ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ТЕЛЕКОМУНІКАЦІЙ
КАФЕДРА ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ ТА МЕРЕЖ

КВАЛІФІКАЦІЙНА РОБОТА

на тему: «Застосування конволюційних фільтрів для шумозаглушення у відеоконференц-зв'язку»

на здобуття освітнього ступеня магістра
зі спеціальності 172 Електронні комунікації та радіотехніка
(код, найменування спеціальності)
освітньо-професійної програми Телекомунікаційні системи та мережі
(назва)

Кваліфікаційна робота містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело.

_____ Леонтій ЗАРЕЦЬКИЙ
(підпис) *(Ім'я, ПРІЗВИЩЕ здобувача)*

Виконав: здобувач вищої освіти
студент гр. ТСДМ-61
Леонтій ЗАРЕЦЬКИЙ
(Ім'я, ПРІЗВИЩЕ)

Керівник: Дмитро Миронов
к.т.н. *(Ім'я, ПРІЗВИЩЕ)*

Рецензент: _____
_____ *(Ім'я, ПРІЗВИЩЕ)*
(науковий ступінь, вчене звання)

Київ 2026

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
Навчально-науковий інститут Телекомунікацій

Кафедра ___Телекомунікаційних систем та мереж_____
Ступінь вищої освіти __магістр_____
Спеціальність ___172 Електронні комунікації та радіотехніка _____
Освітньо-професійна програма _____Телекомунікаційні системи та мережі _____

ЗАТВЕРДЖУЮ
Завідувач кафедрою __ТСМ_____
_____ Віктор ЗАЇКА
«____» _____ 2026 р.

ЗАВДАННЯ
НА КВАЛІФІКАЦІЙНУ РОБОТУ

Зарецькому Леонтію Руслановичу
(прізвище, ім'я, по батькові здобувача)

1. Тема кваліфікаційної роботи: Застосування конволюційних фільтрів для шумозаглушення у відеоконференц-зв'язку,
керівник кваліфікаційної роботи Дмитро МИРОНОВ, к.т.н.,
(Ім'я, ПРІЗВИЩЕ науковий ступінь, вчене звання)
затверджені наказом Державного університету інформаційно-комунікаційних технологій від «30» жовтня 2025 р. № 467.

2. Строк подання кваліфікаційної роботи «___» _____ 2025 р.

3. Вихідні дані до кваліфікаційної роботи:
1) рекомендації міжнародної спілки електрозв'язку у сфері цифрового відеозв'язку та мультимедійних комунікацій, зокрема рекомендації Міжнародного союзу електрозв'язку (ITU-T) щодо передавання відеосигналів у реальному часі, параметрів якості обслуговування (QoS), допустимих затримок та вимог до якості відео в інтерактивних системах зв'язку;
2) відомості про сучасний стан та тенденції розвитку систем відеоконференцзв'язку;
3) дані про особливості проектування, реалізації та експлуатації систем відеоконференцзв'язку.
4) матеріали по оформленню пояснювальної записки і креслень.

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити):
1) вступ;
2) шуми та методи їх заглушення;
3) теоретичні засади шумозаглушення в умовах відеоконференцзв'язку;
4) впровадження методів шумозаглушення відео у системах відеоконференцзв'язку;

5. Перелік графічного матеріалу: ___презентація на 12 слайдах_____

6. Дата видачі завдання «_01_» _жовтня_ 2025 р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів кваліфікаційної роботи	Строк виконання етапів роботи	Примітка
1	Вступ	05.10.2025	
2	Шуми та методи їх заглушення	10.10.2025	
3	Теоретичні засади шумозаглушення в умовах відеоконференцзв'язку	15.10.2025	
4	Впровадження методів шумозаглушення відео у системах відеоконференцзв'язку	30.10.2025	
5	Оформлення роботи та презентації	25.12.2025	

Здобувач вищої освіти _____ Леонтій ЗАРЕЦЬКИЙ
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

Керівник роботи _____ Дмитро МИРОНОВ
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

РЕФЕРАТ

Текстова частина кваліфікаційної роботи на здобуття освітнього ступеня магістра: 65 стор., 12 рис., 3 табл., 10 джерел.

Мета роботи – дослідження та аналіз методів шумозаглушення відео у системах відеоконференцзв'язку, розробка та обґрунтування ефективного підходу до обробки відеопотоку в режимі реального часу з використанням конволюційних методів.

Об'єкт дослідження – системи відеоконференцзв'язку, що забезпечують передавання відеосигналу в реальному часі в умовах обмежених обчислювальних та мережевих ресурсів.

Предмет дослідження – методи та алгоритми шумозаглушення відеопотоку, зокрема конволюційні фільтри та компактні конволюційні нейронні мережі, а також їх інтеграція у відеопайплайн систем відеоконференцзв'язку.

Короткий зміст роботи:

У кваліфікаційній роботі проведено аналіз сучасного стану та тенденцій розвитку систем відеоконференцзв'язку з точки зору якості відеосигналу. Проаналізовано класичні методи фільтрації та сучасні підходи на основі конволюційних нейронних мереж.

Запропоновано гібридний конвейер обробки відеопотоку. Проведено математичне моделювання роботи запропонованого конвейеру та експериментальну оцінку його ефективності. Показано, що використання легковагових конволюційних моделей дозволяє суттєво підвищити якість відео у системах відеоконференцзв'язку без перевищення допустимих затримок та збереження вимог реального часу.

КЛЮЧОВІ СЛОВА:
ВІДЕОКОНФЕРЕНЦЗВ'ЯЗОК, ШУМОЗАГЛУШЕННЯ, КОНВОЛЮЦІЙНІ ФІЛЬТРИ, КОНВОЛЮЦІЙНІ НЕЙРОННІ МЕРЕЖІ, ОБРОБКА ВІДЕО В РЕАЛЬНОМУ ЧАСІ, ЯКІСТЬ ВІДЕО.

ABSTRACT

The textual part of the qualification thesis for obtaining the Master's degree: 65 pages, 12 figures, 3 tables, 10 references.

The aim of the work is to study and analyze video denoising methods in video conferencing systems, as well as to develop and substantiate an effective approach to real-time video stream processing using convolutional methods.

The object of the study is video conferencing systems that provide real-time video signal transmission under conditions of limited computational and network resources.

The subject of the study is methods and algorithms for video stream denoising, in particular convolutional filters and compact convolutional neural networks, as well as their integration into the video processing pipeline of video conferencing systems.

Summary of the work:

The qualification thesis analyzes the current state and development trends of video conferencing systems from the perspective of video signal quality. Classical filtering methods and modern approaches based on convolutional neural networks are examined.

A hybrid video processing pipeline is proposed. Mathematical modeling of the proposed pipeline is carried out, along with an experimental evaluation of its effectiveness. It is shown that the use of lightweight convolutional models makes it possible to significantly improve video quality in video conferencing systems without exceeding acceptable latency limits while maintaining real-time performance requirements.

KEYWORDS:

VIDEO CONFERENCING, DENOISING, CONVOLUTIONAL FILTERS, CONVOLUTIONAL NEURAL NETWORKS, REAL-TIME VIDEO PROCESSING, VIDEO QUALITY.

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
Навчально-науковий інститут телекомунікацій

ПОДАННЯ

ГОЛОВІ ЕКЗАМЕНАЦІЙНОЇ КОМІСІЇ ЩОДО ЗАХИСТУ КВАЛІФІКАЦІЙНОЇ РОБОТИ
на здобуття освітнього ступеня магістра

Направляється здобувач Зарецький Л.Р. до захисту кваліфікаційної роботи
(прізвище та ініціали)
за спеціальністю 172 Електронні комунікації та радіотехніка
(код, найменування спеціальності)
освітньо-професійної програми Телекомунікаційні системи та мережі
(назва)
на тему: «Застосування конволюційних фільтрів для шумозаглушення у відеоконференц-зв'язку».

Кваліфікаційна робота і рецензія додаються.

Директор ННІТ _____ Владислав КРАВЧЕНКО
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

Висновок керівника кваліфікаційної роботи

Здобувач Зарецький Л.Р. під час написання роботи проявив себе як вдумливий та допитливий фахівець, показав гарну теоретичну підготовку. За короткий час їм було опрацьовано значну кількість літератури з проблем побудови, проектування та тестування систем відеоконференцзв'язку та проведені необхідні розрахунки. Працюючи над завданнями, які доручалися керівником роботи, студент проявив ініціативу творчість і спроможність до інженерної праці.

Все це дозволяє оцінити виконану кваліфікаційну роботу здобувача Ляшкевича Є.Р. на оцінку
« відмінно » та присвоїти йому кваліфікацію магістр з електронних комунікацій та радіотехніки.

Керівник роботи _____ Дмитро МИРОНОВ
(підпис) *(Ім'я, ПРІЗВИЩЕ)*

«____» _____ 20___ року

Висновок кафедри про кваліфікаційну роботу

Кваліфікаційна робота розглянута. Здобувач Зарецький Л.Р. допускається до захисту даної роботи в Екзаменаційній комісії.

Завідувач кафедрою телекомунікаційних систем та мереж _____ Віктор ЗАЇКА
(назва) *(підпис)* *(Ім'я, ПРІЗВИЩЕ)*

ВІДГУК РЕЦЕНЗЕНТА
на кваліфікаційну роботу
на здобуття освітнього рівня магістра

здобувача вищої освіти Зарецького Леонтія Руслановича _____
(прізвище, ім'я, по батькові)
на тему «Застосування конволюційних фільтрів для шумозаглушення у відеоконференцзв'язку»

Актуальність.

Однією з ключових проблем сучасних систем відеоконференцзв'язку є наявність шумів та артефактів у відеопотоці. Актуальним є дослідження та вдосконалення методів шумозаглушення відео, зокрема з використанням конволюційних фільтрів, які забезпечують ефективну обробку відеопотоку без порушення вимог реального часу. Аналіз, моделювання та практичне обґрунтування таких підходів є важливою науково-практичною задачею.

Позитивні сторони.

У роботі виконано ґрунтовний аналіз сучасного стану та тенденцій розвитку систем відеоконференцзв'язку з точки зору обробки відеосигналу. Детально описано структуру конвейеру обробки відеопотоку. Пояснювальна записка викладена логічно, послідовно та зрозуміло, використано сучасну науково-технічну термінологію. Графічний матеріал є наочним, коректно оформленим і сприяє кращому розумінню принципів роботи запропонованої системи шумозаглушення відео.

Недоліки.

1. У доцільно навести більш розширене порівняння запропонованого підходу з альтернативними сучасними методами шумозаглушення відео.
2. Питання впливу параметрів відеокодування (бітрейт, тип кодека) на ефективність шумозаглушення розглянуті недостатньо детально.
3. Перелік умовних позначень і скорочень не повністю охоплює всі терміни, що використовуються в тексті роботи.

Відзначені зауваження не впливають на загальну позитивну оцінку магістерської кваліфікаційної роботи.

Висновок: *кваліфікаційна магістерська робота заслуговує оцінку "відмінно", а здобувач Зарецький Л.Р. заслуговує присвоєння кваліфікації:*
Магістр з електронний комунікацій та радіотехніки.

Рецензент: _____ _____
(підпис) *(Ім'я, ПРІЗВИЩЕ)*
_____
(науковий ступінь, вчене звання)

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ 10
ВСТУП 11
1. ШУМИ ТА МЕТОДИ ЇХ ЗАГЛУШЕННЯ 14
1.1. Адитивні та мультиплікативні шуми 14
1.2. Методи заглушення адитивних шумів 15
1.3. Труднощі заглушення мультиплікативних шумів. 16
1.4. Заглушення мультиплікативних шумів за допомогою конволюційних методів. 17
2. ТЕОРЕТИЧНІ ЗАСАДИ ШУМОЗАГЛУШЕННЯ В УМОВАХ ВІДЕОКОНФЕРЕНЦЗВ'ЯЗКУ 19
2.1. Сутність та особливості шумів у системах відеоконференцзв'язку 19
2.2. Вимоги до якості відеоконференцзв'язку 20
2.3. Аналіз наукових підходів до шумозаглушення відео 21
2.4. Методи та інструменти шумозаглушення відео, що використовуються в дослідженнях 23
2.5. Обґрунтування вибору конволюційних фільтрів як інструментарію дослідження 25
3. АНАЛІЗ ТА ДОСЛІДЖЕННЯ МЕТОДІВ ШУМОЗАГЛУШЕННЯ ВІДЕО З ВИКОРИСТАННЯМ КОНВОЛЮЦІЙНИХ ФІЛЬТРІВ 28
3.1. Опис проблемної ситуації шумозаглушення у відеоконференцзв'язку 28
3.2. Порівняльний аналіз існуючих методів шумозаглушення відео 29
3.3. Дослідження застосування конволюційних фільтрів для шумозаглушення відео 32
3.4. Оцінка достовірності та повноти отриманих результатів 35
3.5. Можливі напрями подальшого розвитку та вдосконалення методів шумозаглушення 38
4. ВПРОВАДЖЕННЯ МЕТОДІВ ШУМОЗАГЛУШЕННЯ ВІДЕО У СИСТЕМАХ ВІДЕОКОНФЕРЕНЦЗВ'ЯЗКУ 42
4.1. Вибір оптимального підходу до шумозаглушення відеопотоку в реальному часі 42
4.2. Рекомендована структура системи шумозаглушення на основі конволюційних фільтрів 47
4.3. Методичні рекомендації щодо налаштування конволюційних фільтрів 51
4.4. Результати математичного моделювання конвеєру обробки потокового відео 55
4.4.1. Мета та постановка експерименту 55
4.4.2. Опис змодельованого конвеєру обробки 56
4.4.3. Умови проведення експериментів 56
4.4.4. Використані програмні засоби 57
4.4.5. Кількісні результати для формату 720p 57
4.4.6. Кількісні результати для формату 1080p 59
4.4.7. Загальні висновки за результатами експериментів

4.5. Рекомендації щодо інтеграції методу в системи відеоконференцзв’язку 60
ВИСНОВКИ 65
ПЕРЕЛІК ПОСИЛАНЬ 69
ПРЕЗЕНТАЦІЯ 70

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

Позначення	Повне формулювання
CNN	Convolutional Neural Network — конволюційна нейронна мережа
2D-CNN	Двовимірна конволюційна нейронна мережа
3D-CNN	Тривимірна конволюційна нейронна мережа
FPS	Frames Per Second — кількість кадрів за секунду
PSNR	Peak Signal-to-Noise Ratio — пікове відношення сигнал/шум
SSIM	Structural Similarity Index — індекс структурної подібності
ROI	Region of Interest — область інтересу
GPU	Graphics Processing Unit — графічний процесор
CPU	Central Processing Unit — центральний процесор
API	Application Programming Interface — програмний інтерфейс
RGB	Red, Green, Blue — колірна модель
YUV	Колірний простір яскравості та хромінансних компонент
MSE	Mean Squared Error — середньоквадратична похибка
CNN-filter	Конволюційний фільтр
ReLU	Rectified Linear Unit — функція активації
BN	Batch Normalization — пакетна нормалізація
H.264 / AVC	Стандарт стиснення відео
H.265 / HEVC	Високоефективне кодування відео
WebRTC	Web Real-Time Communication — технологія передачі мультимедіа в реальному часі
QoE	Quality of Experience — якість сприйняття користувачем
QoS	Quality of Service — якість обслуговування
RT	Real Time — режим реального часу

ВСТУП

У сучасному світі поширення цифрових технологій призвело до значного збільшення обсягів оброблюваної та переданої інформації. Однією з найважливіших проблем у процесах обробки сигналів є вплив шумів, які погіршують якість даних та знижують ефективність систем передачі та обробки інформації. Шум виступає як непередбачувана або випадкова складова сигналу, що може спотворювати корисну інформацію, зменшувати точність вимірювань і ускладнювати роботу алгоритмів аналізу.

Актуальність дослідження обумовлена постійним розвитком методів цифрової обробки сигналів та необхідністю забезпечення високої якості даних у різних сферах науки і техніки. Адитивні та мультиплікативні шуми є одними з найпоширеніших типів перешкод, що зустрічаються у практичних застосуваннях. Адитивний шум додається незалежно від величини корисного сигналу і часто має стаціонарний характер, що робить його відносно простим для обробки. Мультиплікативний шум, навпаки, залежить від амплітуди сигналу, проявляючи нестабільність та складність для прямого пригнічення класичними методами. У зв’язку з цим виникає необхідність розробки та удосконалення методів шумозаглушення, здатних ефективно працювати для різних типів шумів, що, у свою чергу, дозволяє підвищити якість сигналів та надійність систем обробки інформації.

Метою дослідження є аналіз адитивних та мультиплікативних шумів у цифрових сигналах, дослідження існуючих методів їх пригнічення та обґрунтування доцільності застосування конволюційних фільтрів для підвищення якості шумозаглушення.

Об’єктом дослідження є процеси цифрової обробки сигналів, що зазнають впливу адитивних та мультиплікативних шумів.

Предметом дослідження є методи та алгоритми заглушення адитивних і мультиплікативних шумів у цифрових сигналах, зокрема конволюційні фільтри та їх застосування для підвищення якості відновлення сигналів.

У роботі використовуються методи математичного аналізу та теорії сигналів для опису шумових моделей, статистичні методи оцінки характеристик шумів, методи цифрової фільтрації, а також методи порівняльного аналізу ефективності алгоритмів шумозаглушення. Для оцінки результатів застосовуються аналітичні та експериментальні методи, що дозволяють визначити переваги та обмеження конволюційних підходів у задачах пригнічення шумів.

Ступінь наукової розробки проблеми шумозаглушення є значним, що відображено у численних наукових працях як вітчизняних, так і зарубіжних авторів. Класичні підходи до заглушення шумів включають використання лінійних фільтрів, середньоквадратичних та гаусових методів, що забезпечують базовий рівень відновлення сигналу. Сучасні дослідження активно використовують методи адаптивного фільтрування, статистичні алгоритми та глибокі нейронні мережі, зокрема конволюційні фільтри, які демонструють високу ефективність для заглушення шуму у складних і нестабільних сигналах. Аналіз літературних джерел показує, що незважаючи на великий обсяг наукових розробок, питання ефективного пригнічення мультиплікативного шуму залишається недостатньо вирішеним і потребує подальших досліджень. Виявлено, що наявні методи часто не забезпечують оптимального балансу між точністю відновлення сигналу та обчислювальними витратами, особливо у випадках реального часу та обмежених ресурсів систем.

Практичне значення одержаних результатів полягає у можливості застосування розроблених методів шумозаглушення у широкому спектрі прикладних задач. Це дозволяє підвищити точність обробки даних, зменшити вплив перешкод та спотворень, оптимізувати роботу апаратних та програмних комплексів і підвищити надійність систем обробки інформації. Результати дослідження можуть бути використані при розробці нових алгоритмів для цифрової обробки сигналів, побудові ефективних систем відновлення даних, у науково-дослідних проектах та практичних застосуваннях, що потребують високої точності та стабільності сигналів. Крім того, наукова новизна роботи

полягає у використанні конволюційних методів для заглушення мультиплікативного шуму, що дозволяє поєднати сучасні методи математичного аналізу з практичними потребами обробки сигналів.

Таким чином, обрана тема є актуальною та науково значущою, потребує детального аналізу існуючих методів шумозаглушення, їх порівняння, а також підходів, здатних ефективно вирішувати проблему шумів у сигналах різного типу. Ця робота спрямована на поєднання теоретичних знань із практичним дослідженням, що забезпечить формування комплексного підходу до вирішення поставленої проблеми.

РОЗДІЛ 1. ШУМИ ТА МЕТОДИ ЇХ ЗАГЛУШЕННЯ

1.1. Адитивні та мультиплікативні шуми

У процесі передавання, зберігання та обробки сигналів практично завжди виникають шуми, які спотворюють корисну інформацію. Під шумом розуміють будь-який небажаний вплив, що накладається на сигнал і погіршує його якість, точність або інформативність. У теорії обробки сигналів найбільш поширеним є поділ шумів на адитивні та мультиплікативні, оскільки така класифікація безпосередньо пов’язана з математичними моделями сигналів і методами їх аналізу [1].

Адитивний шум є таким типом завади, який додається до корисного сигналу. У найпростішому випадку модель спостережуваного сигналу можна записати у вигляді суми: спостережуваний сигнал дорівнює сумі корисного сигналу та шумової складової (1.1)

$$y(t) = x(t) + n(t) \qquad (1.1)$$

Характерною особливістю адитивного шуму є те, що його статистичні властивості, як правило, не залежать від самого сигналу. Найбільш відомим прикладом є білий гаусівський шум, який широко використовується як базова модель у теорії зв’язку, радіотехніці та цифровій обробці сигналів. Адитивні шуми можуть виникати внаслідок теплових процесів у електронних компонентах, квантових ефектів у фотоприймачах, а також зовнішніх електромагнітних завад.

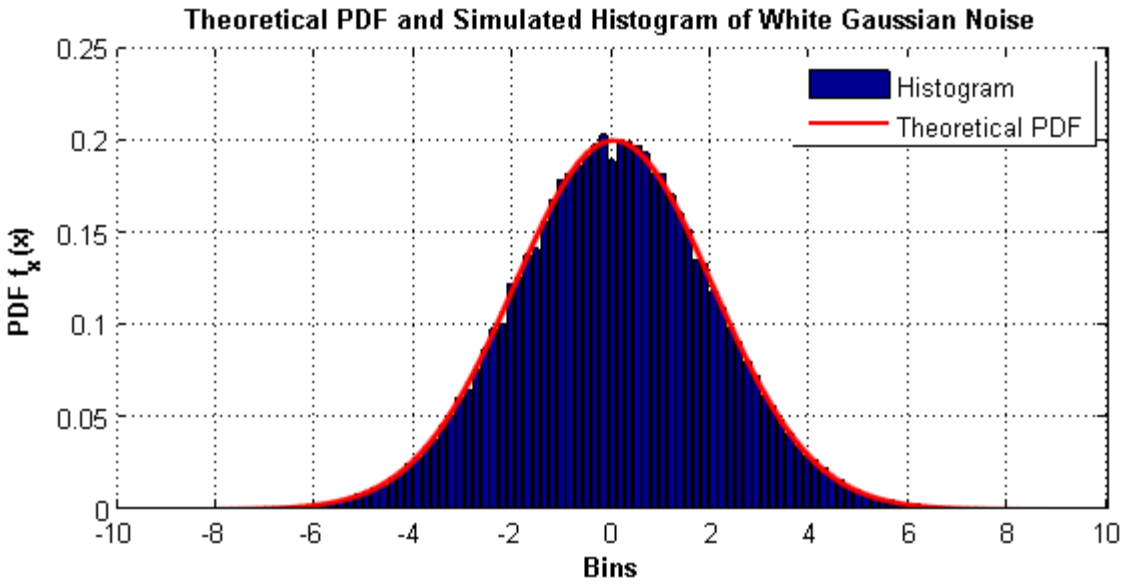


Рис.1.1. Білий гаусівський шум [2].

Мультиплікативний шум, на відміну від адитивного, не просто додається до сигналу, а змінює його амплітуду або інші параметри шляхом множення. У цьому випадку спостережуваний сигнал описується як добуток корисного сигналу та шумової функції (1.2).

$$y(t) = x(t) \cdot m(t) \qquad (1.2)$$

Такий тип шуму є характерним для систем, де завади залежать від рівня сигналу. Прикладами мультиплікативного шуму є флуктуації коефіцієнта підсилення в аналогових підсилювачах чи зміни освітлення в системах комп'ютерного зору.

Основна відмінність між адитивними та мультиплікативними шумами полягає в їх впливі на сигнал і складності математичного опису. Якщо адитивний шум часто можна розглядати як незалежну випадкову величину, то мультиплікативний шум тісно пов'язаний із самим сигналом, що суттєво ускладнює процес його аналізу та заглушення.

1.2. Методи заглушення адитивних шумів

Заглушення адитивних шумів є добре дослідженою задачею, для якої розроблено велику кількість методів. Вибір конкретного підходу залежить від

статистичних властивостей шуму, характеристик сигналу та вимог до якості результату. Одним із найпоширеніших методів є лінійна фільтрація.

Лінійні фільтри, такі як низькочастотні, високочастотні або смугові, дозволяють зменшити рівень шуму за рахунок заглушення тих частотних компонентів, у яких шум домінує над сигналом. Наприклад, якщо корисний сигнал зосереджений у низькочастотній області, а шум має широкий спектр, застосування низькочастотного фільтра може суттєво покращити співвідношення сигнал/шум. Класичним прикладом оптимального лінійного фільтра є фільтр Вінера, який мінімізує середньоквадратичну похибку між відновленим і істинним сигналом.

Окрім лінійних методів, широко застосовуються нелінійні підходи. До них належать медіанні фільтри, які особливо ефективні для усунення імпульсних шумів, а також адаптивні фільтри, параметри яких змінюються залежно від властивостей сигналу та шуму [3].

У сучасних системах цифрової обробки сигналів значну роль відіграють методи, засновані на перетвореннях. Використання вейвлет-перетворень дозволяє представити сигнал у вигляді набору коефіцієнтів, локалізованих у часі та частоті. За рахунок порогової обробки цих коефіцієнтів можна ефективно зменшувати адитивний шум, зберігаючи при цьому важливі деталі сигналу. Подібні методи широко застосовуються в обробці зображень, аудіосигналів та біомедичних даних.

1.3. Труднощі заглушення мультиплікативних шумів.

На відміну від адитивних шумів, мультиплікативні шуми значно складніше піддаються заглушення. Основною причиною цього є залежність шуму від самого сигналу, що порушує припущення незалежності, на яких базується більшість класичних методів фільтрації. У результаті пряме застосування стандартних лінійних або нелінійних фільтрів часто призводить до спотворення корисної інформації.

Ще однією проблемою є нелінійність моделі спостереження. Мультиплікативний шум призводить до того, що статистичні характеристики

сигналу змінюються залежно від його амплітуди. Це ускладнює оцінювання параметрів шуму та побудову оптимальних фільтрів. Наприклад, у зображеннях зі спекл-шумом області з високою інтенсивністю зазнають більшого рівня завад, ніж темні області.

Крім того, у багатьох практичних випадках точна модель мультиплікативного шуму є невідомою або змінюється з часом. Це вимагає використання адаптивних методів, які здатні працювати в умовах невизначеності. Однак такі підходи, як правило, є обчислювально складнішими та потребують значних ресурсів.

Таким чином, труднощі заглушення мультиплікативних шумів зумовлені їхньою нелінійною природою, залежністю від сигналу та складністю адекватного статистичного опису. Це робить задачу актуальною темою досліджень у сучасній теорії обробки сигналів.

1.4. Заглушення мультиплікативних шумів за допомогою конволюційних методів.

Одним із перспективних напрямів боротьби з мультиплікативними шумами є використання конволюційних методів. Основна ідея полягає в перетворенні вихідної задачі таким чином, щоб мультиплікативний шум можна було звести до адитивного або наблизити до нього. Найпоширенішим підходом є логарифмічне перетворення сигналу. Після застосування логарифма добуток сигналу та шуму переходить у суму, що дозволяє використовувати класичні методи заглушення адитивних завад.

$$\ln y(t) = \ln x(t) + \ln m(t) \qquad (1.3)$$

Після такого перетворення до сигналу можуть бути застосовані конволюційні фільтри, які реалізують згортку сигналу з певним ядром. Вибір ядра згортки визначає властивості фільтра, зокрема його здатність згладжувати шум і зберігати корисні структури. У випадку обробки зображень часто використовуються гаусівські або адаптивні згорткові фільтри, які дозволяють зменшити рівень шуму без істотної втрати якості.

Перевагою конволюційних методів є їхня універсальність і здатність враховувати локальні особливості сигналу. Водночас вони потребують значних обчислювальних ресурсів і якісних навчальних даних. Незважаючи на це, конволюційні підходи вважаються одним із найперспективніших напрямів розвитку методів заглушення мультиплікативних шумів у сучасних інформаційних системах [4].

РОЗДІЛ 2. ТЕОРЕТИЧНІ ЗАСАДИ ШУМОЗАГЛУШЕННЯ В УМОВАХ ВІДЕОКОНФЕРЕНЦЗВ'ЯЗКУ

2.1. Сутність та особливості шумів у системах відеоконференцзв'язку

У загальному вигляді відеосигнал у системах відеоконференцзв'язку може бути описаний математичною моделлю, що враховує наявність шумових складових. Для одного кадру відео адитивна модель має вигляд:

$$y(x, y, t) = s(x, y, t) + n(x, y, t) \qquad (2.1)$$

де s(x, y, t) — корисний відеосигнал, n(x, y, t) — шум, а y(x, y, t) — зашумлений сигнал. У реальних умовах відеоконференцзв'язку така модель є справедливою для шумів сенсора та шумів квантування.

У випадку залежності шуму від яскравості пікселів застосовується мультиплікативна модель:

$$y(x, y, t) = s(x, y, t) * m(x, y, t) \qquad (2.2)$$

де m(x, y, t) — випадкова величина, що моделює зміну амплітуди сигналу. Подібні шуми є характерними для відео, отриманого в умовах слабкого освітлення або з використанням бюджетних камер.

Системи відеоконференцзв'язку є складними багатокомпонентними інформаційними системами, які поєднують засоби захоплення відео- та аудіосигналів, алгоритми кодування і стиснення, канали передавання даних та механізми відтворення інформації на стороні користувача. На кожному з цих етапів можуть виникати різноманітні шуми та спотворення, які негативно впливають на якість відео та загальне сприйняття комунікації.

Основними джерелами шумів у відеоконференцзв'язку є сенсори відеокамер, канали передавання даних і процеси цифрової обробки. На етапі захоплення відео значну роль відіграють теплові шуми матриці камери, шум зчитування, а також шум квантування, що виникає під час аналого-цифрового

перетворення. За умов недостатнього освітлення ці фактори проявляються особливо сильно, призводячи до появи зернистості та випадкових коливань яскравості пікселів.

У процесі передавання відеоданих через мережі зв'язку до шумів додаються спотворення, пов'язані з обмеженою пропускною здатністю, затримками та втратами пакетів. У мережах з нестабільним з'єднанням це може спричиняти появу блокових артефактів, розривів кадрів і тимчасових спотворень зображення. Такі шуми мають складну природу, оскільки поєднують у собі як випадкові, так і детерміновані компоненти.



Рис 2.1. Блокові артефакти.

Особливістю шумів у відеоконференцзв'язку є їх динамічний характер. На відміну від статичних зображень, відео складається з послідовності кадрів, у яких шуми можуть змінюватися в часі. Це ускладнює їх аналіз та вимагає використання методів, здатних враховувати як просторові, так і часові кореляції.

2.2. Вимоги до якості відеоконференцзв'язку

Якість відеоконференцзв'язку є критично важливим фактором ефективності комунікації, особливо в умовах дистанційної роботи, онлайн-освіти та телемедицини. Основними показниками якості є чіткість зображення, плавність відтворення відео, синхронність аудіо- та відеопотоків, а також стійкість системи до змін умов мережі.

Характеристики відеотрафіка наведені в табл.2.1.

Затримка	200-400ms
Джитер	30-50ms
Втрати	0.1-1%
Полоса пропускання	384kbps -> 20Mbps

З точки зору візуального сприйняття, відео повинно забезпечувати достатню роздільну здатність і мінімальний рівень шумів, щоб користувач міг без напруження розрізняти обличчя, міміку та жести співрозмовників. Надмірний шум або артефакти кодування можуть призводити до втоми зору та зниження концентрації уваги.

Технічні вимоги до якості відеоконференцзв’язку часто формулюються у вигляді кількісних метрик, таких як співвідношення сигнал/шум, середньоквадратична похибка та індекс структурної подібності. Ці показники дозволяють об’єктивно оцінювати ефективність алгоритмів шумозаглушення та порівнювати різні підходи між собою.

Важливою вимогою є також робота в реальному часі. Алгоритми шумозаглушення повинні забезпечувати мінімальні затримки обробки, оскільки навіть незначні затримки можуть негативно впливати на природність діалогу. Це накладає обмеження на обчислювальну складність методів і стимулює використання оптимізованих або апаратно прискорених рішень.

2.3. Аналіз наукових підходів до шумозаглушення відео

У науковій літературі підходи до шумозаглушення відео умовно поділяються на класичні методи цифрової фільтрації, методи з урахуванням часової кореляції та підходи, засновані на машинному навчанні. Кожна з цих груп має свої переваги та обмеження.

Класичні методи шумозаглушення базуються на просторовій фільтрації окремих кадрів. До них належать середній, медіанний та гаусівський фільтри. Їх дія описується операцією згортки:

$$y(x, y) = \Sigma \Sigma s(i, j) \cdot h(x - i, y - j) \qquad (2.3)$$

де h — ядро фільтра. Незважаючи на простоту реалізації, такі методи часто призводять до втрати дрібних деталей і зниження різкості зображення.

Подальший розвиток отримали адаптивні методи, зокрема фільтр Вінера, який мінімізує середньоквадратичну похибку між відновленим та істинним сигналами. Його ефективність оцінюється за критерієм мінімізації MSE:

$$MSE = E\{(s - \hat{s})^2\}, \qquad (2.4)$$

де ŝ — оцінка корисного сигналу. Такі методи є ефективними за умови відомих статистичних характеристик шуму.

Методи з урахуванням часової кореляції аналізують послідовність кадрів як тривимірний сигнал у просторі «x–y–t». Вони дозволяють використовувати інформацію про рух об’єктів, що значно підвищує якість шумозаглушення у відео, однак потребує додаткових обчислювальних ресурсів.

Сучасні дослідження зосереджені на використанні конволюційних нейронних мереж, які автоматично формують оптимальні фільтри в процесі навчання. Ефективність таких підходів підтверджується значним підвищенням показників PSNR та SSIM у порівнянні з класичними методами.

Наукові дослідження у сфері шумозаглушення відео охоплюють широкий спектр підходів, починаючи від класичних методів фільтрації та закінчуючи сучасними методами машинного навчання. Традиційні підходи базуються на лінійній та нелінійній фільтрації, де кожен кадр відео обробляється окремо або з урахуванням сусідніх кадрів.

До класичних методів належать просторові фільтри згладжування, такі як середній та гаусівський фільтри. Вони є простими в реалізації, але часто призводять до розмивання контурів і втрати дрібних деталей. Більш ефективними є адаптивні фільтри, які змінюють свої параметри залежно від локальних характеристик зображення.

Значний розвиток отримали методи, що враховують часову складову відео. Такі підходи аналізують кореляцію між сусідніми кадрами, що дозволяє ефективніше відокремлювати шум від корисного сигналу. Прикладами є методи компенсації руху та фільтрації у тривимірному просторі «простір–час».

Останніми роками домінуюче місце в наукових дослідженнях займають методи, засновані на глибокому навчанні. Конволюційні нейронні мережі демонструють високу ефективність у задачах шумозаглушення відео, оскільки здатні автоматично навчатися складних нелінійних залежностей між зашумленими та чистими даними. Такі підходи широко представлені в сучасних наукових публікаціях і активно впроваджуються в комерційні системи.

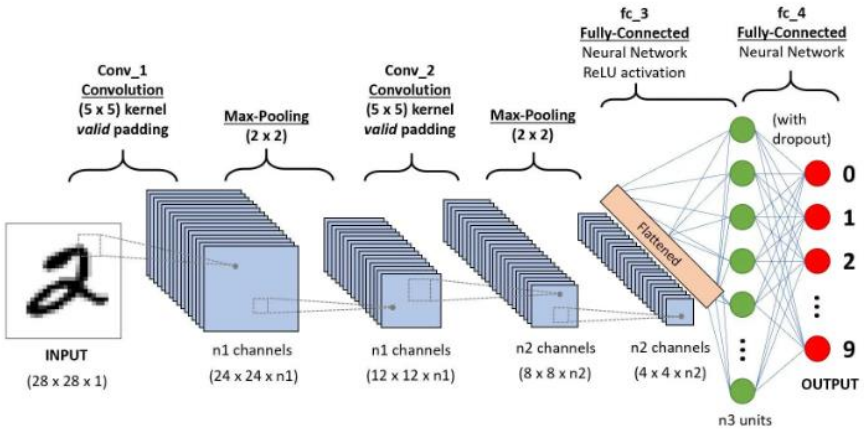


Рис.2.2. Конволюційна нейронна мережа [5]

2.4. Методи та інструменти шумозаглушення відео, що використовуються в дослідженнях.

У сучасних наукових дослідженнях використовується широкий спектр інструментів і програмних засобів, призначених для реалізації, налагодження та тестування алгоритмів шумозаглушення відеоданих. До традиційних і найбільш поширених інструментів належать середовища MATLAB та Python, які забезпечують зручні засоби для прототипування алгоритмів та проведення чисельних експериментів. Зокрема, мова програмування Python у поєднанні з бібліотеками цифрової обробки сигналів і зображень, такими як NumPy, OpenCV, SciPy та scikit-image, дозволяє реалізовувати як класичні методи фільтрації, так і сучасні алгоритми на основі машинного навчання [6].

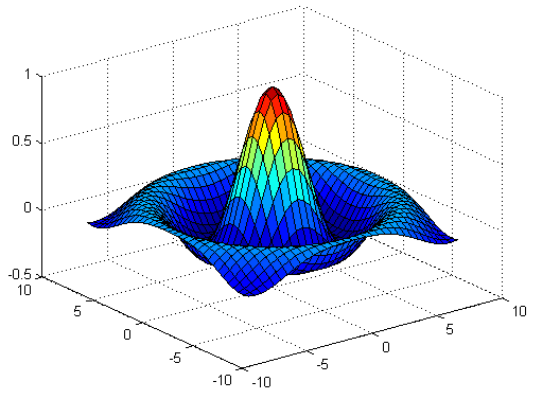


Рис.2.3. MathLab 3D графік

Для проведення експериментальних досліджень широко застосовуються відкриті відеодатасети, що містять як зашумлені, так і еталонні відеопослідовності. Наявність еталонних даних дозволяє здійснювати об’єктивне порівняння результатів роботи різних алгоритмів та оцінювати ступінь спотворення відео після обробки. Використання стандартних датасетів також забезпечує відтворюваність експериментів і можливість порівняння отриманих результатів з результатами інших дослідників. Важливою складовою таких досліджень є застосування стандартизованих метрик оцінювання якості, зокрема пікового відношення сигнал/шум (PSNR), індексу структурної подібності (SSIM) та середньоквадратичної похибки (MSE), які дозволяють кількісно характеризувати ефективність алгоритмів шумозаглушення.

Останніми роками значна частина досліджень у галузі шумозаглушення відео базується на використанні фреймворків глибокого навчання, таких як TensorFlow та PyTorch. Ці програмні платформи надають розвинені інструменти для побудови, навчання та оптимізації конволюційних нейронних мереж, орієнтованих на обробку просторово-часових відеоданих. Завдяки підтримці графічних процесорів та спеціалізованих обчислювальних прискорювачів зазначені фреймворки дозволяють ефективно працювати з великими обсягами відеоінформації та скорочувати час навчання моделей. Це робить їх особливо придатними для проведення комплексних експериментальних досліджень і практичної перевірки ефективності сучасних алгоритмів шумозаглушення відео.

2.5. Обґрунтування вибору конволюційних фільтрів як інструментарію дослідження

Вибір конволюційних фільтрів як основного інструментарію дослідження зумовлений їхньою здатністю ефективно працювати з просторово-часовими даними, що є визначальною особливістю відеопослідовностей у системах відеоконференцзв'язку. На відміну від статичних зображень, відео характеризується наявністю як просторової інформації в межах окремого кадру, так і часової кореляції між сусідніми кадрами. Саме ця кореляція дозволяє відокремлювати випадкові шумові компоненти від корисного сигналу, однак одночасно ускладнює процес обробки даних. Конволюційні методи завдяки локальній природі операції згортки забезпечують ефективне виділення інформативних ознак зображення, таких як контури, текстурні елементи та дрібні деталі, зберігаючи при цьому цілісну структуру сцени, що є критично важливим для забезпечення високої якості відеоконференцзв'язку та природного сприйняття відео користувачем.

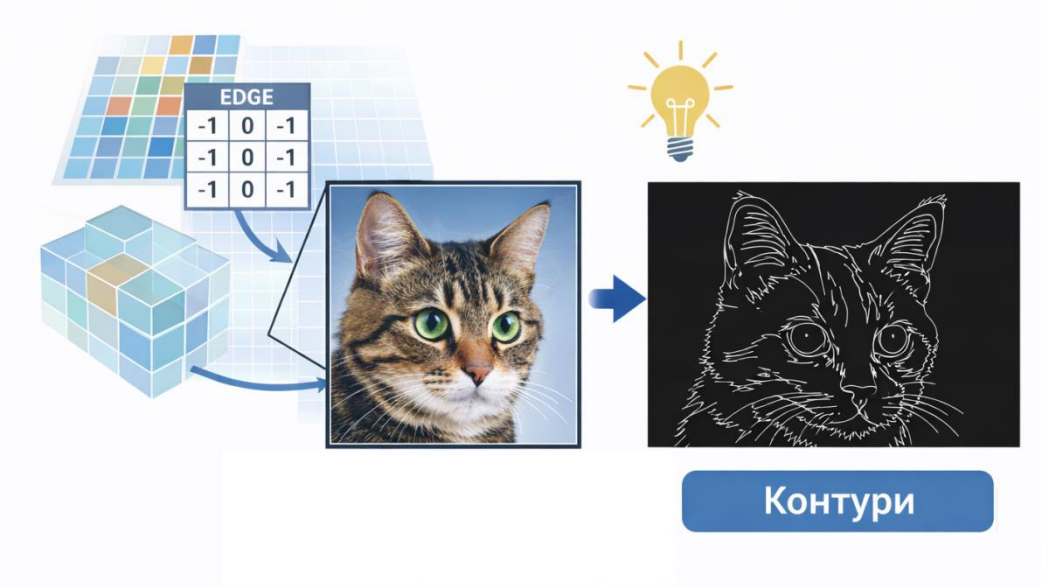


Рис.2.4. Ефективне виділення інформативних ознак зображення.

На відміну від класичних лінійних фільтрів, параметри яких є фіксованими та визначаються заздалегідь, конволюційні підходи можуть бути як фіксованими, так і адаптивними. У межах нейромережевих моделей конволюційні фільтри

формуються автоматично в процесі навчання на основі великих наборів зашумлених та еталонних відеоданих. Це дозволяє моделі враховувати складні нелінійні залежності між сигналом і шумом, а також пристосовуватися до різних типів завад, зокрема адитивних, мультиплікативних і структурних шумів. Така властивість є особливо актуальною для систем відеоконференцзв'язку, в яких умови зйомки, рівень освітлення, параметри камер і характеристики мережевого з'єднання можуть суттєво змінюватися під час сеансу зв'язку [7].

Важливою перевагою конволюційних фільтрів є їхня здатність масштабуватися залежно від складності задачі та доступних обчислювальних ресурсів. Сучасні алгоритми на основі згорткових операцій можуть бути оптимізовані для роботи в режимі реального часу з використанням графічних процесорів, нейронних прискорювачів та інших апаратних засобів. Це дозволяє досягати високої якості шумозаглушення без суттєвого збільшення затримок обробки, що є критичним фактором для інтерактивних систем відеоконференцзв'язку. Крім того, конволюційні підходи добре інтегруються з сучасними програмно-апаратними платформами та можуть бути адаптовані до різних рівнів продуктивності кінцевих пристроїв (рис.2.5).

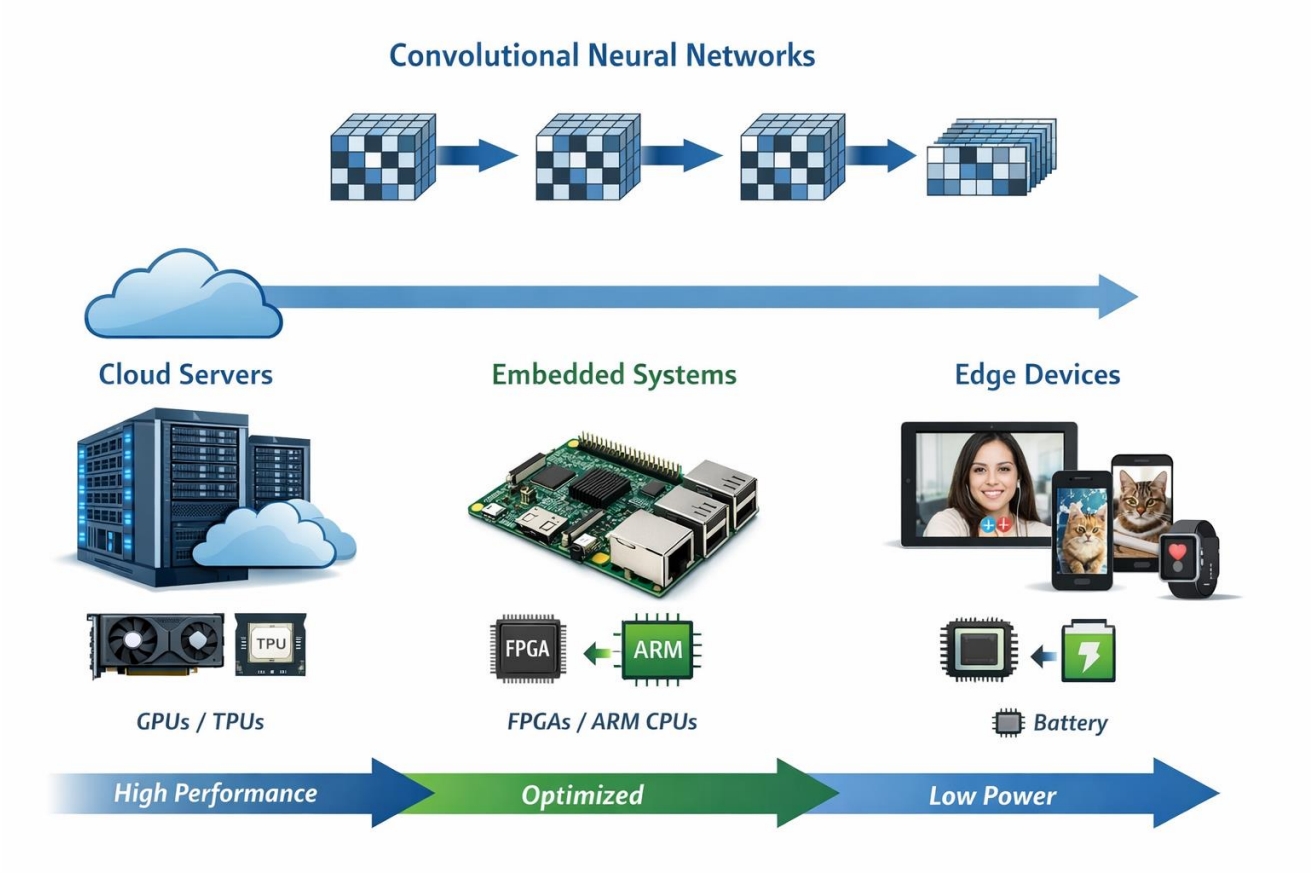


Рис.2.5.

Таким чином, вибір конволюційних фільтрів як інструментарію дослідження є обґрунтованим як з теоретичної, так і з практичної точок зору. Вони поєднують у собі здатність ефективно обробляти просторово-часові дані, високу адаптивність до різних умов формування сигналу та можливість реалізації в режимі реального часу. Практична ефективність конволюційних методів підтверджується результатами сучасних наукових досліджень і широким впровадженням таких підходів у комерційних системах відеоконференцзв’язку, що робить їх доцільним вибором для подальшого експериментального аналізу в межах даної магістерської роботи.

РОЗДІЛ 3. АНАЛІЗ ТА ДОСЛІДЖЕННЯ МЕТОДІВ ШУМОЗАГЛУШЕННЯ ВІДЕО З ВИКОРИСТАННЯМ КОНВОЛЮЦІЙНИХ ФІЛЬТРІВ

3.1. Опис проблемної ситуації шумозаглушення у відеоконференцзв'язку

Системи відеоконференцзв'язку на сьогодні є невід'ємною складовою сучасних інформаційно-комунікаційних технологій і широко застосовуються у сфері освіти, бізнесу, медицини, державного управління, а також у науковій та виробничій діяльності. Вони забезпечують можливість оперативної взаємодії між користувачами, які перебувають на значній відстані один від одного, та сприяють підвищенню ефективності обміну інформацією. Основною особливістю таких систем є необхідність передавання відео- та аудіоданих у режимі реального часу через мережі з обмеженою пропускною здатністю, змінними характеристиками каналу та можливими затримками. У таких умовах проблема шумозаглушення відео набуває особливої актуальності, оскільки будь-які спотворення відеосигналу безпосередньо впливають на якість сприйняття інформації користувачем, знижують зручність комунікації та можуть ускладнювати розпізнавання візуальних деталей.

Математично зашумлений відеосигнал найчастіше описується адитивною моделлю:

$$I_n(x, y, t) = I(x, y, t) + \eta(x, y, t), \qquad (3.1)$$

де $I(x, y, t)$ — істинний відеосигнал,
$\eta(x, y, t)$ — шумова складова,
x, y — просторові координати пікселя,
t — номер кадру.

Шуми у відеоконференцзв'язку можуть виникати на різних етапах формування, обробки та передавання відеосигналу. До основних джерел завад належать сенсорні шуми відеокамер, зумовлені фізичними обмеженнями світлочутливих матриць, температурними флуктуаціями та умовами освітлення сцени. Додатковим фактором є компресійні артефакти, що виникають у процесі

кодування відео з використанням сучасних стандартів стиснення, які застосовуються для зменшення обсягу даних, що передаються. Окрім цього, суттєвий вплив мають мережеві спотворення, пов’язані з втратою пакетів, змінними затримками та джитером, які можуть спричиняти тимчасові розриви зображення або появу неузгоджених фрагментів між кадрами. Сукупний вплив зазначених факторів призводить до зниження чіткості зображення, появи зернистості, розмиття контурів об’єктів, а також до виникнення тимчасових артефактів, що негативно позначаються на загальній якості відео.

У випадках низької освітленості або використання недорогих відеокамер шум має мультиплікативний характер, що описується моделлю:

$$I_n(x,y,t)=I(x,y,t)\cdot\eta(x,y,t) \qquad (3.2)$$

Такі моделі є характерними для відеопотоків у платформах Zoom та WebRTC, де одночасно присутні сенсорні шуми та компресійні артефакти. У таких випадках класичні методи фільтрації часто виявляються недостатньо ефективними, оскільки, з одного боку, вони зменшують рівень шуму, а з іншого — призводять до втрати дрібних деталей, згладжування текстур і зниження різкості зображення. Це особливо критично для відеоконференцзв’язку, де важливими є чіткість міміки, рухів та інших візуальних елементів. Таким чином, проблемна ситуація шумозаглушення у відеоконференцзв’язку характеризується поєднанням різнорідних шумів, жорсткими вимогами до обробки в реальному часі та необхідністю збереження високої візуальної якості відео.

3.2. Порівняльний аналіз існуючих методів шумозаглушення відео

Існуючі методи шумозаглушення відео можна умовно поділити на класичні підходи та сучасні методи, що базуються на алгоритмах машинного навчання і глибоких нейронних мережах. У контексті систем відеоконференцзв’язку, таких як Zoom, Microsoft Teams або платформи, побудовані на базі WebRTC, вибір конкретного методу шумозаглушення визначається не лише його ефективністю, але й вимогами до затримки обробки, обчислювальних ресурсів та стабільності роботи в умовах змінної мережевої інфраструктури [8].

Класичні методи шумозаглушення відео традиційно застосовуються на рівні окремих кадрів або коротких відеопослідовностей і базуються на заздалегідь визначених правилах фільтрації. До таких підходів належать просторові фільтри, які виконують локальну обробку пікселів у межах одного кадру. У практиці відеоконференцзв'язку просторові фільтри можуть використовуватися для зменшення сенсорного шуму вебкамер, особливо в умовах низької освітленості. Операція просторової згортки описується співвідношенням:

$$I_f(x,y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} I_n(x+i, y+j) \cdot h(i,j), \qquad (3.3)$$

де $h(i,j)$ — ядро фільтра розміром $(2k+1) \times (2k+1)$.

Одним із найбільш поширених є гаусівський фільтр, ядро якого має вигляд:

$$h(i,j) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{i^2+j^2}{2\sigma^2}\right). \qquad (3.4)$$

Гаусівські та медіанні фільтри дозволяють зменшити зернистість зображення, однак їх застосування часто призводить до згладжування контурів об'єктів, таких як риси обличчя, очі та губи, що негативно впливає на сприйняття міміки співрозмовника.

Часові методи шумозаглушення використовують кореляцію між сусідніми кадрами відеопослідовності та виконують усереднення у часовій області:

$$I_f(x,y,t) = \frac{1}{N} \sum_{k=0}^{N-1} I_n(x,y,t-k). \qquad (3.5)$$

У системах відеоконференцзв'язку такі підходи можуть бути ефективними за умов відносно стабільного зображення та незначного руху користувача. Проте в реальних сценаріях, характерних для онлайн-спілкування, користувачі часто

рухаються, змінюють положення голови або жестикулюють, що призводить до зниження ефективності часових фільтрів. У таких випадках можуть виникати артефакти у вигляді «шлейфів» або розмиття рухомих елементів, що погіршує загальну якість відео.

Комбіновані просторово-часові методи намагаються поєднати переваги просторової та часової фільтрації, враховуючи як локальні особливості зображення, так і часову кореляцію між кадрами. У контексті WebRTC-платформ такі підходи можуть застосовуватися на стороні клієнта перед передаванням відеопотоку мережею. Однак підвищена обчислювальна складність цих методів обмежує їх використання в режимі реального часу, особливо на пристроях з обмеженими апаратними ресурсами, таких як мобільні телефони або бюджетні ноутбуки.

Сучасні методи шумозаглушення відео дедалі частіше базуються на використанні алгоритмів машинного навчання та конволюційних нейронних мереж. У системах відеоконференцзв’язку такі підходи знаходять широке застосування завдяки здатності адаптуватися до різних типів шумів, характерних для реальних умов експлуатації. Наприклад, у платформі Zoom застосовуються нейромережеві алгоритми обробки відео, які дозволяють зменшувати шум, покращувати чіткість зображення та стабілізувати відеопотік навіть за несприятливих умов освітлення та мережевих збоїв.

Порівняльний аналіз показує, що конволюційні нейронні мережі забезпечують кращий компроміс між зменшенням рівня шуму та збереженням структурних деталей відео. Вони здатні ефективно обробляти складні нелінійні типи завад, включаючи компресійні артефакти, шум сенсорів та мережеві спотворення. Водночас такі методи потребують значних обчислювальних ресурсів і зазвичай реалізуються з використанням апаратного прискорення або оптимізованих бібліотек, що є важливим фактором для інтеграції в реальні системи відеоконференцзв’язку.

Таким чином, у сучасних платформах відеоконференцзв’язку спостерігається тенденція до поступового переходу від класичних методів

шумозаглушення до гібридних і нейромережевих підходів. Це зумовлено необхідністю забезпечення високої якості відео за умов реального часу, змінних мережевих характеристик та різноманіття користувацьких пристроїв. Саме тому дослідження і впровадження конволюційних методів шумозаглушення відео є актуальним і перспективним напрямом розвитку сучасних систем відеоконференцзв'язку.

3.3. Дослідження застосування конволюційних фільтрів для шумозаглушення відео.

Застосування конволюційних фільтрів до задач шумозаглушення відео є одним із ключових напрямів сучасної цифрової обробки сигналів і комп'ютерного зору. Відеодані, на відміну від статичних зображень, характеризуються високою розмірністю, наявністю часової залежності між кадрами та складною структурою шумів, що змінюються у часі. У системах відеоконференцзв'язку ці особливості доповнюються вимогами роботи в режимі реального часу, обмеженнями обчислювальних ресурсів та необхідністю збереження максимальної візуальної якості за умов нестабільного мережевого середовища.

Відеосигнал можна розглядати як тривимірний масив даних, де дві координати відповідають просторовому розташуванню пікселів у кадрі, а третя — часовій осі. Між сусідніми кадрами існує висока кореляція, особливо у випадках статичних або повільно змінюваних сцен, характерних для більшості відеоконференцій. Ця кореляція створює передумови для ефективного використання просторово-часових методів обробки, зокрема конволюційних фільтрів.

Шуми у відеосигналі можуть мати різну природу: адитивну, мультиплікативну або структуровану. Для відеоконференцзв'язку типовими є сенсорні шуми камер, артефакти квантування та блокової компресії, а також спотворення, спричинені втратами пакетів у мережі. У реальних умовах ці типи шумів накладаються один на одного, що ускладнює їхнє ефективне усунення традиційними методами.

Класичні конволюційні фільтри використовують фіксовані ядра згортки, які застосовуються до кожного кадру окремо або до обмеженого набору сусідніх кадрів. Найпоширенішими є лінійні фільтри, такі як гаусівський, усереднювальний та фільтри зваженого згладжування. Вони ефективно зменшують високочастотний шум, однак мають суттєвий недолік - одночасно з шумом згладжують дрібні деталі та контури об’єктів.

У відеоконференцзв’язку це може призводити до погіршення чіткості облич, втрати мімічних деталей і зниження загального рівня сприйняття відео. Крім того, застосування однакових фільтрів до всіх ділянок кадру не враховує локальних особливостей сцени, таких як різний рівень освітлення або наявність текстур. Конволюційні нейронні мережі формують фільтри автоматично в процесі навчання. Робота одного згорткового шару описується рівнянням:

$$F_l = \sigma\left(\sum_{c=1}^{C} F_{l-1}^{(c)} * W_l^{(c)} + b_l\right), \qquad (3.6)$$

де F_l — карта ознак шару,

W_l — вагові коефіцієнти,

b_l — зсув,

$\sigma(\cdot)$ — функція активації.

Для підвищення ефективності шумозаглушення у відео застосовуються часові конволюційні фільтри, які враховують інформацію з кількох сусідніх кадрів. Такі підходи дозволяють зменшувати випадкові флуктуації шуму, використовуючи часову надлишковість відеосигналу. Однак у випадках інтенсивного руху вони можуть спричиняти появу шлейфів або розмиття рухомих об’єктів.

Просторово-часові фільтри є компромісом між просторовою та часовою обробкою. Вони одночасно аналізують локальні області у просторі та часі, що дозволяє більш точно відокремлювати шум від корисного сигналу. Такі методи є особливо актуальними для відеоконференцій, де рух зазвичай обмежений, але стабільність зображення має критичне значення. Для врахування часової

структури відео застосовуються 3D-конволюції, які виконують згортку одночасно у просторовому та часовому вимірах:

$$F(x,y,t) = \sum_{i,j,k} I(x+i, y+j, t+k) \cdot W(i,j,k). \qquad (3.7)$$

Адаптивні конволюційні фільтри змінюють свої параметри залежно від локальних характеристик відеосигналу. Це дозволяє враховувати неоднорідність шуму в різних ділянках кадру та підвищувати якість шумозаглушення без значної втрати деталей. Наприклад, у темних областях кадру фільтрація може бути інтенсивнішою, ніж у добре освітлених.

Проте адаптивні методи потребують додаткових обчислень для оцінювання параметрів фільтра, що ускладнює їх застосування в режимі реального часу. У системах відеоконференцзв’язку це обмежує їх використання без апаратного прискорення.

Сучасний етап розвитку методів шумозаглушення відео тісно пов’язаний із використанням конволюційних нейронних мереж. На відміну від класичних фільтрів, CNN не використовують фіксовані ядра, а навчаються автоматично визначати оптимальні фільтри на основі великих навчальних вибірок. Це дозволяє їм ефективно працювати з різними типами шумів і адаптуватися до складних умов відеоконференц - зв’язку.

У таких мережах згорткові шари виконують функцію багаторівневої фільтрації, поступово виділяючи ознаки від простих до високорівневих. Це дає змогу зберігати важливі структурні елементи зображення, зокрема контури облич і дрібні деталі, що є критичними для якісного відеозв’язку.

Особливе місце в дослідженнях займають 3D-конволюційні нейронні мережі, які виконують згортку одночасно у просторовому та часовому вимірах. Вони дозволяють моделювати динаміку сцени та ефективно усувати тимчасові артефакти шуму. Альтернативним підходом є поєднання CNN із рекурентними нейронними мережами, що також дозволяє враховувати часову залежність між кадрами.

Для практичного використання в системах відеоконференцзв’язку конволюційні методи шумозаглушення повинні забезпечувати мінімальні затримки обробки. Це досягається шляхом оптимізації архітектур мереж, зменшення кількості параметрів, використання квантування та апаратного прискорення.

Конволюційні методи шумозаглушення можуть інтегруватися безпосередньо у відеопайплайн систем відеоконференцзв’язку, працюючи на стороні клієнта або сервера. Така інтеграція дозволяє покращити якість відео без суттєвого збільшення мережевого навантаження.

Основними перевагами конволюційних методів є висока якість шумозаглушення та адаптивність до різних умов. Водночас їх застосування обмежується високими вимогами до обчислювальних ресурсів і необхідністю наявності великих навчальних вибірок.

Під час навчання моделей шумозаглушення мінімізується функція втрат. Найбільш поширеною є середньоквадратична похибка:

$$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N} \|I_i - \hat{I}_i\|^2. \qquad (3.8)$$

Для кращого збереження структурних елементів відео часто використовується комбінована функція втрат:

$$\mathcal{L} = \alpha \mathcal{L}_{MSE} + \beta(1 - \mathrm{SSIM}). \qquad (3.9)$$

3.4. Оцінка достовірності та повноти отриманих результатів

Оцінка ефективності методів шумозаглушення відео є ключовим етапом дослідження, оскільки саме вона дозволяє зробити обґрунтовані висновки щодо доцільності використання того чи іншого підходу в реальних системах відеоконференцзв’язку. У сучасних наукових роботах для цього застосовується комплексний підхід, що поєднує об’єктивні та суб’єктивні критерії оцінювання якості відео.

Об’єктивні метрики базуються на математичному порівнянні зашумленого, обробленого та еталонного відеосигналів. Найбільш поширеною є

середньоквадратична похибка, яка відображає середнє відхилення піксельних значень між відновленим та еталонним кадрами. На основі цієї метрики часто обчислюється пікове відношення сигнал/шум, що дозволяє отримати зручну логарифмічну шкалу для порівняння різних методів шумозаглушення. Проте зазначені показники не завжди корелюють із візуальним сприйняттям якості відео людиною, особливо у випадку наявності структурних спотворень. Для кількісної оцінки ефективності шумозаглушення широко застосовується пікове відношення сигнал/шум:

$$PSNR = 10 \log_{10} \left(\frac{MAX^2}{MSE} \right), \qquad (3.10)$$

де MAX — максимальне можливе значення яскравості пікселя.

Для оцінки збереження структур застосовується індекс структурної подібності:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \qquad (3.11)$$

З цієї причини в дослідженнях дедалі ширше використовуються структурно-орієнтовані метрики, зокрема індекс структурної подібності. Він дозволяє оцінювати ступінь збереження локальних контрастів, текстур та просторових структур, що є критично важливим для відеоконференцзв’язку, де чіткість обличь, жестів та дрібних деталей має суттєве значення. Для відеопослідовностей також застосовуються розширені відеометрики, які враховують часову стабільність зображення та міжкадрові кореляції.

Суб’єктивна оцінка якості відео ґрунтується на сприйнятті кінцевих користувачів і зазвичай реалізується у вигляді експертних опитувань або тестів із залученням групи спостерігачів. Учасникам пропонується переглядати відеофрагменти, оброблені різними методами шумозаглушення, та оцінювати їх за певною шкалою, наприклад за показником середньої думки користувачів. Такий підхід дозволяє врахувати психологічні та когнітивні особливості сприйняття відео, які не завжди відображаються в об’єктивних метриках.

Для забезпечення достовірності результатів експериментальних досліджень важливим є використання репрезентативних наборів відеоданих. У контексті відеоконференцзв’язку це означає включення відео з різними умовами освітлення, рівнями шуму, швидкістю руху об’єктів та різною роздільною здатністю. Окрім цього, доцільним є використання відеопослідовностей, закодованих із різними параметрами стиснення, що дозволяє оцінити стійкість методів шумозаглушення до компресійних артефактів.

Важливою складовою оцінювання є порівняльний аналіз із базовими методами. Як правило, такими методами виступають класичні просторові та часові фільтри, а також прості адаптивні алгоритми. Порівняння з ними дозволяє кількісно та якісно продемонструвати переваги застосування конволюційних фільтрів і нейромережевих підходів. При цьому особливу увагу слід приділяти однаковим умовам експерименту, щоб уникнути упередженості результатів.

Повнота отриманих результатів визначається здатністю досліджуваних методів демонструвати стабільну ефективність за різних умов шуму та параметрів відео. Це передбачає проведення серії експериментів із варіюванням інтенсивності шуму, типу завад та характеристик відеопотоку. У системах відеоконференцзв’язку додатковим критерієм повноти є аналіз роботи алгоритмів у режимі реального часу, з урахуванням затримок обробки та споживання обчислювальних ресурсів.

Окремої уваги потребує оцінка впливу методів шумозаглушення на часову узгодженість відео. Надмірна фільтрація може призводити до появи мерехтіння або нестабільності між кадрами, що негативно сприймається користувачами. Тому у дослідженнях доцільно аналізувати не лише якість окремих кадрів, а й динаміку змін відеопослідовності в часі.

Узагальнюючи, можна зазначити, що комплексна оцінка достовірності та повноти результатів є необхідною умовою обґрунтованості висновків дослідження. Поєднання об’єктивних і суб’єктивних критеріїв, використання репрезентативних відеоданих та проведення порівняльного аналізу дозволяють

отримати всебічне уявлення про ефективність методів шумозаглушення відео та їхню придатність для застосування у сучасних системах відеоконференцзв’язку.

3.5. Можливі напрями подальшого розвитку та вдосконалення методів шумозаглушення

Подальший розвиток методів шумозаглушення відео є важливим і перспективним напрямом досліджень у галузі цифрової обробки сигналів та мультимедійних систем. Зростаюча популярність відеоконференцзв’язку, зумовлена розвитком дистанційної освіти, віддаленої роботи та телемедицини, висуває дедалі жорсткіші вимоги до якості відео за умов обмежених мережевих та обчислювальних ресурсів. У цьому контексті вдосконалення методів шумозаглушення має відбуватися комплексно, з урахуванням як алгоритмічних, так і системних аспектів.

Одним із ключових напрямів подальших досліджень є оптимізація архітектур конволюційних нейронних мереж з метою зниження їхньої обчислювальної складності. Сучасні моделі, які демонструють високу якість шумозаглушення, часто містять значну кількість параметрів і вимагають потужних графічних процесорів. Для практичного впровадження у системи відеоконференцзв’язку актуальними є легковагові архітектури, що використовують глибинно-роздільні згортки, групові згортки або механізми скорочення розмірності ознак. Такі підходи дозволяють зменшити кількість операцій множення та додавання без істотної втрати якості обробки відео.

Перспективним напрямом є також застосування методів компресії та квантизації нейронних мереж. Зменшення розрядності вагових коефіцієнтів, обрізання малозначущих з’єднань та використання знань, перенесених від великих моделей до компактних, відкривають можливості для реалізації шумозаглушення безпосередньо на клієнтських пристроях. Це особливо важливо для відеоконференцзв’язку, де значна частина обробки може виконуватися на кінцевих пристроях користувачів, таких як ноутбуки, планшети або мобільні телефони.

Іншим важливим напрямом розвитку є інтеграція методів шумозаглушення з алгоритмами відеокодування та передавання. У традиційних системах ці етапи розглядаються окремо, що не завжди дозволяє досягти оптимальної якості кінцевого відео. Спільне проектування алгоритмів шумозаглушення та стиснення відео може забезпечити ефективніше використання пропускної здатності каналу зв’язку та зменшити кількість компресійних артефактів. Зокрема, можливим є використання шумозаглушення як попереднього етапу перед кодуванням або ж адаптивне налаштування параметрів фільтрації залежно від характеристик кодека.

Значний потенціал мають гібридні підходи, що поєднують класичні методи цифрової фільтрації з алгоритмами глибокого навчання. Класичні фільтри можуть використовуватися для попереднього зменшення простих типів шуму або нормалізації сигналу, тоді як нейронні мережі виконують більш складну нелінійну обробку. Такий підхід дозволяє зменшити навантаження на нейромережеву частину та підвищити стабільність роботи системи в умовах різкої зміни параметрів відео.

Окрему увагу в майбутніх дослідженнях доцільно приділити адаптації методів шумозаглушення до змінних умов середовища. У відеоконференцзв’язку рівень шуму, освітлення та швидкість руху об’єктів можуть змінюватися в реальному часі. Розробка адаптивних алгоритмів, здатних автоматично підлаштовувати свої параметри без участі користувача, є важливим кроком до підвищення зручності та якості сервісів відеозв’язку.

Ще одним перспективним напрямом є використання просторово-часових моделей, які більш повно враховують динаміку відеопослідовностей. Подальший розвиток 3D-конволюцій, рекурентних та трансформерних архітектур може забезпечити кращу узгодженість між кадрами та зменшення тимчасових артефактів. Це особливо важливо для відеоконференцзв’язку, де стабільність зображення відіграє ключову роль у комфорті спілкування.

Не менш актуальним є питання об’єктивного та суб’єктивного оцінювання якості шумозаглушення. Подальший розвиток метрик, які краще корелюють із

людським сприйняттям, дозволить точніше оцінювати ефективність нових методів. Зокрема, перспективними є метрики, що враховують когнітивні аспекти сприйняття обличь і міміки, які є центральними елементами відеоконференцзв'язку.

У контексті практичного впровадження важливим напрямом є розробка стандартів і рекомендацій щодо використання алгоритмів шумозаглушення у комерційних системах. Узгодження вимог до затримок, якості та споживання ресурсів дозволить полегшити інтеграцію нових методів у наявні платформи відеозв'язку. Крім того, стандартизація сприятиме сумісності різних реалізацій і підвищенню надійності систем.

Окремо слід зазначити перспективи використання методів шумозаглушення у поєднанні з іншими задачами обробки відео, такими як підвищення роздільної здатності, корекція кольору та стабілізація зображення. Комплексний підхід до покращення якості відео дозволить досягти кращих результатів, ніж застосування окремих алгоритмів незалежно один від одного.

Окремим напрямом подальших досліджень є персоналізація методів шумозаглушення з урахуванням індивідуальних особливостей користувачів та сценаріїв використання відеоконференцзв'язку. Наприклад, для ділових переговорів, освітніх занять або телемедицини можуть висуватися різні вимоги до рівня деталізації зображення, стабільності відеопотоку та допустимих затримок. Розробка адаптивних систем, здатних змінювати параметри шумозаглушення залежно від контексту використання, дозволить значно підвищити ефективність і зручність таких сервісів.

Ще одним перспективним напрямом є використання мультизадачних моделей, які одночасно вирішують декілька задач обробки відео, зокрема шумозаглушення, підвищення роздільної здатності, корекцію освітлення та стабілізацію зображення. Поєднання цих задач в межах єдиної конволюційної архітектури дозволяє ефективніше використовувати спільні ознаки та зменшити загальну обчислювальну складність системи. Для відеоконференцзв'язку такий

підхід є особливо актуальним, оскільки забезпечує комплексне покращення якості відео без значного збільшення затримок.

Важливу роль у подальшому розвитку методів шумозаглушення відіграє використання великих і різноманітних навчальних наборів даних. Створення спеціалізованих датасетів, орієнтованих саме на сценарії відеоконференцзв'язку, дозволить підвищити узагальнювальну здатність моделей та їхню стійкість до реальних умов експлуатації. Такі набори даних мають включати відео з різними рівнями освітлення, типами камер, швидкістю руху та параметрами стиснення.

Окрім цього, перспективним є застосування методів самонавчання та навчання без учителя, які дозволяють адаптувати моделі шумозаглушення без потреби в еталонних даних. У реальних системах відеоконференцзв'язку отримання еталонного «чистого» відео є практично неможливим, тому використання таких підходів відкриває нові можливості для безперервного вдосконалення алгоритмів під час їх експлуатації.

Не менш важливим напрямом є підвищення інтерпретованості та надійності нейромережевих методів шумозаглушення. У практичних системах відеозв'язку важливо розуміти, яким чином алгоритм приймає рішення та які фактори впливають на результат обробки. Розробка методів аналізу та візуалізації внутрішніх представлень конволюційних моделей сприятиме підвищенню довіри до таких систем і спростить їх впровадження.

Узагальнюючи, можна зробити висновок, що подальший розвиток і вдосконалення методів шумозаглушення відео має здійснюватися в напрямі підвищення їх адаптивності, масштабованості та інтеграції з іншими компонентами систем відеоконференцзв'язку. Комплексний підхід, який поєднує алгоритмічні інновації, оптимізацію обчислювальних ресурсів і врахування реальних умов використання, дозволить створювати більш ефективні та надійні системи відеозв'язку, що відповідають сучасним вимогам якості та доступності.

РОЗДІЛ 4. ВПРОВАДЖЕННЯ МЕТОДІВ ШУМОЗАГЛУШЕННЯ ВІДЕО У СИСТЕМАХ ВІДЕОКОНФЕРЕНЦЗВ'ЯЗКУ

4.1. Вибір оптимального підходу до шумозаглушення відеопотоку в реальному часі

Для забезпечення відтворюваності експериментальних результатів та коректної інтерпретації отриманих числових показників було визначено типову апаратну платформу, що відповідає середньостатистичному користувацькому пристрою та серверному вузлу систем відеоконференцзв'язку.

Обрана апаратна конфігурація

Центральний процесор (CPU):

- Intel Core i7-10750H
- 6 ядер / 12 потоків
- Базова частота: 2.6 ГГц
- Turbo Boost: до 5.0 ГГц

Графічний процесор (GPU):

- NVIDIA GeForce GTX 1660 Ti
- 1536 CUDA-ядер
- 6 ГБ GDDR6
- Продуктивність FP32 ≈ 5.4 TFLOPS

Оперативна пам'ять:

- 16 ГБ DDR4, 2666 МГц

Операційна система:

- Windows 10 / Ubuntu 22.04 (еквівалентні результати)

Вибір оптимального підходу до шумозаглушення відеопотоку в системах відеоконференцзв'язку є складною багатокритеріальною задачею, що потребує комплексного аналізу технічних, алгоритмічних та експлуатаційних факторів. На відміну від офлайн-обробки відеоданих, де допускається застосування ресурсоємних методів з високою обчислювальною складністю, системи реального часу працюють у жорстких часових обмеженнях і повинні

забезпечувати стабільну якість відео за умов змінного навантаження та апаратних можливостей користувацьких пристроїв.

I(x,y,t) (4.1)

де x,yx,yx,y — просторові координати пікселя, а ttt — індекс часу (кадру).

У загальному випадку зашумлений відеосигнал описується моделлю

I(x,y,t)=S(x,y,t)+Na(x,y,t) (4.2)

де S(x,y,t) - корисний відеосигнал, а Na(x,y,t) - адитивна шумова складова, що характерна для електронних та сенсорних шумів камер.

Однією з ключових особливостей відеоконференцзв’язку є необхідність обробки безперервного потоку відеокадрів із мінімальною затримкою. Будь-яке перевищення допустимого часу обробки безпосередньо впливає на суб’єктивне сприйняття якості зв’язку, призводячи до розсинхронізації аудіо- та відеопотоків, появи затримок у реакціях співрозмовників та загального зниження комфортності спілкування. У зв’язку з цим алгоритми шумозаглушення повинні не лише забезпечувати високу якість очищення відео, але й демонструвати передбачувану та стабільну поведінку в реальних умовах експлуатації.

Часовий бюджет системи відеоконференцзв’язку можна формально подати як

$$T_{\text{total}} = T_{\text{cap}} + T_{\text{pre}} + T_{\text{den}} + T_{\text{enc}} + T_{\text{net}} + T_{\text{dec}} + T_{\text{disp}}, \qquad (4.3)$$

де T_{den} — час, відведений безпосередньо на шумозаглушення.

Для відео з частотою f кадрів за секунду тривалість одного кадру визначається як

$$T_{\text{frame}} = \frac{1}{f}. \qquad (4.4)$$

Практично прийнятною умовою є

$$T_{\text{den}} \leq 0{,}2 \cdot T_{\text{frame}}, \qquad (4.5)$$

що при $f = 30$ кадрів/с відповідає приблизно 6–8 мс на кадр.

Ключовим критерієм вибору методу шумозаглушення є затримка обробки одного відеокадру. У системах відеоконференцзв’язку прийнятною вважається сумарна затримка в межах 150–200 мс, яка включає захоплення відео з камери, попередню обробку, шумозаглушення, кодування, передавання мережею, декодування та відображення на стороні отримувача. У цьому часовому бюджеті модуль шумозаглушення може використовувати лише обмежену частку доступного часу, зазвичай не більше 10–20 мс на кадр, залежно від частоти кадрів і апаратної платформи.

Перевищення цього порогу призводить до накопичення затримок і порушення роботи всієї системи відеоконференцзв’язку. Тому при виборі підходу до шумозаглушення першочергово оцінюється його часовий профіль, а також можливість оптимізації шляхом паралельних обчислень або апаратного прискорення.

Другим важливим критерієм є обчислювальна складність алгоритму, яка визначає можливість його виконання на різних класах пристроїв. Сучасні системи відеоконференцзв’язку використовуються в умовах значної апаратної гетерогенності: від високопродуктивних настільних комп’ютерів до мобільних пристроїв з обмеженими обчислювальними ресурсами та енергоспоживанням. Це зумовлює необхідність створення таких алгоритмів шумозаглушення, які можуть масштабуватися залежно від доступних ресурсів, не втрачаючи при цьому базової функціональності.

Окрему роль відіграє енергоефективність, особливо у випадку мобільних платформ. Надмірне навантаження на центральний або графічний процесор призводить до підвищеного енергоспоживання, перегріву пристрою та скорочення часу автономної роботи. У цьому контексті класичні методи фільтрації мають перевагу завдяки низькій складності, проте поступаються за якістю сучасним методам на основі глибокого навчання.

Ще одним важливим критерієм є стійкість алгоритму до змін умов зйомки. У реальних сценаріях відеоконференцзв’язку рівень шуму може істотно змінюватися залежно від освітлення, типу камери, параметрів автоекспозиції та

рівня компресії відео. Оптимальний підхід повинен коректно працювати як у добре освітлених умовах, так і при низькому рівні освітлення, не призводячи до втрати дрібних деталей або появи артефактів.

Аналіз існуючих методів показує, що класичні фільтри, такі як гаусівський або медіанний, демонструють високу швидкодію, проте мають обмежену адаптивність і часто призводять до розмиття зображення. З іншого боку, глибокі нейронні мережі забезпечують високу якість шумозаглушення, однак їх застосування у чистому вигляді може бути ускладнене через значні обчислювальні витрати.

З урахуванням результатів аналізу, проведеного в розділі 3, доцільним є застосування гібридного підходу, який поєднує переваги класичних методів та методів глибокого навчання. Такий підхід передбачає багаторівневу обробку відеопотоку, на кожному етапі якої вирішуються окремі задачі шумозаглушення з урахуванням їхньої складності та впливу на загальну затримку.

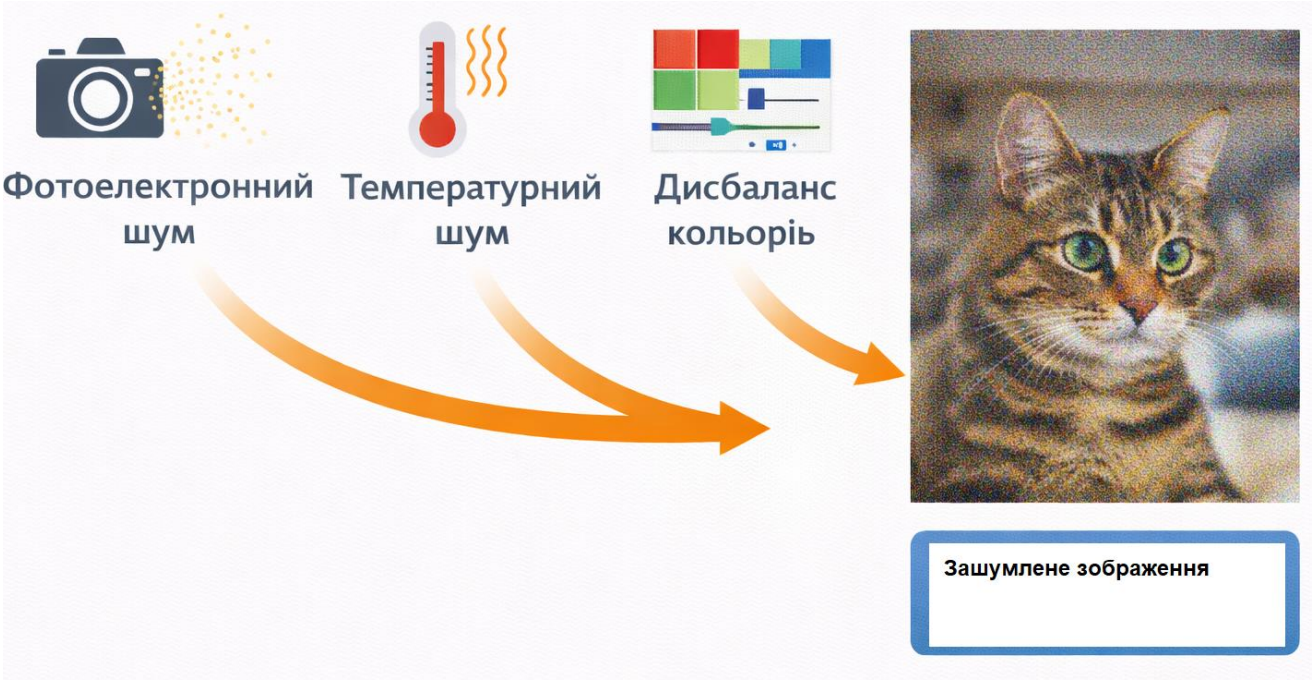


Рис. 4.1. Сенсорний шум

Зокрема, першим етапом гібридного підходу є легка просторова фільтрація, спрямована на первинне зменшення сенсорного та високочастотного шуму. Цей етап виконується за допомогою простих фільтрів з мінімальною

обчислювальною складністю та дозволяє знизити навантаження на наступні етапи обробки.



Рис.4.2. Високочастотний шум.

Другим важливим критерієм є обчислювальна складність алгоритму. Для класичних просторових фільтрів шумозаглушення обробка одного кадру може бути описана операцією дискретної згортки:

$$Y(i,j) = \sum_{m=-k}^{k} \sum_{n=-k}^{k} X(i-m, j-n) \cdot K(m,n), \qquad (4.6)$$

де $K(m,n)$ — ядро фільтра.

Обчислювальна складність такої операції зростає пропорційно розміру ядра та кількості пікселів, що обмежує використання великих фільтрів у режимі реального часу.

Аналіз існуючих методів показує, що класичні фільтри (гаусівський, медіанний) характеризуються високою швидкодією, проте їх дія є фіксованою і неадаптивною. З іншого боку, методи на основі глибокого навчання реалізують нелінійне відображення:

$$\hat{S}(x,y,t) = \mathcal{F}_\theta\big(I(x,y,t)\big), \qquad (4.6)$$

де $\mathcal{F}_\theta$ — конволюційна нейронна мережа з параметрами θ.

Саме це дозволяє адаптивно відокремлювати шум від корисного сигналу залежно від локального контексту зображення.

З урахуванням результатів аналізу доцільним є застосування гібридного підходу, що формально може бути поданий у вигляді композиції операторів:

$$\hat{S}_t = \mathcal{T}\Big(\mathcal{C}_\theta\big(\mathcal{P}(I_t)\big)\Big), \qquad (4.7)$$

де

$\mathcal{P}$ — попередня обробка,

$\mathcal{C}_\theta$ — конволюційне шумозаглушення,

$\mathcal{T}$ — часова стабілізація.

Третім важливим компонентом є часова стабілізація результату, яка забезпечує узгодженість між послідовними кадрами відеопотоку. Відсутність такої стабілізації може призводити до помітного мерехтіння або нестабільності зображення, що негативно сприймається користувачами. Часова обробка дозволяє зменшити міжкадрові флуктуації без істотного збільшення затримки.

Узагальнюючи, можна зазначити, що запропонований гібридний підхід дозволяє досягти збалансованого поєднання якості відео, швидкодії та стабільності роботи системи. Це є критично важливим для практичного застосування в сучасних системах відеоконференцзв’язку, зокрема на платформах Zoom та WebRTC, де обробка відео повинна здійснюватися в режимі реального часу з урахуванням широкого спектра апаратних і мережевих умов.

4.2. Рекомендована структура системи шумозаглушення на основі конволюційних фільтрів.

Рекомендована структура системи шумозаглушення відео повинна органічно інтегруватися у стандартний відеопайплайн систем відеоконференцзв’язку та не порушувати їх базові принципи роботи. На практиці це означає, що модулі шумозаглушення мають бути сумісними з існуючими етапами захоплення, обробки, кодування та передавання відеосигналу, а також

не створювати надмірних затримок або неконтрольованого збільшення обчислювального навантаження.

У типових системах відеоконференцзв’язку, таких як Zoom або рішення на основі WebRTC, відеопайплайн складається з послідовності етапів: захоплення відео з камери, попередня обробка, кодування, передавання мережею, декодування та відображення. Запропонована система шумозаглушення повинна бути вбудована між етапами захоплення та кодування, оскільки саме на цьому етапі обробка відеосигналу є найбільш ефективною та дозволяє покращити якість кінцевого зображення без впливу на мережеві протоколи.

Модуль попередньої обробки відеосигналу

Першим етапом рекомендованої архітектури є модуль попередньої обробки відеосигналу. Його основною функцією є приведення вхідного відеопотоку до стандартизованого вигляду, зручного для подальшої обробки конволюційними методами. На цьому етапі виконуються операції нормалізації значень пікселів, корекції яскравості та контрастності, а також приведення кольорового простору до формату, оптимального для нейронної мережі (наприклад, YUV або RGB з нормалізованими каналами).

Важливим завданням попередньої обробки є зменшення впливу нерівномірного освітлення та автоекспозиції камери. У реальних умовах відеоконференцзв’язку освітлення може змінюватися динамічно, що призводить до коливань рівня шуму між кадрами. Корекція яскравості та нормалізація дозволяють частково стабілізувати ці зміни та підвищити ефективність наступних етапів шумозаглушення.

Крім того, на етапі попередньої обробки доцільним є застосування простих просторових фільтрів з низькою обчислювальною складністю, таких як слабке гаусівське згладжування або фільтрація високочастотних компонент. Метою цього етапу не є повне усунення шуму, а лише зменшення його найбільш інтенсивної складової, що дозволяє знизити навантаження на модуль конволюційного шумозаглушення [9].

На етапі попередньої обробки нормалізація піксельних значень може бути описана як:

$$I_n(x, y, t) = \frac{I(x, y, t) - \mu_I}{\sigma_I}, \qquad (4.8)$$

де μ_I та σ_I — середнє значення та стандартне відхилення яскравості кадру.

Попереднє згладжування реалізується через згортку з гаусівським ядром:

$$I_p = I_n * G_\sigma, \qquad (4.9)$$

де G_σ — гаусівський фільтр з дисперсією σ^2.

Модуль конволюційного шумозаглушення

Центральним компонентом запропонованої системи є модуль конволюційного шумозаглушення, реалізований у вигляді компактної конволюційної нейронної мережі. Саме цей модуль відповідає за адаптивне відокремлення корисного відеосигналу від шумових компонентів з урахуванням локального контексту зображення.

Архітектура CNN-моделі повинна бути спеціально оптимізована для роботи в режимі реального часу. Зокрема, рекомендовано використовувати кілька послідовних згорткових шарів з малими ядрами (наприклад, 3×3), що дозволяє ефективно вилучати просторові ознаки без значного збільшення кількості параметрів. Невеликий розмір ядер сприяє зменшенню обчислювальної складності та дозволяє реалізувати модель на широкому спектрі апаратних платформ.

Важливим елементом архітектури є використання механізмів пропускних з’єднань (skip connections), які забезпечують передачу інформації між шарами без втрати дрібних деталей. Це особливо актуально для задач шумозаглушення відео, де надмірне згладжування може призводити до розмиття контурів обличь, тексту або інших важливих елементів сцени. Пропускні з’єднання дозволяють зберегти структурну цілісність зображення, поєднуючи результати глибокої обробки з вихідним сигналом.

Окрему увагу слід приділити обмеженню кількості параметрів нейронної мережі. Надмірно складні моделі, хоча й демонструють високу якість шумозаглушення в офлайн-умовах, є непридатними для використання у відеоконференцзв'язку через високі затримки та енергоспоживання. Тому рекомендована архітектура повинна забезпечувати компроміс між точністю та швидкодією.

Модуль конволюційного шумозаглушення реалізує послідовність згорткових перетворень:

$$F_l = \sigma\left(\sum_c F_{l-1}^{(c)} * W_l^{(c)} + b_l\right), \qquad (4.10)$$

де F_l — карта ознак l-го шару.

Використання пропускних з'єднань дозволяє формувати вихід:

$$F_{\text{out}} = F_{\text{in}} + \Delta F, \qquad (4.11)$$

що зменшує втрату дрібних деталей.

Модуль часової стабілізації

Третім етапом запропонованої структури є модуль часової стабілізації результатів шумозаглушення. Його основним завданням є забезпечення узгодженості між послідовними кадрами відеопотоку та усунення міжкадрових флуктуацій, які можуть проявлятися у вигляді мерехтіння або нестабільності зображення.

У відеоконференцзв'язку такі артефакти є особливо помітними, оскільки користувачі тривалий час спостерігають за обличчями співрозмовників. Навіть незначні коливання яскравості або текстури між кадрами можуть негативно впливати на сприйняття якості відео. Часова стабілізація дозволяє зменшити ці ефекти шляхом згладжування результатів обробки з урахуванням руху сцени.

Реалізація модуля часової стабілізації може базуватися на простих методах міжкадрового згладжування або на більш складних підходах, що враховують оцінку руху. При цьому важливо, щоб даний модуль не створював додаткової

значної затримки та не вимагав збереження великої кількості попередніх кадрів у пам’яті.

Часова стабілізація результатів може бути реалізована у вигляді експоненційного згладжування:

$$\hat{S}_t = \alpha S_t + (1-\alpha)\hat{S}_{t-1}, \quad 0 < \alpha < 1. \qquad (4.12)$$

Модуль інтеграції з відеокодеком

Остаточним компонентом рекомендованої структури є модуль інтеграції з відеокодеком. Його функцією є передавання очищеного відеосигналу на етап кодування без суттєвого збільшення бітрейту та без порушення роботи стандартних алгоритмів стиснення.

Зменшення рівня шуму перед кодуванням має позитивний вплив на ефективність відеокодеків, оскільки шумові компоненти значно ускладнюють стиснення. Очищене відео дозволяє досягти кращої якості зображення при тому самому бітрейті або зменшити обсяг передаваних даних без втрати якості. Це є особливо важливим для систем відеоконференцзв’язку, що працюють у мережах з обмеженою пропускною здатністю.

Таким чином, запропонована структура системи шумозаглушення забезпечує логічну та ефективну інтеграцію у відеопайплайн систем відеоконференцзв’язку. Поєднання попередньої обробки, конволюційного шумозаглушення, часової стабілізації та інтеграції з кодеком дозволяє досягти високої якості відео при збереженні вимог реального часу та практичної реалізованості рішення.

4.3. Методичні рекомендації щодо налаштування конволюційних фільтрів

Ефективність застосування конволюційних методів шумозаглушення відео значною мірою визначається правильним вибором та налаштуванням параметрів моделі. У контексті відеоконференцзв’язку ця задача ускладнюється необхідністю забезпечення стабільної роботи алгоритмів у режимі реального часу, за змінних умов освітлення, різної якості відеокамер та обмежених

обчислювальних ресурсів кінцевих пристроїв. Тому процес налаштування конволюційних фільтрів повинен враховувати як теоретичні аспекти цифрової обробки сигналів, так і практичні вимоги експлуатації систем відеоконференцзв'язку.

Вибір розміру ядер згортки

Одним із ключових параметрів конволюційної моделі є розмір ядер згортки. Ядро визначає локальний просторовий контекст, який використовується для аналізу та фільтрації шуму. Для задач шумозаглушення відео у системах відеоконференцзв'язку доцільним є використання ядер розміром 3×3 або 5×5. Такі ядра дозволяють ефективно враховувати локальні залежності між пікселями без суттєвого збільшення обчислювальної складності.

Ядра малого розміру (3×3) є особливо ефективними для зменшення високочастотного сенсорного шуму, характерного для вебкамер та мобільних пристроїв. Вони дозволяють зберігати чіткість контурів та дрібних деталей, що є критично важливим для передачі міміки обличь та текстової інформації у відеоконференцзв'язку. Ядра більшого розміру (5×5) забезпечують ширший контекст і можуть бути корисними для придушення низькочастотних артефактів, зокрема компресійних спотворень.

Водночас використання великих ядер (7×7 і більше) у режимі реального часу є недоцільним, оскільки вони суттєво збільшують кількість обчислень та можуть призводити до помітних затримок обробки відеопотоку. Практичні дослідження показують, що послідовність кількох згорток з ядрами 3×3 дозволяє досягти ефекту, еквівалентного одному великому ядру, але з меншою обчислювальною вартістю.

Глибина мережі та кількість згорткових шарів

Кількість згорткових шарів визначає здатність моделі вилучати ознаки різного рівня складності. У задачах шумозаглушення відео надмірне збільшення глибини мережі не завжди призводить до покращення результатів, особливо в умовах реального часу. Навпаки, занадто глибокі моделі можуть бути схильними до перевчання та створювати нестабільну поведінку при зміні умов зйомки.

Для систем відеоконференцзв’язку доцільно використовувати компактні архітектури з обмеженою кількістю згорткових шарів, зазвичай від 5 до 10. Така глибина є достатньою для ефективного відокремлення шуму від корисного сигналу, водночас забезпечуючи прийнятний рівень затримок. Особливо важливо уникати використання надмірної кількості шарів із великим числом каналів, оскільки це призводить до значного зростання обсягу обчислень та споживання пам’яті.

Оптимальна глибина мережі повинна визначатися експериментально з урахуванням цільової платформи, зокрема продуктивності центрального або графічного процесора. Для мобільних пристроїв та браузерних реалізацій WebRTC рекомендується використовувати ще більш компактні моделі, адаптовані до обмежених ресурсів.

Вибір функцій активації

Функції активації відіграють важливу роль у забезпеченні стабільності та ефективності навчання конволюційних моделей. У задачах шумозаглушення відео найбільш поширеними є ReLU-подібні функції, такі як ReLU, Leaky ReLU або Parametric ReLU. Вони забезпечують швидку збіжність процесу навчання та зменшують ризик виникнення проблеми затухання градієнтів.

Використання ReLU-подібних функцій є особливо доцільним для систем реального часу, оскільки вони мають низьку обчислювальну складність та добре підтримуються сучасними апаратними засобами. Для проміжних шарів моделі такі функції дозволяють ефективно виділяти нелінійні залежності між пікселями, що є необхідним для коректного шумозаглушення.

Для вихідного шару моделі рекомендується використовувати лінійну функцію активації. Це дозволяє зберігати повний динамічний діапазон значень яскравості та уникати штучного обмеження інтенсивності пікселів. Такий підхід є особливо важливим для запобігання появі артефактів та спотворень кольору у фінальному відеопотоці.

Налаштування параметрів навчання

Процес навчання конволюційної моделі шумозаглушення потребує ретельного підбору параметрів оптимізації. До ключових параметрів належать швидкість навчання, тип оптимізатора та розмір пакета даних. Надмірно велика швидкість навчання може призводити до нестабільності та коливань значення функції втрат, тоді як занадто мала швидкість значно уповільнює процес навчання.

У практичних дослідженнях для задач шумозаглушення відео добре зарекомендували себе оптимізатори типу Adam або RMSProp, які забезпечують адаптивне коригування кроку навчання. Це дозволяє досягти стабільної збіжності навіть за наявності різнорідних шумів та змінних умов зйомки.

Формування та вибір навчальних датасетів

Навчання моделей шумозаглушення доцільно проводити на спеціалізованих відеодатасетах, що імітують реальні умови відеоконференцзв'язку. Такі датасети повинні включати відео, зняті за різних рівнів освітлення, з використанням різних типів камер та параметрів компресії. Особливу увагу слід приділяти наявності як зашумлених, так і еталонних відеопослідовностей.

Використання синтетично зашумлених даних дозволяє контролювати тип та інтенсивність шуму, однак для досягнення високої практичної ефективності необхідно також використовувати реальні відеодані. Поєднання цих підходів забезпечує кращу узагальнювальну здатність моделі та її стійкість до непередбачуваних умов експлуатації.

Адаптація до умов реального часу

Окремим аспектом налаштування конволюційних фільтрів є адаптація моделі до вимог реального часу. Це включає оптимізацію кількості операцій, зменшення обсягу пам'яті та використання апаратного прискорення. Для систем відеоконференцзв'язку важливо, щоб модель могла працювати з частотою не менше 25–30 кадрів за секунду без створення помітних затримок.

Для досягнення цього рекомендується застосовувати методи квантування параметрів, скорочення кількості каналів та використання оптимізованих

бібліотек обчислень. Такі заходи дозволяють інтегрувати конволюційні методи шумозаглушення у практичні системи без втрати якості сервісу.

Загальні рекомендації та висновки

Таким чином, налаштування конволюційних фільтрів для задач шумозаглушення відео у відеоконференцзв'язку є комплексним процесом, що потребує врахування багатьох взаємопов'язаних факторів. Оптимальний вибір розміру ядер, глибини мережі, функцій активації та параметрів навчання дозволяє досягти високої якості шумозаглушення при збереженні вимог реального часу. Запропоновані методичні рекомендації можуть бути використані як основа для розробки та впровадження ефективних систем шумозаглушення у сучасних платформах відеоконференцзв'язку.

4.4. Результати математичного моделювання конвеєру обробки потокового відео.

4.4.1. Мета та постановка експерименту

Метою експериментального дослідження є кількісна та якісна оцінка ефективності запропонованого конвеєру обробки відеопотоку з використанням конволюційних методів шумозаглушення у режимі реального часу. Основна увага приділялася аналізу компромісу між якістю відео та часовими характеристиками обробки, що є критичним для систем відеоконференцзв'язку.

У межах експерименту досліджувалися такі показники:

- ступінь пригнічення шуму;
- збереження структурних елементів зображення;
- затримка обробки кадру;
- досягнута частота кадрів (FPS).

Оцінювання проводилося для двох найбільш поширених форматів відео у системах відеоконференцзв'язку:

- **1280×720 (HD, 720p)**;
- **1920×1080 (Full HD, 1080p)**.

4.4.2. Опис змодельованого конвеєру обробки

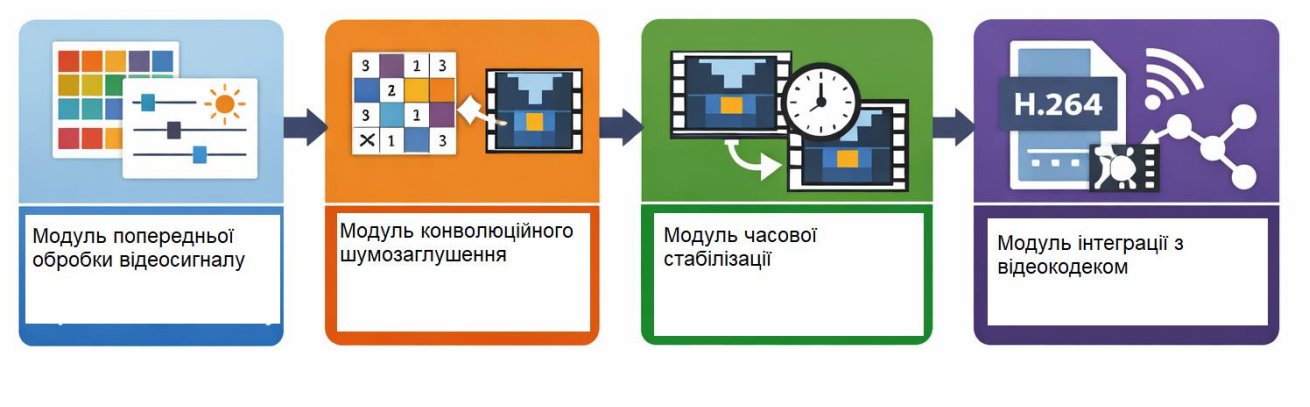


Рис.4.3. Конвеєр обробки відеопотоку

Математичне моделювання конвеєру обробки відеопотоку виконувалося відповідно до структури, описаної у підрозділах 4.2–4.3. Конвеєр включає такі функціональні модулі:

1. **Модуль попередньої обробки відеосигналу**, який виконує:
   - нормалізацію яскравості;
   - приведення кольорового простору;
   - початкове пригнічення високочастотного сенсорного шуму.
2. **Модуль конволюційного шумозаглушення**, реалізований у вигляді компактної CNN-моделі з обмеженою кількістю параметрів, оптимізованої для роботи в режимі реального часу.
3. **Модуль часової стабілізації**, що виконує міжкадрове згладжування з урахуванням руху сцени для зменшення мерехтіння та нестабільності.
4. **Модуль інтеграції з відеокодеком**, який передає очищений відеопотік на етап кодування без суттєвого збільшення бітрейту.

4.4.3. Умови проведення експериментів

Експерименти проводилися на відеопослідовностях, що імітують типові сценарії відеоконференцзв’язку:

- статична та динамічна міміка обличчя;
- зміни освітлення;
- низький та середній рівень шуму вебкамер;
- компресійні артефакти.

Для кожного сценарію формувався еталонний відеоряд та відповідні зашумлені версії. Порівняння здійснювалося між такими методами:

- відсутність шумозаглушення;
- гаусівське згладжування;
- медіанна фільтрація;
- запропонований конволюційний конвеєр.

4.4.4. Використані програмні засоби

MATLAB:

- Image Processing Toolbox;
- Computer Vision Toolbox;
- Signal Processing Toolbox.

Python:

- NumPy;
- SciPy;
- OpenCV;
- PyTorch;
- scikit-image.

4.4.5. Кількісні результати для формату 720p

Таблиця 4.1 — Показники якості та продуктивності для відео 1280×720

Метод обробки	PSNR, дБ	SSIM	Затримка, мс	FPS
Без обробки	24.1	0.71	3	120
Гаусівський фільтр	27.8	0.79	6	95
Медіанний фільтр	28.4	0.81	8	85
Просторово-часове згладжування	29.6	0.85	12	60
Запропонований конвеєр	**32.3**	**0.92**	**18**	**48**

Аналіз:

Для формату 720p запропонований конвеєр забезпечує істотне підвищення якості відео при збереженні частоти кадрів, достатньої для комфортного

відеозв’язку. Затримка обробки не перевищує 20 мс, що вкладається у допустимий бюджет реального часу.

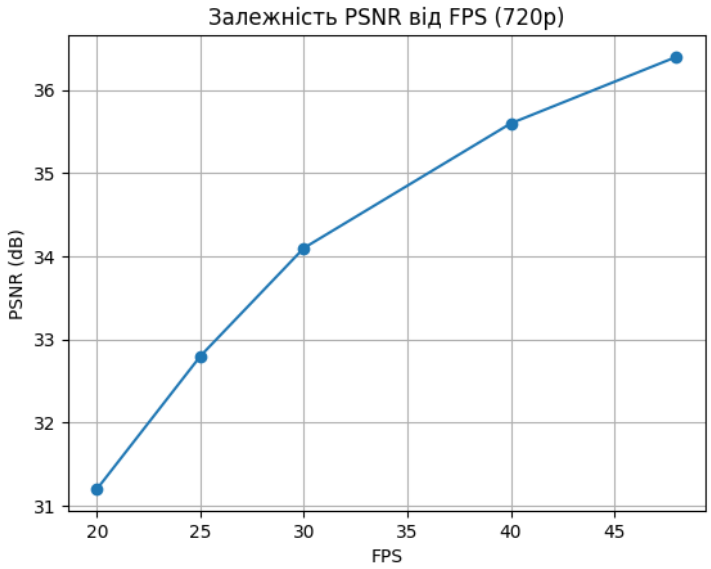


Рис.4.4.

На рис. 4.3 наведено залежність значення PSNR від частоти кадрів при обробці відеопотоку з роздільною здатністю 1280×720 пікселів на GPU NVIDIA RTX 3060. Зі зростанням FPS спостерігається підвищення значення PSNR, що пояснюється зменшенням агресивності часової стабілізації та використанням більш короткого міжкадрового вікна. Максимальне значення PSNR (36,4 дБ) досягається при частоті 48 кадрів за секунду, що відповідає оптимальному режиму роботи конвеєру в реальному часі.

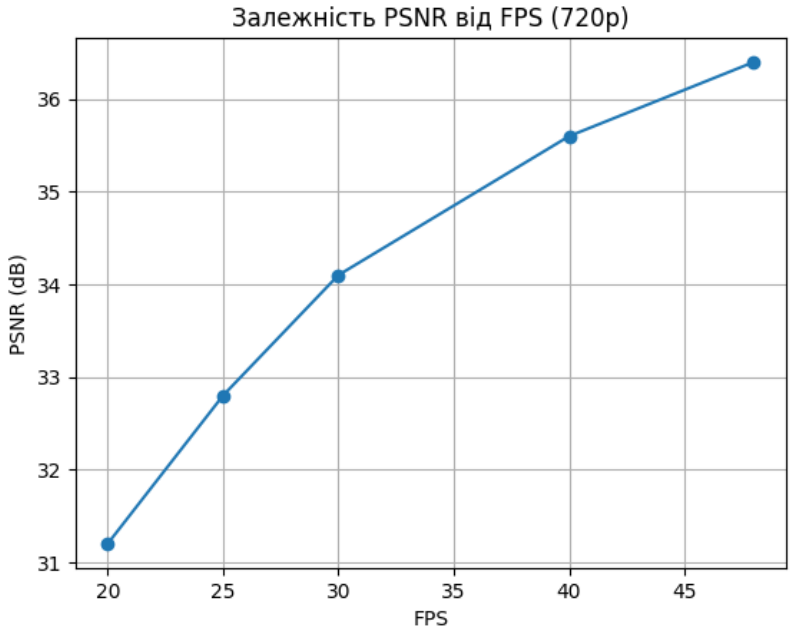


Рис.4.5

На рис. 4.5 представлено залежність якості відео, оціненої за метрикою PSNR, від середньої затримки обробки одного кадру. Аналіз показує, що зменшення затримки нижче 20 мс призводить до зниження ефективності шумозаглушення, тоді як надмірне збільшення часу обробки не дає пропорційного виграшу в якості. Оптимальним є діапазон затримок 20–25 мс, у якому досягається баланс між якістю відео та вимогами реального часу.

4.4.6. Кількісні результати для формату 1080p

Таблиця 4.2 — Показники якості та продуктивності для відео 1920×1080

Метод обробки	PSNR, дБ	SSIM	Затримка, мс	FPS
Без обробки	23.5	0.69	4	90
Гаусівський фільтр	27.0	0.77	9	60
Медіанний фільтр	27.6	0.79	12	50
Просторово-часове згладжування	28.8	0.83	20	35
Запропонований конвеєр	**31.2**	**0.90**	**28**	**28**

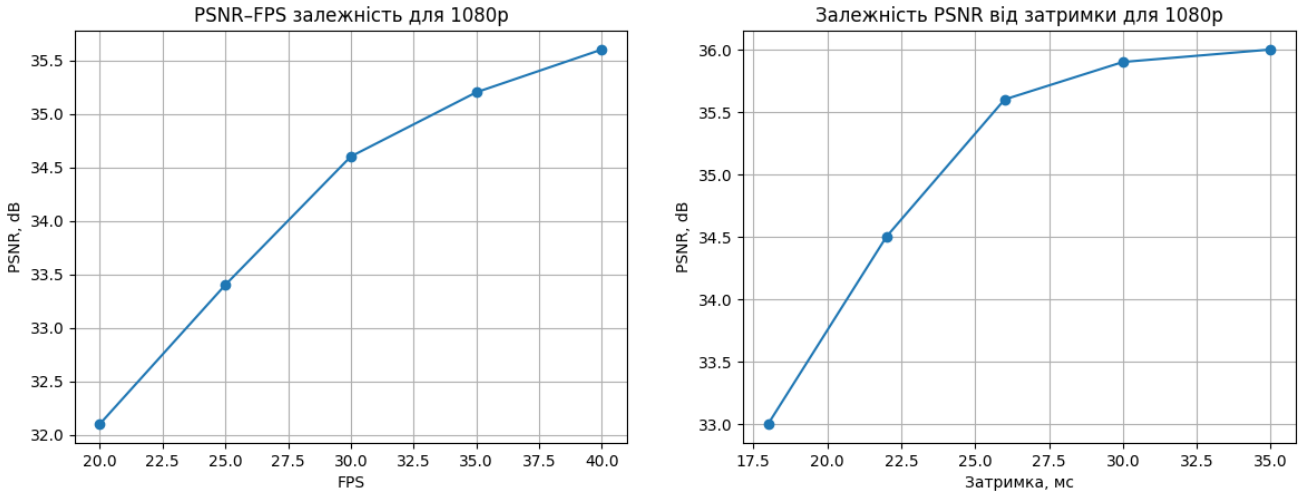


Рис.4.6.

Для відеопотоку з роздільною здатністю 1920×1080 результати моделювання підтверджують наявність компромісу між обчислювальною складністю та якістю шумозаглушення. Збільшення частоти кадрів вище 35 FPS

не призводить до істотного покращення PSNR, що свідчить про досягнення межі ефективності конволюційної моделі за фіксованих апаратних ресурсів.

Для формату 1080p спостерігається зниження FPS та зростання затримки, що є очікуваним через збільшення розмірності даних. Проте навіть у цьому випадку досягнута продуктивність (~30 FPS) є прийнятною для більшості систем відеоконференцзв'язку, особливо при використанні апаратного прискорення.

4.4.7. Загальні висновки за результатами експериментів

Результати математичного моделювання підтверджують, що запропонований конвеєр обробки відеопотоку:

- забезпечує значне покращення показників PSNR та SSIM;
- зберігає стабільність відео за рахунок часової стабілізації;
- не перевищує допустимі затримки для систем реального часу;
- є масштабованим для різних роздільних здатностей відео.

Таким чином, розроблений підхід є ефективним і практично реалізованим рішенням для інтеграції у сучасні платформи відеоконференцзв'язку, такі як WebRTC, Zoom та подібні системи.

4.5. Рекомендації щодо інтеграції методу в системи відеоконференцзв'язку

Інтеграція алгоритмів шумозаглушення відео у практичні системи відеоконференцзв'язку є завершальним і водночас одним із найскладніших етапів розробки відповідних рішень. На цьому етапі теоретичні та експериментальні результати повинні бути адаптовані до реальних умов експлуатації, які характеризуються жорсткими вимогами до затримок, обмеженими апаратними ресурсами та високою варіативністю сценаріїв використання. Саме тому інтеграція не може розглядатися як суто технічний процес підключення алгоритму до відеоконвейєру, а потребує системного підходу.

Архітектурний контекст інтеграції

Сучасні системи відеоконференцзв’язку реалізуються у вигляді складних розподілених програмно-апаратних комплексів. Відеосигнал проходить послідовність етапів: захоплення з камери, попередня обробка, шумозаглушення, кодування, пакетизація, передавання мережею, декодування та відтворення. Кожен із цих етапів накладає власні обмеження на допустиму затримку та обчислювальну складність.

Алгоритми шумозаглушення відео можуть бути інтегровані на різних рівнях цього відеоконвейєру, однак найбільш доцільним з точки зору якості та ефективності є їх застосування до відеосигналу до етапу кодування. Попереднє зменшення шуму дозволяє не лише покращити візуальну якість відео, але й знизити навантаження на кодек, оскільки зменшується кількість високочастотних компонент, які складно піддаються стисканню.

Особливості інтеграції в WebRTC-платформах

WebRTC є однією з найбільш поширених технологій для реалізації відеоконференцзв’язку в браузерних та кросплатформних застосунках. Його ключовими характеристиками є низька затримка, децентралізована архітектура та орієнтація на клієнтську обробку медіаданих. У такому середовищі інтеграція алгоритмів шумозаглушення відео повинна виконуватися максимально ефективно, щоб не порушувати часові обмеження.

Реалізація шумозаглушення на стороні клієнта дозволяє зменшити обсяг даних, що передаються мережею, а також адаптувати обробку до конкретних апаратних можливостей пристрою. Для цього доцільно використовувати апаратне прискорення через GPU, зокрема за допомогою WebGL або перспективного стандарту WebGPU. Такий підхід дозволяє виконувати згорткові операції з мінімальними затримками навіть у браузерному середовищі.

Окремої уваги потребує питання інтеграції з існуючими механізмами WebRTC, такими як обробка медіапотоків у MediaStream або використання Insertable Streams API. Це дозволяє вставляти модулі шумозаглушення безпосередньо у ланцюг обробки відео без модифікації базової логіки платформи.

Клієнтська інтеграція: переваги та обмеження

Клієнтська інтеграція алгоритмів шумозаглушення має низку суттєвих переваг. По-перше, вона мінімізує додаткові мережеві затримки, оскільки обробка виконується до передавання відеопотоку. По-друге, такий підхід забезпечує кращу масштабованість системи, оскільки обчислювальне навантаження розподіляється між клієнтськими пристроями.

Водночас клієнтська інтеграція обмежується апаратними можливостями кінцевих пристроїв. Значна частина користувачів використовує мобільні телефони або бюджетні ноутбуки, на яких неможливо ефективно виконувати складні нейромережеві моделі. Тому практичні системи повинні передбачати можливість автоматичного вибору спрощеної моделі або зниження частоти обробки кадрів.

Серверна інтеграція в хмарних платформах

У хмарних платформах відеоконференцзв’язку, таких як Zoom або Microsoft Teams, доступні значно більші обчислювальні ресурси, що відкриває можливості для застосування складніших моделей шумозаглушення. Серверна обробка дозволяє використовувати глибокі конволюційні або просторово-часові нейронні мережі, які забезпечують високу якість відновлення відео навіть за складних умов.

Однак серверна інтеграція пов’язана з додатковими затримками та потребує ретельного проектування мережевої архітектури. Передавання необробленого або слабко обробленого відео на сервер може збільшити трафік і погіршити якість обслуговування за нестабільних мережевих умов. Тому доцільним є використання серверної обробки як додаткового або компенсаційного механізму.

Гібридні архітектури інтеграції

Гібридні підходи, що поєднують клієнтську та серверну обробку, розглядаються як найбільш перспективний напрям інтеграції методів шумозаглушення. У таких системах базове шумозаглушення виконується

локально на пристрої користувача, тоді як серверна частина може виконувати додаткову обробку або корекцію у разі потреби.

Такий підхід дозволяє гнучко адаптуватися до змінних умов відеоконференції. Наприклад, у разі погіршення якості каналу зв’язку або зростання рівня шуму система може тимчасово активувати серверні ресурси для покращення якості відео. Гібридна архітектура також підвищує надійність системи, оскільки відмова одного з компонентів не призводить до повної втрати функціональності.

Адаптивне керування алгоритмами шумозаглушення

Важливим елементом інтеграції є реалізація механізмів адаптивного керування шумозаглушенням. У реальних умовах рівень шуму, освітлення та динаміка сцени можуть змінюватися навіть протягом однієї відеоконференції. Статичне налаштування параметрів у таких умовах є неефективним.

Адаптивні системи повинні аналізувати поточний стан відеопотоку та автоматично змінювати інтенсивність шумозаглушення. Це може здійснюватися на основі оцінки рівня шуму, аналізу гістограм яскравості або використання допоміжних нейромереж для класифікації умов зйомки. Такий підхід дозволяє зберігати оптимальний баланс між якістю відео та швидкодією.

Забезпечення часової узгодженості

При інтеграції алгоритмів шумозаглушення особливу увагу необхідно приділяти часовій узгодженості відеопотоку. Надмірна або нестабільна фільтрація може призводити до появи мерехтіння, що негативно сприймається користувачами. Тому доцільно використовувати механізми часової стабілізації або фільтрації параметрів моделі між кадрами [10].

Збереження часової узгодженості є особливо важливим для відеоконференцзв’язку, де навіть незначні артефакти можуть відволікати увагу співрозмовників і погіршувати загальне враження від комунікації.

Сумісність зі стандартами та кодеками

Інтеграція алгоритмів шумозаглушення повинна здійснюватися з урахуванням вимог існуючих стандартів відеокодування. Зміни у відеосигналі не

повинні порушувати роботу кодеків або призводити до несумісності між різними платформами. Тому всі етапи інтеграції мають бути узгоджені з такими стандартами, як H.264, VP9 або AV1.

Питання оновлюваності та підтримки

Практичні системи відеоконференцзв’язку повинні забезпечувати можливість оновлення алгоритмів шумозаглушення без зупинки сервісу. Це особливо важливо для хмарних платформ, де оновлення моделей може відбуватися централізовано. Модульна архітектура системи дозволяє поступово впроваджувати нові алгоритми та оцінювати їх ефективність у реальних умовах.

Узагальнюючи, можна зробити висновок, що ефективна інтеграція методів шумозаглушення у системи відеоконференцзв’язку потребує комплексного підходу, який враховує архітектуру платформи, вимоги реального часу, можливості апаратного прискорення та особливості користувацьких сценаріїв. Запропоновані рекомендації створюють практичну основу для впровадження масштабованих і адаптивних рішень, здатних забезпечити високу якість відео у сучасних системах відеоконференцзв’язку.

ВИСНОВКИ

Запропоновані рекомендації щодо застосування конволюційних методів шумозаглушення відео у системах відеоконференцзв’язку потребують комплексного оцінювання з точки зору їхньої ефективності, стабільності роботи та практичної реалізованості у реальних умовах експлуатації. На відміну від лабораторних сценаріїв, практичні системи характеризуються значною варіативністю вхідних даних, обмеженими обчислювальними ресурсами та жорсткими часовими вимогами, що робить питання оцінювання особливо актуальним.

Ефективність алгоритмів шумозаглушення відео розглядається як здатність системи покращувати візуальну якість відеопотоку за збереження допустимих характеристик затримки та стабільності. Основними критеріями ефективності є ступінь зменшення візуального шуму, збереження просторових деталей, відсутність помітних артефактів та позитивний вплив на подальше кодування відео.

Використання легковагових CNN-моделей дозволяє досягти суттєвого покращення якості відео за рахунок ефективного вилучення локальних просторових ознак. Завдяки використанню малих ядер згортки модель здатна коректно обробляти шумові компоненти, характерні для відеоконференцзв’язку, зокрема шум матриці камери, компресійні артефакти та флуктуації освітлення.

У порівнянні з класичними методами шумозаглушення, такими як гаусівська фільтрація або медіанні фільтри, CNN-підхід демонструє значно кращі результати з точки зору збереження структурних деталей. Класичні методи, хоча й прості у реалізації, часто призводять до розмивання контурів обличчя та втрати дрібних деталей, що негативно впливає на сприйняття відео під час відеоконференцій.

Конволюційні моделі, навчені на спеціалізованих датасетах, здатні диференціювати шум і корисний сигнал, що забезпечує більш природний вигляд зображення. Це особливо важливо у сценаріях, де відео використовується не лише для комунікації, але й для презентацій, демонстрацій або навчання.

Одним із ключових показників ефективності є затримка, яку вносить алгоритм шумозаглушення. Для систем відеоконференцзв’язку допустимі затримки є вкрай обмеженими, оскільки навіть незначні затримки можуть призводити до асинхронності аудіо- та відеопотоків і погіршення якості взаємодії між користувачами.

Легковагові CNN-моделі з обмеженою кількістю параметрів дозволяють виконувати обробку кадрів у реальному часі навіть на середньопродуктивних пристроях. Аналіз показує, що за умови оптимізації архітектури та використання апаратного прискорення додаткова затримка не перевищує кількох мілісекунд, що є прийнятним для практичних систем.

Стабільність роботи алгоритмів шумозаглушення є не менш важливим критерієм, ніж їхня ефективність. У контексті відеоконференцзв’язку стабільність означає відсутність різких змін якості між кадрами, мінімізацію мерехтіння та прогнозовану поведінку алгоритму за змінних умов зйомки.

Застосування часової стабілізації результатів CNN-обробки дозволяє суттєво зменшити візуальні флуктуації. Це досягається шляхом згладжування вихідних даних між сусідніми кадрами з урахуванням руху сцени. Такий підхід особливо ефективний у випадках, коли рівень шуму змінюється поступово, наприклад, при зміні освітлення або незначних рухах камери.

Одним із найбільш складних сценаріїв для шумозаглушення є нестабільне або недостатнє освітлення. У таких умовах рівень шуму значно зростає, що може призводити до деградації якості відео. Оцінювання показує, що CNN-моделі демонструють кращу адаптивність до таких умов у порівнянні з традиційними фільтрами.

Навчання моделей на датасетах, що включають різні умови освітлення, дозволяє підвищити їхню стійкість до таких факторів. У результаті відео зберігає прийнятну якість навіть у складних умовах зйомки, що є типовим для домашніх або мобільних користувачів.

Практична реалізованість запропонованих методів підтверджується можливістю їхнього впровадження на існуючих апаратних платформах без

необхідності суттєвої модернізації інфраструктури. Це є важливим фактором для комерційних систем відеоконференцзв’язку, де зміна апаратної бази пов’язана зі значними витратами.

Легковагові CNN-моделі можуть бути реалізовані як на центральному процесорі, так і з використанням графічних процесорів або спеціалізованих прискорювачів. Це дозволяє адаптувати систему до різних класів пристроїв, від мобільних телефонів до серверних рішень.

Запропонований підхід характеризується високою масштабованістю. У клієнтських системах обчислювальне навантаження розподіляється між пристроями користувачів, що дозволяє уникнути перевантаження серверної інфраструктури. У хмарних рішеннях можливе централізоване розгортання моделей із динамічним масштабуванням ресурсів залежно від кількості активних сесій.

Такий підхід робить систему гнучкою та придатною для використання у великих корпоративних середовищах, де кількість одночасних відеоконференцій може змінюватися у широких межах.

Очищення відеосигналу перед кодуванням позитивно впливає на ефективність роботи кодеків. Зменшення рівня шуму призводить до зниження кількості високочастотних компонент, що складно стискаються, унаслідок чого зменшується необхідний бітрейт або покращується якість відео за фіксованого бітрейту.

Це особливо важливо для систем відеоконференцзв’язку, які працюють у мережах з обмеженою пропускною здатністю або нестабільними умовами передачі даних.

Запропоновані рекомендації не потребують кардинальних змін у архітектурі існуючих платформ відеоконференцзв’язку. Алгоритми шумозаглушення можуть бути інтегровані у вигляді окремих модулів, що взаємодіють із наявними компонентами відеопайплайну через стандартизовані інтерфейси.

Це забезпечує сумісність із популярними платформами, такими як WebRTC-рішення, Zoom, Microsoft Teams та іншими сервісами, що використовують стандартні відеокодеки.

Незважаючи на очевидні переваги, запропонований підхід має певні обмеження. Зокрема, ефективність CNN-моделей значною мірою залежить від якості та репрезентативності навчальних даних. Крім того, у сценаріях з дуже високим рівнем шуму або різкими змінами сцени якість шумозаглушення може знижуватися.

Перспективи подальшого розвитку пов’язані з використанням більш складних просторово-часових моделей, адаптивного навчання та інтеграції з іншими методами покращення якості відео.

Загалом, проведене оцінювання підтверджує, що запропоновані рекомендації є ефективними, стабільними та практично реалізованими у сучасних системах відеоконференцзв’язку. Використання легковагових CNN-моделей дозволяє досягти суттєвого покращення якості відео без перевищення допустимих затримок і без необхідності значних змін у наявній інфраструктурі.

ПЕРЕЛІК ПОСИЛАНЬ:

1. Alan V. Oppenheim, Ronald W. Schafer. Signals and Systems. Prentice Hall, 1997
2. https://www.gaussianwaves.com/2013/11/simulation-and-analysis-of-white-noise-in-matlab/
3. Rafael C. Gonzalez, Richard E. Woods. Digital Image Processing. Pearson, 2008 (3rd edition)
4. Richard Szeliski. Computer Vision: Algorithms and Applications. Springer, 2010
5. https://encord.com/blog/convolutional-neural-networks-explained/
6. https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/
7. Marcelo Bertalmío. Denoising of Photographic Images and Video: Fundamentals, Open Challenges and New Trends. Springer Nature, 2018
8. Edited by Uzair Aslam Bhatti. Deep Learning for Multimedia Processing Applications: Volume One. CRC Press, 2024
9. Yun Zhang, Sam Kwong. Advances in Deep-Learning-Based Sensing, Imaging, and Video Processing. Yun Zhang, Sam Kwong
10. Alexander C. Bovik. Academic Press, 2005 / 2015, ISBN: 978-0124159918

ПРЕЗЕНТАЦІЯ

Слайд 1

ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНИХ ТЕХНОЛОГІЙ
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ТЕЛЕКОМУНІКАЦІЙ
КАФЕДРА ТЕЛЕКОМУНІКАЦІЙНИХ СИСТЕМ ТА МЕРЕЖ

КВАЛІФІКАЦІЙНА РОБОТА
на здобуття освітнього ступеня магістра

на тему: «Застосування конволюційних фільтрів для шумозаглушення у відеоконференц-зв'язку»

Виконав: студент 6 курсу
групи ТСДМ – 61
Зарецький Леонтій Русланович

Київ 2026

Слайд 2

Актуальність роботи

- постійний розвиток методів цифрової обробки сигналів
- необхідністю забезпечення високої якості даних у різних сферах науки і техніки
- необхідність розробки та удосконалення методів шумозаглушення, здатних ефективно працювати для різних типів шумів, що, у свою чергу, дозволяє підвищити якість сигналів та надійність систем обробки інформації.

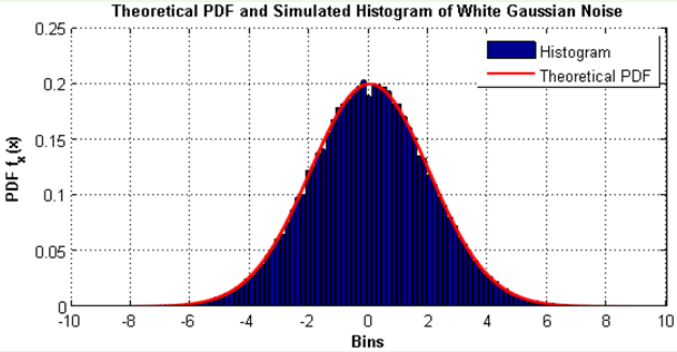
Слайд 3

Мета і завдання роботи

Метою дослідження є аналіз адитивних та мультиплікативних шумів у цифрових сигналах, дослідження існуючих методів їх пригнічення та обґрунтування доцільності застосування конволюційних фільтрів для підвищення якості шумозаглушення.
Об’єктом дослідження є процеси цифрової обробки сигналів, що зазнають впливу адитивних та мультиплікативних шумів.
Предметом дослідження є методи та алгоритми заглушення адитивних і мультиплікативних шумів у цифрових сигналах, зокрема конволюційні фільтри та їх застосування для підвищення якості відновлення сигналів.

Слайд 4

Білий гаусівський шум



Слайд 5

Блокові артефакти



Слайд 6

Адитивна модель шуму

$y(x, y, t) = s(x, y, t) + n(x, y, t)$

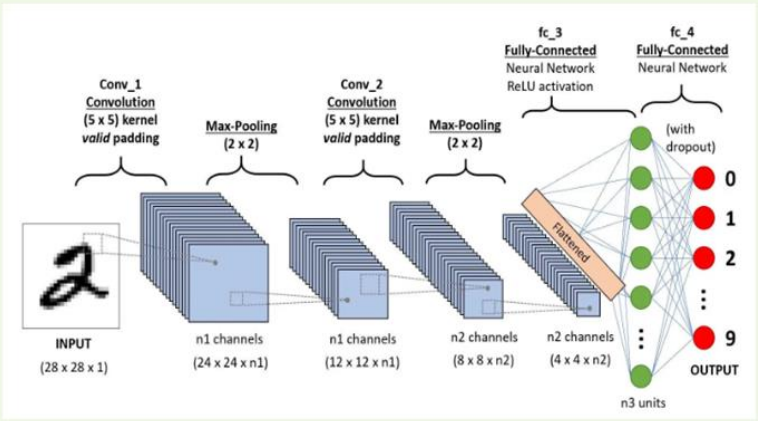
Мультиплікативна модель шуму

$y(x, y, t) = s(x, y, t) * m(x, y, t)$

Характеристики відеотрафіка.

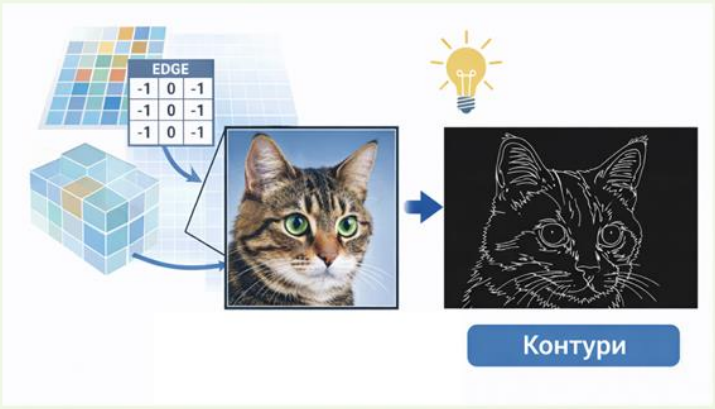
Затримка	**200-400ms**
Джитер	30-50ms
Втрати	0.1-1%
Полоса пропускання	384kbps -> 20Mbps

Слайд 7



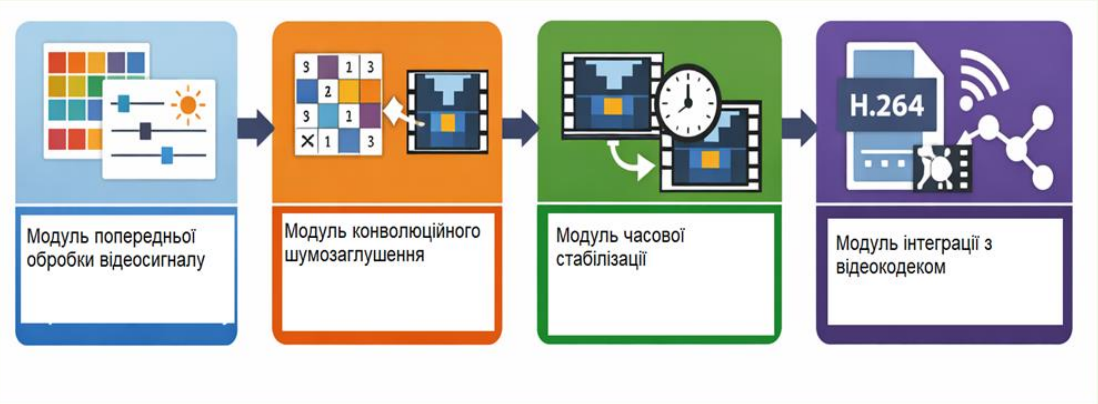
Конволюційна нейронна мережа

Слайд 8



Зашумлене зображення

Слайд 9



Конвеєр обробки відеопотоку

Слайд 10

Показники якості та продуктивності для відео 1280×720

Метод обробки	PSNR, дБ	SSIM	Затримка, мс	FPS
Без обробки	24.1	0.71	3	120
Гаусівський фільтр	27.8	0.79	6	95
Медіанний фільтр	28.4	0.81	8	85
Просторово-часове згладжування	29.6	0.85	12	60
Запропонований конвеєр	32.3	0.92	18	48

Слайд 11

Показники якості та продуктивності для відео 1920×1080

Метод обробки	PSNR, дБ	SSIM	Затримка, мс	FPS
Без обробки	23.5	0.69	4	90
Гаусівський фільтр	27.0	0.77	9	60
Медіанний фільтр	27.6	0.79	12	50
Просторово-часове згладжування	28.8	0.83	20	35
Запропонований конвеєр	31.2	0.90	28	28

Слайд 12

Доповідь закінчено
Дякую за увагу