Оскільки середнє арифметичне завжди менше або дорівнює середньому квадратичному, то $S_{\Omega} \leq 1$. У даній роботі S_{Ω} завжди обчислюється для спектрограм, амплітудні коефіцієнти яких більші або рівні

0. Отже, $0 \leq S_\Omega \leq 1$. Чим менше $|a_i - a_j|$ для будь-яких $i, j \in \Omega$ (тобто всі значення досить близькі один до одного), тим сильніше S_Ω наближається до 1. Це відповідає розмитому спектру, як, наприклад, у білого шуму.

За наявності малої кількості великих за величиною коефіцієнтів a_i міра S_Ω стає малою, що показує компактну локалізацію енергії спектра в малій кількості коефіцієнтів, як, наприклад, для тональних сигналів або коротких сплесків енергії.

Поведінка міри розмиття від двох змінних показана на рис. 3.5. При великій кількості значень в області Ω мінімум S_Ω прагне до 0.

Для кожної із спектрограм обчислюється свій коефіцієнт міксування:

$$\alpha = 1/S_\Omega \text{ и } \sum \alpha_i = 1, \quad (3.6)$$

де N – кількість аналізованих спектрограм (тобто кількість ефектів).

3.5 Висновки по розділу 3

Таким чином, для довгих тональних сигналів буде підбиратися FFT з найбільшою роздільною здатністю за частотою, а для різких сплесків енергії (кляцань, подій з швидкою атакою) – з найменшим за частотою, але найкращим за часом. Використання цього підходу дуже ефективно, і його робота навіть у сукупності із простим алгоритмом спектрального віднімання виявляється високоякісною.

4 ДОСЛІДЖЕННЯ АЛГОРИТМІВ ШУМОПОГНИКАННЯ НА ОСНОВІ СПЕКТРАЛЬНОГО ВІДНІМАННЯ

4.1 Алгоритм спектрального віднімання

Алгоритм спектрального віднімання широко використовується як сам по собі, так і у складі складніших алгоритмів, у багатьох програмах – редакторах звуку [11-14]. Він складається з наступних етапів:

- розкладання сигналу за допомогою короточасного перетворення Фур'є (STFT) або іншого перетворення, що компактно локалізує енергію сигналу;
- оцінка діапазону шуму;
- "віднімання" амплітудного спектра шуму з амплітудного спектра сигналу;
- зворотне перетворення STFT – синтез результуючого сигналу.

При короточасному віконному перетворенні сигнал заданої тривалості розбивається на ряд інтервалів за допомогою ковзного вікна того чи іншого типу. Це дозволяє отримувати, обробляти і будувати у вигляді спектрограм динамічні спектри та аналізувати їхню поведінку в часі. Спектрограма будується у трьох координатах – частота, час та амплітуда. При цьому амплітуда задається кольором або кольором кожного прямокутника спектрограми.

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

$$w(t) = 1 - \cos \frac{2\pi(t + 0,5)}{N}, \quad 0 \leq t \leq N \quad (4.1)$$

тривалістю близько 50 мс та ступенем перекриття 75%. Амплітуду вагового вікна масштабують так, щоб при вибраному ступені перекриття вікон банк фільтрів не змінював загальну амплітуду сигналу без обробки.

Оцінка спектру шуму може здійснюватися як автоматично шляхом пошуку ділянок мінімальної енергії в кожній частотній смузі, так і вручну шляхом аналізу спектра на часовому сегменті, який користувач ідентифікував як шум.

Зліпок шуму отримуємо усередненням за часом амплітуд спектру, взятих з ділянки шуму, що не містить корисного сигналу:

$$foot(f) = \frac{\sum_{t=1}^K n(f,t)}{K}, \quad (4.2)$$

де $n(f,t)$ – спектр шуму;

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

t – номер поточного STFT вікна;

K – кількість вікон на ділянці з шумом.

Віднімання амплітудних спектрів може здійснюватися за формулою

$$Y(f,t) = \max\{X(f,t) - k \cdot W(f,t), 0\}, \quad (4.3)$$

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

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

k – коефіцієнт подавлення.

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

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

Вираз (4.3) еквівалентно поданій функції подавлення:

$$G(f, t) = \frac{Y(f, t)}{X(f, t)} = \max \left\{ 1 - k \frac{W(f, t)}{X(f, t)}, 0 \right\}. \quad (4.4)$$

Алгоритм віднімання амплітудних спектрів наведений на рис.4.1.

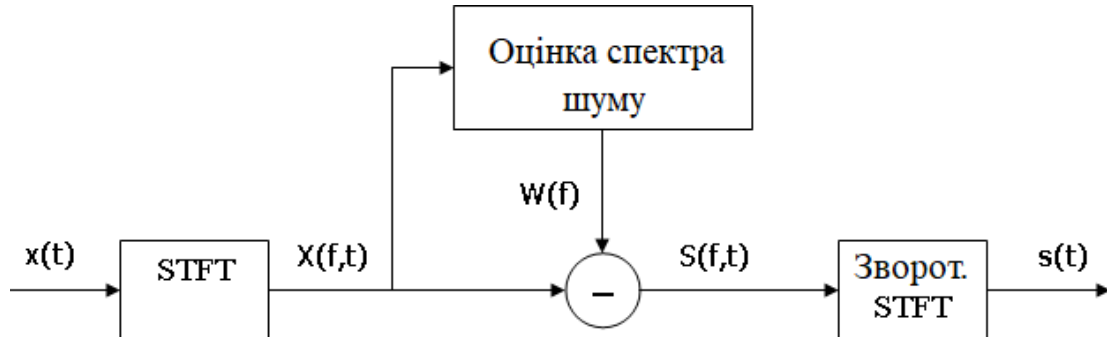


Рисунок 4.1 – Алгоритм віднімання амплітудних спектрів

Розглянутий алгоритм має близьку спорідненість із багатосмуговим гейтом. Спектрограму сигналу можна трактувати як набір горизонтальних рядків, кожен із яких є окремим субсмуговим сигналом. І над кожною такою частотною смугою проводиться операція віднімання з її амплітуди певної константи – рівня шуму у цій смузі. Коли рівень сигналу смуги великий, віднімання невеликої константи мало змінює сигнал у смузі. Коли рівень сигналу близький до рівня шуму, зменшення рівня сигналу на величину шуму фактично еквівалентно застосуванню гейта з «м'яким порогом». Тому весь процес спектрального віднімання можна як багатосмуговий гейт з великою кількістю смуг і «м'яким порогом».

Існує безліч різних реалізацій алгоритму спектрального віднімання. Вони можуть відрізнятися типами передавальної характеристики у індивідуальних гейтів (правилом «віднімання» спектра шуму з спектра сигналу), числом частотних смуг, способами боротьби з різними артефактами. Оскільки системи на основі спектрального віднімання є найбільш сучасними та поширеними, дослідимо ці аспекти докладніше.

Досліджуємо роботу алгоритму віднімання амплітудних спектрів у програмах обробки звуку.

4.2 Дослідження алгоритму спектрального віднімання

Чистий алгоритм спектрального віднімання реалізований у звуковому редакторі Audacity. Це програма, що вільно розповсюджується, з відкритою платформою та описаними алгоритмами обробки [14]. Тому редактор Audacity обрано для проведення досліджень.

Тестова запис є мова на тлі гулу (основний тон і гармоніки) і шипіння (широкополосний шум). Спектрограма вхідного зашумленого запису отримано у програмі Right Mark Audio Analyzer і показано на рис. 4.2.

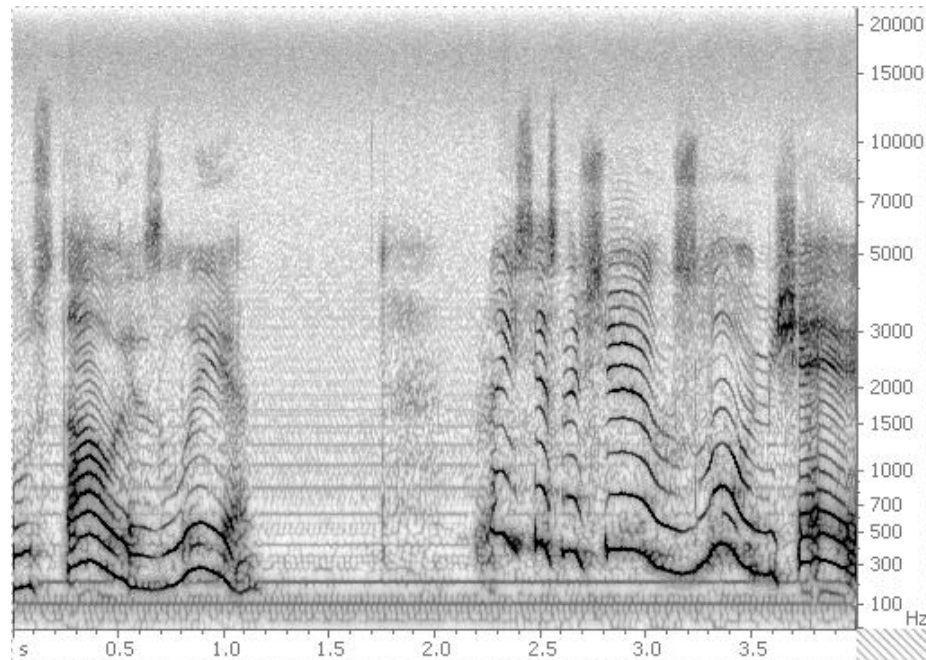


Рисунок 4.2 – Спектрограма вхідного зашумленого запису

На рис.4.2 горизонтальна спектрограма дана в трьох координатах - час (горизонтальна вісь), частота (вертикальна вісь) та амплітуда (задана градацією сірого, чим темніше, тим більше амплітуда). На спектрограмі можна візуально побачити шуми та сигнал. Проведемо аналіз спектрограми на рис.4.2.

Особливості, які можна побачити на спектрограмі:

- горизонтальні лінії на низьких і середніх частотах – це наведення, що гудуть (частина шуму);
- візерунчастий фон на всіх частотах – це широкосмуговий шум;
- хвилясті лінії – це гармоніки (обертони) голосу;
- вертикальні світлі області на СЧ та ВЧ – це шиплячі та свистячі приголосні;
- шумова хмара в районі 2 с – це вдих;
- вертикальні штрихи в районі 1,7 с – це «плямкання», що супроводжує відкривання рота.

Тестовий зашумлений запис оброблений у програмі Audacity.

При обробці вибрано такі параметри шумоподавлення:

- поріг шумоподавлення: -20 дБ;
- подавлення шуму: 48 дБ;
- спектральне розрізнення: 20 Гц;
- час атаки / відновлення: 120 мс.

Параметри шумоподавлення були обрані з таких міркувань. Типова кількість частотних смуг в алгоритмі спектрального віднімання – близько тисячі. Це означає, що ширина смуг становить близько 20 Гц. Така смуга дозволяє досить точно розбивати гармоніки багатьох сигналів по окремим частотним смугам. А це, у свою чергу, дозволяє більш надійно відокремити смуги, що містять сигнал, від шумових смуг і буквально відфільтрувати шум між гармоніками сигналу.

Збільшення числа смуг має і негативний ефект. Разом із збільшенням частотного розрізнення спектрального аналізу погіршується його часове розрізнення. Це пов'язано з тим, що імпульси відповідних розділових фільтрів стають довшими за часом і захоплюють інформацію про більш довгі ділянки сигналу. Це призводить до того, що спектрограма розмивається за часом і втрачається можливість точно локалізувати початок і кінець звучання тих чи інших подій аудіосигналу. Це ускладнює подавлення шуму поблизу транз'єнтів – різких сплесків енергії аудіосигналу, які відповідають атакам

нот або ударам перкусійних інструментів. В результаті навколо транзйентів з'являються області слабого шуму, а самі транзйенти можуть розмиватися за часом і втрачати чіткість.

Спектрограма обробленого сигналу показана на рис.4.3, б).

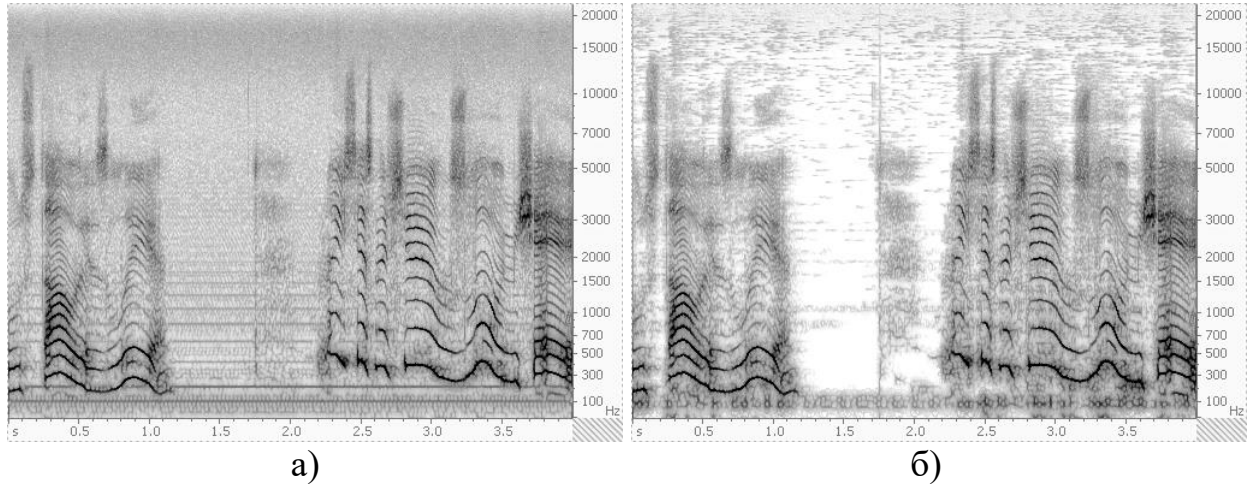


Рисунок 4.3 – Спектрограми вихідного зашумленого запису (а) та обробленої алгоритмом спектрального віднімання (б)

При аналізі спектрограми на рис.4.3 б і при прослуховуванні запису обробленого звуку (файл Audacity.wav на диску додатку) можна відзначити наступне:

- загальний рівень шуму у записі зменшився, але з'явилися деякі спотворення (артефакти);
- з'явився гул на низьких частотах (горизонтальна хмара внизу спектрограми);
- з'явилося клацання в районі 1,7 с., пов'язане з пропусканням та посиленням таким чином «плямкання»;
- з'явилася велика кількість "музичного шуму" – локальних сплесків енергії на спектрограмі.

"Музичний шум" є однією з основних проблем алгоритму спектрального віднімання. Він виникає внаслідок того, що коефіцієнти STFT шумових сигналів статистично випадкові, що призводить до їх нерівномірного подавлення. В результаті, очищений сигнал містить короточасні та обмежені по частоті сплески енергії, які на слух

сприймаються як "дзвіночки" або "вода, що ллється". У деяких випадках цей ефект навіть менш бажаний, ніж вхідний шум, що подавлюється. "Музичний шум" часто більш помітний, ніж вхідний рівний шум, так як слуху легше адаптуватися до рівного шуму і відключитися від нього.

Для дослідження характеру «музичного шуму» було досліджено роботу алгоритму спектрального віднімання на білому шумі. Для цього у програмі Audacity було створено шумову доріжку з білим шумом тривалістю 2 с. Після цього ця доріжка була оброблена шумоподавником спектрального віднімання з параметрами, зазначеними вище. Спектрограми вхідної та обробленої доріжок отримані у програмі Right Mark Audio Analyzer.

На рис. 4.4 показана спектрограма білого шуму.

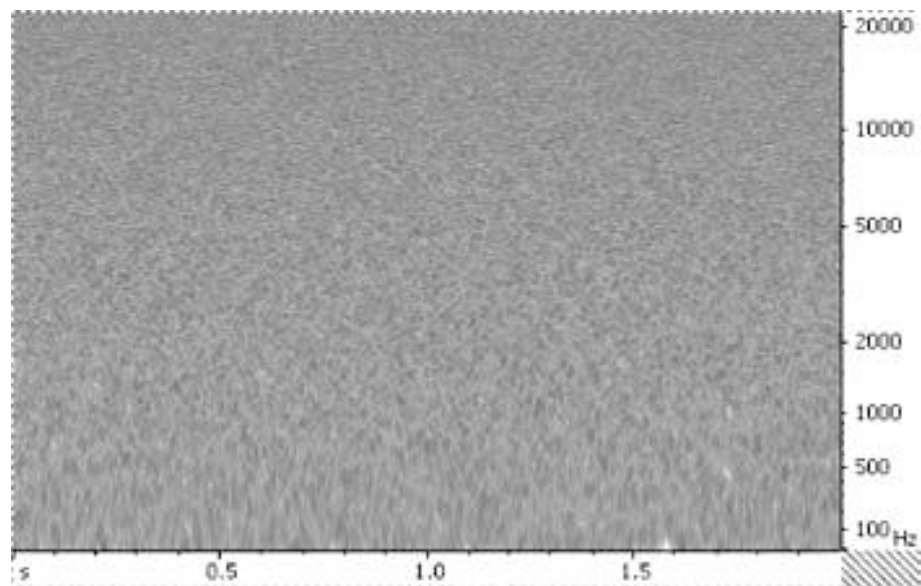


Рисунок 4.4 – Спектрограма білого шуму

Як очевидно з рис.4.4, спектрограма білого шуму перестав бути константою. Спектрограма сама статистично випадковою і коливається навколо деякої середньої спектральної щільності шуму. Це призводить до того, що в процесі спектрального віднімання окремі точки на спектрограмі пригнічуються слабкіше за інших. Вони й утворюватимуть «музичний шум», спектрограма якого показана на рис. 4.5.

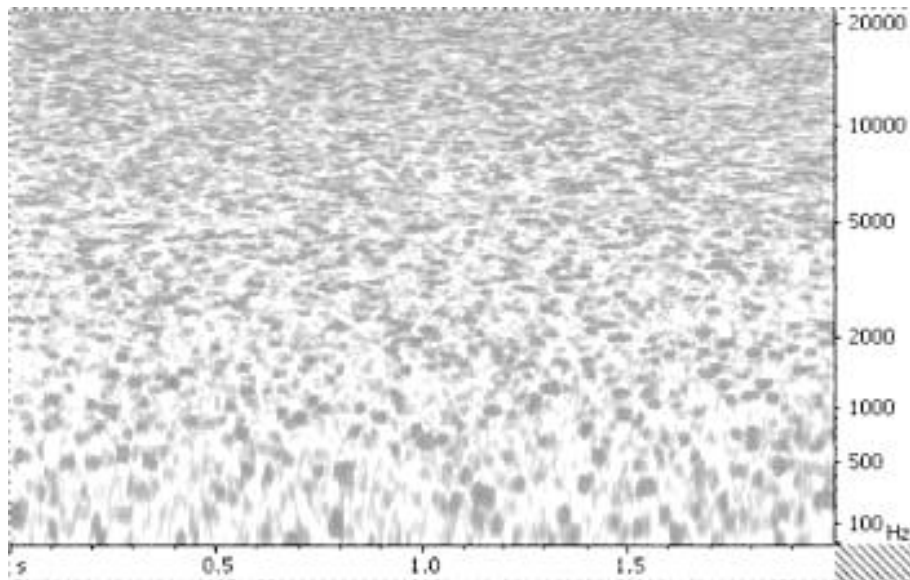


Рисунок 4.5 – Спектрограма «музыкального шуму»

Записи вхідної доріжки з білим шумом та обробленої доріжки з «музичним шумом» знаходяться у файлах `WhiteNoise.wav` та `MusicNoise.wav` на диску програми.

Існує кілька способів боротьби з «музичним шумом». Найпростіший спосіб – завищення оцінки спектру шуму, внаслідок чого шум буде подавлятися сильніше і ефект музичного шуму не виникатиме. Однак, разом з цим будуть сильніше подавлюватись низькорівневі компоненти корисного сигналу, що негативно позначиться на тембрі.

Інший спосіб – неповне подавлення шуму. Це дозволяє залишити в сигналі частину шуму, що подавляється, для маскуванню музичного шуму, що виникає. Цей спосіб не усуває музичний шум, а лише робить його менш помітним при слабкому подавленні шуму.

Найпопулярніший спосіб боротьби з музичним шумом – використання згладжених оцінок спектра сигналу. Згладжування зазвичай здійснюється простим рекурсивним фільтром вздовж осі часу. Якщо розглядати алгоритм спектрального віднімання як багатосмуговий гейт, то таке згладжування аналогічне до введення часу атаки/відновлення в роботу гейта. На рис. 4.6 показаний результат застосування алгоритму спектрального віднімання у програмі Audacity зі збільшеним до 500 мс часом атаки/відновлення.

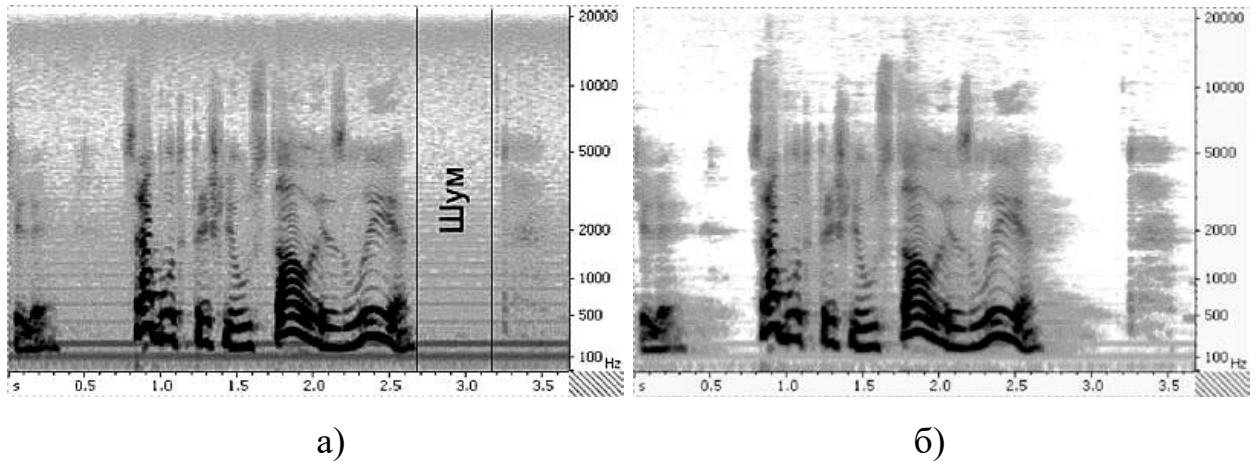


Рисунок 4.6 – Результат застосування алгоритму спектрального віднімання зі збільшеним до 500 мс часом атаки/відновлення (спектрограми зашумленого (а) та очищеного (б) сигналів)

Як видно з рис .4.6, збільшення часу атаки/відновлення призводить до уповільнення реакції гейту на сплески низького рівня, а також утворення ехоподібних шлейфів непригніченого шуму після спадів енергії сигналу.

Досліджуємо алгоритми шумоподавлення, що застосовуються у сучасних звукових редакторах.

4.3 Дослідження алгоритмів шумоподавлення у звукових редакторах

Більшість сучасних редакторів звуку не дають жодної інформації про особливості алгоритмів, що використовуються для подавлення широкосмугового шуму. Відомо лише, що в основі лежить алгоритм спектрального віднімання. Аналіз великої кількості Інтернет-матеріалів дозволив зробити такі узагальнення.

Для досягнення найкращого компромісу між частотною та часовою роздільною здатністю спектрограми більшість алгоритмів фіксують число частотних смуг на деякій середній величині (порядку 1000). Проте є і складніші алгоритми, що здійснюють так званий аналіз мультирозрізнення (multiresolution), адаптивно змінюючи розрізнення аналізу фонограми задля

досягнення найкращої чіткості спектрограми на кожній її ділянці. Таким чином, вони досягають високої часової роздільної здатності поблизу транз'єнтів і високої частотної роздільної здатності для тональних сигналів.

Складні алгоритми подавлення музичного шуму можуть проводити як згладжування оцінок спектру за часом, а також й згладжувати за частотою. Найбільш складні алгоритми згладжування спектрограм запозичують алгоритми з області обробки зображень [8]. Вони аналізують двовимірну структуру спектрограми, намагаючись виявити ознаки корисного сигналу, втраченого у шумі. Це дозволяє усунути "шумові ехо" після спадів енергії сигналу.

Тестовим записом є мова на фоні гулу (основний тон і гармоніки) і шипіння (широкополосний шум). Спектрограму вихідного зашумленого запису отримано у програмі Right Mark Audio Analyzer і показано на рис. 4.2.

Досліджено роботу шумоподавлювача у VST-плагіні Voxengo Redunoise 1.5. Плагін було підключено до звукового редактора Cubase.

На рис. 4.7 показано результат роботи плагіна Voxengo Redunoise 1.5.

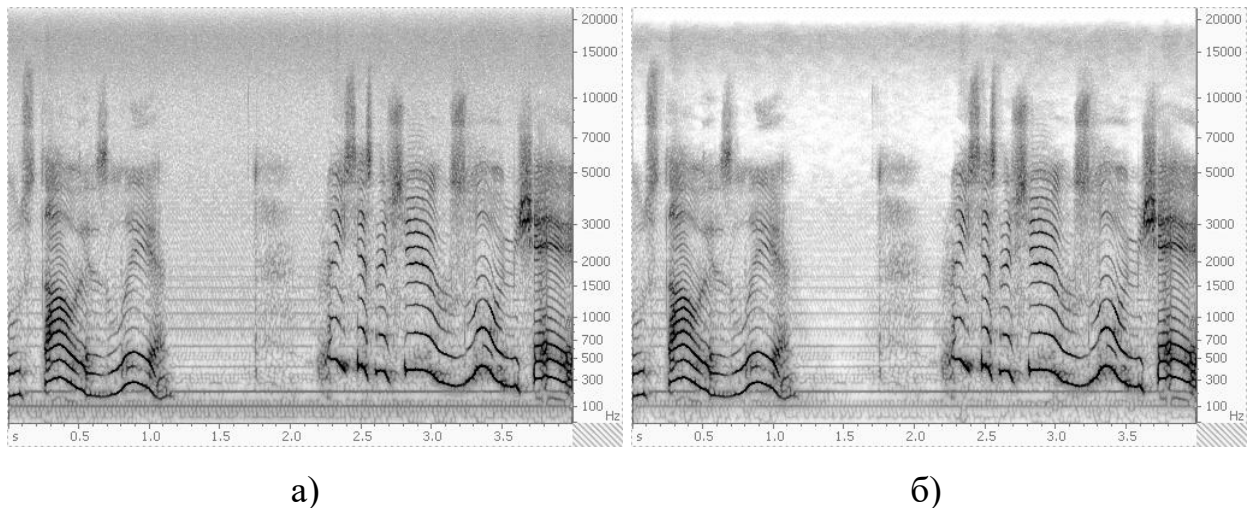


Рисунок 4.7 – Спектрограми вихідного зашумленого запису (а) та обробленою плагіном Voxengo Redunoise 1.5 (б)

Прослухати результати роботи фільтра можна у файлі Voxengo.wav на диску додатку.

Для шумоподавлення були вибрані такі параметри фільтра: linear phase (лінійна фаза), max suppression (максимальне подавлення). Зі спектрограми на рис.4.7 та при прослуховуванні аудіофайлу можна відзначити недолік плагіна Voxengo Redunoise – слабе подавлення гулу. Подавлення широкосмугового шуму також невисоке.

Досліджено роботу шумоподавлювача у плагіні Sonic Foundry Noise Reduction 2.0. Плагін було підключено до звукового редактора Sound Forge.

На рис. 4.8 показано результат роботи плагіна Sonic Foundry Noise Reduction 2.0.

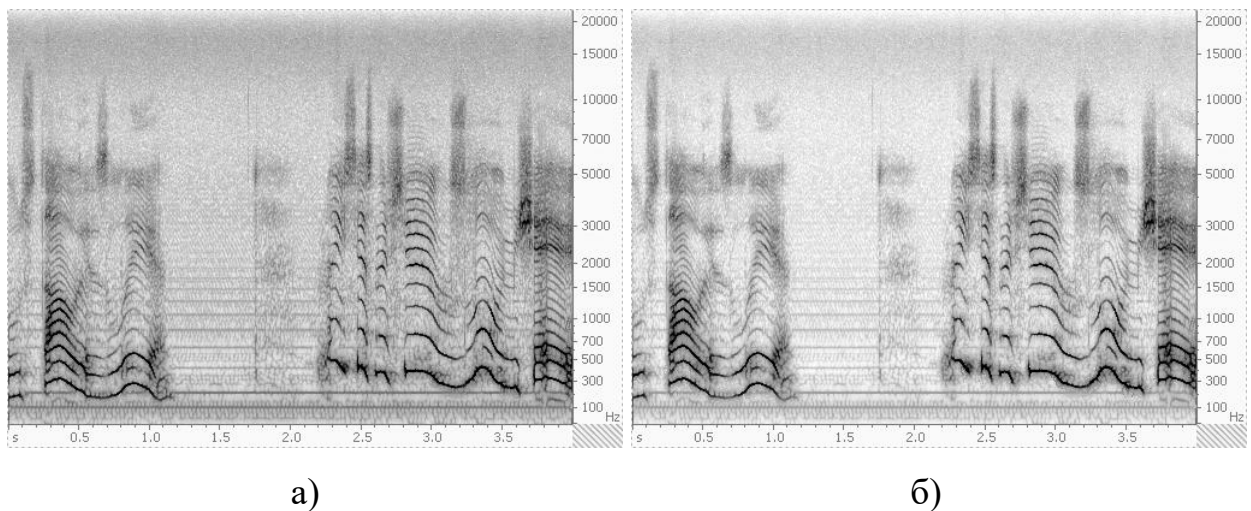


Рисунок 4.8 –Спектрограми вхідного зашумленого запису (а) та обробленої плагіном Sonic Foundry Noise Reduction 2.0 (б)

Прослухати результат роботи фільтра можна у файлі Sony.wav на диску програми.

Для шумоподавлення були обрані такі параметри фільтра: mode 2 (режим 2), reduction 25 dB (подавлення 25 дБ). Зі спектрограми на рис.4.8 і при прослуховуванні аудіофайлу можна відзначити недолік плагіна Voxengo Redunoise – неповне подавлення гулу. Подавлення широкосмугового шуму також невисоке.

Досліджено шумоподавлення у програмі iZotope RX (алгоритм C). У цьому алгоритмі використовується аналіз з мультироздільною здатністю (multiresolution), який адаптивно варіює частотне і часове розрізнення аналізу

фонограми для досягнення найкращої чіткості спектрограми на кожній її ділянці. Таким чином намагаються досягти високої часової роздільної здатності поблизу сплесків енергії та високої частотної роздільної здатності для тональних сигналів.

На рис. 4.9 показаний результат роботи алгоритму шумоподавлення у програмі iZotope RX. Для шумоподавлення були обрані такі параметри фільтра: алгоритм, пригнічення 15 дБ.

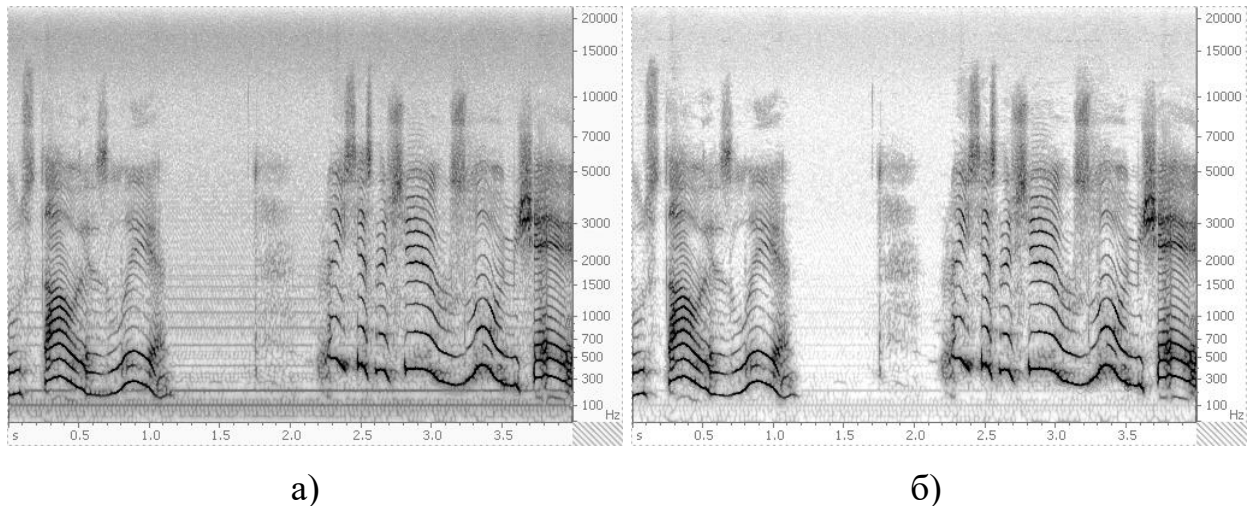


Рисунок 4.9 – Спектрограми вихідного зашумленого запису (а) та обробленого програмою iZotope RX(б)

Прослухати результат роботи фільтра можна у файлі iZotope_RX.wav на диску додатку.

4.4 Висновки по розділу 4

Аналіз спектрограм на рис.4.7 – 4.9 (а також тих, що надані в додатку А) та отриманих аудіозаписів показує, що алгоритм спектрального віднімання з мультирозрізненням дозволяє отримати досить якісне подавлення гулу і широкосмугового шуму при мінімальних спотвореннях сигналу в області сплеску і згасання.

Результати тестування шумоподавлювачів в інших звукових редакторах наведено в додатку А.

ВИСНОВКИ

Зменшення шуму у звукозаписі є дуже актуальною задачею. З нею звукорежисери стикаються з ранніх часів і до цього дня. В даний час існує багато програм – звукових редакторів, випускається безліч плагінів до цих програм, які здійснюють шумоподавлення тим чи іншим методом. Опис для цих плагінів та програм носить в основному рекламний характер і не розкриває всіх подробиць алгоритму та детальних результатів обробки.

У кваліфікаційній роботі виконано прикладні дослідження більше десятиох популярних шумоподавлювачів на тестових фонограмах та проведено детальний порівняльний аналіз результатів обробки.

Розглядаються адитивні стаціонарні шуми. Тобто, потужність та спектр шуму не змінюються в часі. Це шум магнітної стрічки, шум наведень електромережі, шипіння мікрофонного підсилювача, шум вентиляції в студії і т.д.

Мета дослідження – визначення числових і якісних характеристик різних алгоритмів шумоподавлення.

В якості початкового матеріалу використано закадровий голос диктора, записаний на телецентрі ХНУРЕ. Потім цей звук був зашумлений записом гулу електродвигуна. Змікшований запис використовувався для подальших досліджень. Відношення сигнал-шум склало близько 30 дБ.

Проаналізовано та досліджено базові методи подавлення шуму. Найпростішим є пороговий шумоподавлювач (гейт). Він зменшує підсилення тракту в паузах, коли шуми проявляються найсильніше. Налаштування гейту – поріг спрацьовування, час атаки та відновлення. Поріг спрацьовування встановлюють так, щоб гейт надійно відкривався сигналом і не відкривався шумом. Атака та відновлення допомагають згладити процеси відкриття та закриття. Ступінь подавлення шуму може залежати від рівня вхідного сигналу, а передавальна характеристика може мати "м'який поріг".

Дослідження роботи гейту у програмі Sound Forge. Можна зробити висновок, що Noise Gate повністю подавляє шум у паузах. До недоліків можна віднести ефект модуляції шуму сигналами низької амплітуди та подавлення корисних сигналів низької амплітуди. Фактично гейт подавив шум у паузах, а під час звучання його дії немає. Це істотний недолік, особливо при малому відношенні сигнал/шум. В цілому, гейт дозволяє зменшити сумарний рівень шуму в каналі відтворення на 1...2 дБ.

Широко відомі динамічний шумоподавлювач (DNR) та динамічний обмежувач шуму (DNL), які змінюють ефективну смугу фільтра в області ВЧ (а іноді і НЧ) залежно від наявності сигналу. Своєрідною комбінацією гейта та динамічного шумоподавлювача є багатосмуговий гейт. У ньому завдяки наявності декількох частотних смуг можна вибірково подавити тільки ті смуги, які в даний момент містять шум.

Проведено дослідження роботи трисмугового гейту Multiband NoiseGate NX3, випущеного фірмою 123creative.com у форматі VST-плагіну. Плагін було підключено до звукового редактора FL Studio. Частотні смуги мали межі: 0...700 Гц, 0,7...3,2 кГц, 3,2 кГц і вище. У всіх смугах поріг -20 дБ, час атаки 3 мс, час відновлення 100 мс.

За результатами обробки можна зробити висновки, що трисмуговий гейт погіршує якість звучання менше ніж односмуговий. Однак чутно, як у паузах широкосмугові гейти відкриваються від слабких звуків, тембр звучання приглушений у порівнянні з вихідним записом. Тобто, основні недоліки трисмугового гейту пов'язані з малим числом смуг.

Розвитком багатосмугового нелінійного фільтра є алгоритм спектрального віднімання. Він складається з наступних етапів: розкладання сигналу за допомогою короткочасного перетворення Фур'є (STFT); оцінка діапазону шуму; «віднімання» амплітудного спектра шуму з амплітудного спектра сигналу; зворотне перетворення STFT та синтез результуючого сигналу.

Сигнал розбивається на ряд інтервалів за допомогою ковзного вікна. Найчастіше використовують вікно Хана тривалістю близько 50 мс та ступенем перекриття 75%. Вибір сегмента шуму може здійснюватись як автоматично, так і вручну.

Такий алгоритм дозволяє аналізувати та обробляти динамічні спектри. Спектрограма будується у трьох координатах – частота (по вертикалі), час (по горизонталі) та амплітуда (відтінок кольору).

Алгоритм спектрального віднімання широко використовується як сам по собі, так і у складі складніших методів, у багатьох редакторах звуку. Чистий метод спектрального віднімання реалізований у програмі Audacity.

Спектрограма вихідного зашумленого запису отримано у програмі Right Mark Audio Analyzer. На спектрограмі можна побачити шуми та сигнал. Горизонтальні лінії на низьких та середніх частотах – це наведення, що гудуть. Візерунковий фон на всіх частотах – це широкопasmовий шум. Хвилясті лінії – це гармоніки голосу. Вертикальні світлі області на середніх та верхніх частотах – це шиплячі та свистячі приголосні. Шумова хмара в районі 2 с – це вдих. Вертикальні штрихи в районі 1,7 с – це «плямкання» при відкритті рота.

Тестовий зашумлений запис оброблений у програмі Audacity. Можна відзначити, що загальний рівень шуму в записі зменшився, але з'явилися деякі спотворення: гул на низьких частотах (горизонтальна хмара внизу спектрограми); клацання в районі 1,7 с, пов'язане з пропусканням та посиленням «плямкання»; з'явилася велика кількість "музичного шуму" – локальних сплесків енергії на спектрограмі.

"Музичний шум" є однією з проблем методу спектрального віднімання. Він з'являється внаслідок того, що очищений сигнал містить короткочасні та обмежені за частотою сплески енергії, які на слух сприймаються як "дзвіночки" або "вода, що ллється".

Існує кілька способів боротьби з «музичним шумом». Перший спосіб – це завищення оцінки спектру шуму, внаслідок чого шум подавлятиметься

сильніше і ефект музичного шуму не виникатиме. Але разом з цим сильніше будуть подавлені низькорівневі компоненти корисного сигналу, що негативно позначиться на тембрі. Другий спосіб – це неповне подавлення шуму. Цей спосіб не усуває музичний шум, лише маскує його при слабкому шумоподавленні. Третій спосіб – збільшити час атаки/відновлення, при подавленні шуму. Це запобігає короткочасному проходженню сплесків шуму, але погіршить реакцію фільтра на різкі сплески звукової енергії.

Найбільш ефективний спосіб боротьби з музичним шумом – використання алгоритму мультирозрізнення, що адаптивно змінює дозвіл аналізу фонограми для досягнення найкращої чіткості спектрограми на кожній її ділянці. Таким чином досягають високої часової роздільної здатності поблизу сплесків звукової енергії і високої частотної роздільної здатності для тональних сигналів.

Отримано результат роботи алгоритму мультирозрізнення у програмі iZotope RX. Аналіз спектрограм та отриманих аудіозаписів показує, що алгоритм спектрального віднімання з мультирозрізненням дозволяє отримати досить хороше подавлення гулу і широкосмугового шуму при мінімальних спотвореннях сигналу в області сплеску і згасання.

Взагалі, в роботі отримані результати тестування шумоподавлювачів в 11 звукових редакторах. Отримані в роботі результати можуть бути використані фахівцями зі створення аудіопродукції, а також у лабораторних роботах з дисциплін аудіоциклу.

ПЕРЕЧЕНЬ ССЫЛОК

1. Электроакустика и звуковое вещание: Учебное пособие для вузов / И.А. Алдошина, Э.И. Вологдин и др.; Под ред. Ю.А. Ковалгина. – М.: Горячая линия-Телеком, Радио и связь, 2007. – 872 с.
2. Цифровая звукозапись. Под редакцией Дж. Мааса и М. Веркамена. Пер. с англ. – М.: Мир, 2004. – 352 с.
3. Никамин В.А. Цифровая звукозапись. Технология и стандарты. – СПб: Наука и техника, 2002. – 256 с.
4. Ковалгин Ю.А., Вологдин Э.И. Цифровое кодирование звуковых сигналов. – СПб.: Корона Принт., 2004. – 240 с.
5. Рабинер Л., Гоулд Б. Теория и применение цифровой обработки сигналов. – М.: Мир, 1978. – 848 с.
6. Чернецкий М. Системы шумоподавления [Электронный ресурс] // Звукорежиссер. № 9. – 2001. <http://rus.625-net.ru/archive/z0901/r1.htm> (дата звернения 01.12.2021).
7. Лукин А. Подавление широкополосного шума: история и новые разработки [Электронный ресурс] // Звукорежиссер. №10. – 2008. – <http://audioproducer.625-net.ru/files/601/523> (дата звернения 01.12.2021).
8. Кен Гендри. Системы шумоподавления [Электронный ресурс] // Звукорежиссёр, №6, 2004. – <http://rus.625-net.ru/archive/z0604/noise.htm> (дата звернения 01.12.2021).
9. Лукин А. Введение в цифровую обработку сигналов (Математические основы). Методическое пособие, 2002, 44 с.
10. Лядов М. Лукин А. RightMark Audio Analyzer 5.5. Руководство пользователя. – IXBT.com, 2005. – 17 с.
11. Квинт И. Sound Forge 9. – СПб.: Питер, 2009. – 167 с.
12. Самоучитель по музыкальным редакторам: WaveLab, Sound Forge, Cubase, Acid. – М.: ИДДК, 2007. – 212 с.

13. Степаненко О.С. Adobe Audition 3.0 создание фонограмм и обработка звука. – М.: Диалектика, 2010. – 414 с.

14. Горюнов В.А., Стась А.Н. Обработка и монтаж аудиозаписей с использованием Audacity (ПО для обработки и монтажа аудиозаписей): Учебное пособие. — Москва: 2008. – 40 с.

15. В.Н. Олейников, О.В. Зубков, В.М. Карташов, И.В. Корытцев, С.И. Бабкин, С.А. Шейко, И.С. Селезнев. Экспериментальная оценка эффективности алгоритмов пеленгования беспилотных летательных аппаратов по акустическому излучению. Радиотехника: Всеукр. межвед. науч.-техн. сб. – 2019. – Вып. 199. – С. 29 – 37.

16. V. Kartashov, V. Oleynikov, I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin, I. Selieznov. Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles. 2020 IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). 2020. 4 p.

17. Kartashov V.M., Oleynikov V.N, Zubkov O.V., Koryttsev I.V., Babkin S. I., Sheiko S.A., Kolendovskaya M.M. Spatial-temporal Processing of acoustic Signals of Unmanned Aerial Vehicles/ Telecommunications and Radio Engineering. – New York. – 2020. – Vol. 79, №9. – P.769-780.

18. V. Kartashov, V. Oleynikov , I. Koryttsev, S. Sheiko, O. Zubkov, S. Babkin. Processing of Wide Band Acoustic Signals During Detection of Unmanned Aerial Vehicles // 2020 IEEE Ukrainian Microwave Week (UkrMW). Kharkiv, Ukraine, September 21 - 25, 2020. Volume 1 on 2020 IEEE 12th International Conference on Antenna Theory and Techniques (ICATT). pp. 35-39.

19. V.M. Kartashov, G.I. Sidorov, S.A. Sheiko, M.M. Kolendovskaya, O.Yu. Sergienko. Principles of construction and assessment of technical characteristics of multi-frequency atmospheric sodar in the humidity measurement mode. Telecommunications and Radio Engineering. Vol. 79. N.4. 2020. – pp. 323-333.

20. S. Sheiko. Study of the method for assessing atmospheric turbulence by the envelope of sodar signals // *Eastern-European Journal of Enterprise Technologies.* – 2/5 (92). – April, 2018. – p. 33–40.

21. Сідоров Г.І., Шейко С.О., Шаповалов С.В., Полонська А.С., Дмитренко А.І. Акустичний метод вимірювання турбулентного стану атмосферного прикордонного шару // *Радиотехніка: Всеукр. міжвід. наук.-техн. зб.* 2018. – Вип. 192. – С. 46–50.

22. Valerii V. Semenets, V. M. Kartashov, V. I. Leonidov. Registration of refraction phenomenon in the problem of acoustic sounding of atmosphere in airports zone. *Telecommunications and Radio Engineering.* Volume 77, Issue 5, 2018. – P. 461-468.

23. Буйницький Д.В. Апаратно-программний комплекс для дослідження акустических пристроїв // 23-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». – Харків, 16–18 квітня 2019 р. – с. 92-93.

24. Ашихмин В.О. Исследование системы коррекции звука с учётом характеристик помещения // 23-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». – Харків, 16–18 квітня 2019 р. – с. 96-97.

25. Тушев В.О. Исследование влияния фазовых искажений аудиоаппаратуры на качество звучания // 23-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». – Харків, 16–18 квітня 2019 р. – с. 84-85.

26. Чернов К.А. Исследование акустических характеристик студии звукозаписи // 23-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». – Харків, 16–18 квітня 2019 р. – с. 94-95.

27. Ашихмін В.О. Адаптивні методи подавлення акустичного зворотного зв'язку: атестаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2019 р. – 69 с.

28. Буйницький Д.В. Дослідження методів компенсації ревербераційних завад: атестаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2019 р. – 67 с.

29. Шевченко П.Д. Система вимірювання та аналізу акустичних характеристик приміщень: атестаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2019 р. – 68 с.

30. Лісовий Р.О. Комплексне дослідження акустичних характеристик студійного приміщення: атестаційна робота на здобуття освітнього ступеня магістра. – Х.: ХНУРЕ. – 2019 р. – 81 с.

31. Медінцев Р.В. Використання моделей лінійного передбачення для факторизації параметричних спектрів /25-й Міжнародний молодіжний форум «Радіоелектроніка та молодь у ХХІ столітті». Зб. матеріалів форуму. Т. 3. – Харків: ХНУРЕ. 2021. – с. 121 – 122.

32. Методичні вказівки з виконання атестаційної магістерської роботи за спеціальністю 8.05090102 «Апаратура радіозв'язку, радіомовлення і телебачення». Освітньо-кваліфікаційний рівень – магістр / Упоряд. В.М. Карташов, В.А. Тихонов, І.В. Савченко – Харків: ХНУРЕ, 2012. – 68 с.