

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет комп'ютерної інженерії та управління
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Метод нанесення цифрових водяних знаків в аудіофайли

(тема)

Виконав:

студент II курсу, групи СПМ-22-3
Растегаєв Р. І.
(прізвище, ініціали)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва спеціальності)

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

Освітня програма Системне програмування
(повна назва освітньої програми)

Керівник: доц. Мартовицький В.О.
(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри ЕОМ

(підпис)

Коваленко А.А.

(прізвище, ініціали)

2024 р.

Харківський національний університет радіоелектроніки

Факультет _____ комп'ютерної інженерії та управління _____

Кафедра _____ електронних обчислювальних машин _____

Рівень вищої освіти _____ другий (магістерський) _____

Спеціальність _____ 123 «Комп'ютерна інженерія» _____
(код і повна назва)

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

Освітня програма _____ Системне програмування _____
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студенту _____ Растегаєв Роман Іванович _____
(прізвище, ім'я, по батькові)

1. Тема роботи Метод нанесення цифрових водяних знаків в аудіофайли

затверджена наказом по університету від “ 01 ” квітня 2024 р. № 257 Ст

2. Термін подання студентом роботи до екзаменаційної комісії 15 червня 2024 р.

3. Вхідні дані до роботи Набір аудіофайлів

4. Перелік питань, що потрібно опрацювати у роботі _____

1) Огляд предметної області _____

2) Огляд методів стеганографії _____

3) Розробка запропонованого методу _____

4) Проведення експерименту _____

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) 12 слайдів

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

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

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Аналіз сучасних досліджень в предметній області	02.04.24-08.04.24	
2	Огляд підходів до ЦВЗ	09.04.24-16.04.24	
3	Розробка методу	17.04.24-22.04.24	
4	Проведення експериментів	23.04.24-06.05.24	
5	Оформлення матеріалів кваліфікаційної роботи	07.05.24-23.05.24	
6	Подання кваліфікаційної роботи керівникові та її попередній захист	24.05.24-03.06.24	
7	Подання кваліфікаційної роботи на рецензування	04.06.24-07.06.24	

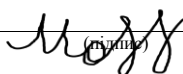
Дата видачі завдання 01 квітня 2024 р.

Студент

(підпис)



Керівник роботи



доц. Мартовицький В.О.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 53 с., 10 рис., 9 табл., 1 дод., 34 джерел.

СТЕГАНОГРАФІЯ, ДКТ, ЗАХИСТ АВТОРСЬКИХ ПРАВ, АУДІО, ЗАХИСТ ІНФОРМАЦІЇ.

Метою кваліфікаційної роботи є розробка та дослідження ефективного методу нанесення цифрових водяних знаків в аудіофайли для захисту авторських прав, забезпечення автентичності та цілісності аудіоконтенту, а також для відстеження та ідентифікації джерел незаконного розповсюдження аудіофайлів.

Завдання роботи:

- аналіз існуючих методів цифрових водяних знаків;
- розробка власного методу нанесення цифрових водяних знаків;
- моделювання та реалізація розробленого методу;
- дослідження стійкості та ефективності розробленого методу.

ABSTRACT

Master's thesis: 53 pages, 10 figures, 9 tables, 1 appendices, 34 sources.

STEGANOGRAPHY, DCT, COPYRIGHT PROTECTION, AUDIO, INFORMATION PROTECTION.

The major goal of this thesis is to develop and research an effective method for applying digital watermarks to audio files to protect copyright, ensure the authenticity and integrity of audio content, and track and identify sources of illegal distribution of audio files.

Objectives:

- analysis of existing digital watermarking methods;
- development of an own method for applying digital watermarks;
- modelling and implementation of the developed method;
- studying the stability and efficiency of the developed method.

Translated with www.DeepL.com/Translator (free version)

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКРОЧЕНЬ І ТЕРМІНІВ	7
ВСТУП	8
1.1 Постановка проблеми та її зв'язок із важливими науковими чи практичними завданнями	10
1.2 Кількісна оцінка продуктивності методів нанесення цифрових водяних знаків на аудіофайли	11
1.3 Огляд алгоритмів нанесення водяних знаків на аудіо	17
2 СПЕКТРАЛЬНИЙ АНАЛІЗ СИГНАЛІВ	27
2.1 Комплексна еволюція спектральної фази	27
2.2 Математичне представлення комплексної еволюція спектральної фази	28
2.3 Демонстрація роботи CSPE	31
2.4 Швидке перетворення Фур'є	32
3 НАНЕСЕННЯ ВОДЯНИХ ЗНАКІВ НА АУДІО	34
3.1 Запропонована схема нанесення водяних знаків	34
3.2 Експериментальні результати та обговорення	36
ВИСНОВКИ	41
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	43
ДОДАТОК А Графічний матеріал кваліфікаційної роботи	47

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СИМВОЛІВ, ОДИНИЦЬ, СКОРОЧЕНЬ
І ТЕРМІНІВ

- ЦВЗ – Цифрові водяні знаки
ФП – Фур'є-перетворення
ДФП – Дробове Фур'є-перетворення
STFT – Короткочасне Фур'є-перетворення (Short-Time Fourier Transform)
DWT – Дискретне хвилькове перетворення (Discrete Wavelet Transform)
FFT – Швидке Фур'є-перетворення (Fast Fourier Transform)
PSNR – Пікове співвідношення сигнал/шум (Peak Signal-to-Noise Ratio)
SNR – Співвідношення сигнал/шум (Signal-to-Noise Ratio)
MSE – Середньоквадратична помилка (Mean Squared Error)
BER – Коефіцієнт бітових помилок (Bit Error Rate)
DCT – Дискретне косинусне перетворення (Discrete Cosine Transform)
MDCT – Модифіковане дискретне косинусне перетворення (Modified Discrete Cosine Transform)
MP3 – MPEG-1

ВСТУП

Стрімкий розвиток інфокомунікаційних та комп'ютерних технологій та їхня конвергенція призвели до стрімкого зростання обсягів цифрового контенту, який створюється, зберігається, поширюється та застосовується в різних сферах. Термін "контент" в широкому сенсі означає будь-яку цифрову інформацію, таку як цифрове аудіо, відео, графіка, анімація, зображення, текст або будь-які комбінації цих типів. До цього цифрового контенту можна легко отримати доступ, скопіювати, швидко поширювати і масово використовувати без втрати якості, на відміну від ситуації з більш ранніми аналоговими носіями, такими як аудіокасети і касети для відеомагнітофонів. Однак ці переваги цифрових медіа-форматів над аналоговими перетворюються на недоліки з точки зору управління авторськими правами, оскільки можливість необмеженого копіювання без втрати автентичності призвела до значних фінансових втрат для власників авторських прав [1].

Щоб зменшити фінансові втрати від несанкціонованого копіювання, власники контенту здебільшого звертаються до криптографії, яка є одним з найпоширеніших методів захисту цифрового контенту. При використанні методів криптографії контент шифрується перед наданням споживачеві, а потім ключ для розшифровки надається лише тим, хто придбав легальні копії контенту. Однак криптографічні методи не пропонують надійного рішення для боротьби з піратством контенту. Наприклад, пірат може придбати зашифрований контент легальним способом, а потім використати ключ дешифрування для незаконного виготовлення та розповсюдження копій контенту. Іншими словами, після розшифрування контент не має подальшого захисту.

Таким чином, існує гостра потреба в альтернативі або доповненні криптографічних методів захисту аудіо контенту. Для вирішення проблем, з якими стикається криптографія, було запропоновано маркування водяними

знаками, оскільки вони мають потенціал для забезпечення більшої надійностію. Водяні знаки можуть захистити цифровий контент під час його звичайного використання, оскільки інформація про авторське право розміщується в цифровому контенті таким чином, що її неможливо видалити. Ця унікальна особливість водяних знаків робить їх одним із найперспективніших методів захисту цифрового контенту, що стало мотивуючим фактором для більшості досліджень останніх двох десятиліть.

1 ОГЛЯД МЕТОДІВ НАНЕСЕННЯ ЦИФРОВИХ ВОДЯНИХ ЗНАКІВ ДЛЯ ЗАХИСТУ АУДІОФАЙЛІВ

1.1 Постановка проблеми та її зв'язок із важливими науковими чи практичними завданнями

Цифрові водяні знаки (ЦВЗ) в аудіофайлах є однією з сучасних технологій захисту авторського права, що знаходить все ширше застосування у сфері музичної індустрії, медіа та цифрових продуктів. Основна проблема полягає у забезпеченні надійного захисту авторських прав, зберігаючи при цьому якість аудіоконтенту і забезпечуючи стійкість до різних видів атак.

Для того, щоб краще зрозуміти процес нанесення водяних знаків у цифровому аудіо, а потім виявити деякі невирішені питання в поточних реалізаціях, було проведено ретельний огляд літератури. Розглянуті алгоритми є різноманітними і тому поділяються на різні категорії, такі як алгоритми на основі часової області, алгоритми на основі перетворень і гібридні алгоритми відповідно до методології, яку використовує кожен алгоритм. Переваги та недоліки важливих алгоритмів у кожній категорії розглядаються відповідно до критеріїв, визначених нижче:

продуктивність з точки зору непомітності, надійності, пропускну здатності, обчислювальної складності;

надійність результатів, представлених для кожного алгоритму. Це передбачає дослідження використовуваної методології тестування;

встановлення того, чи можна видалити вбудований цифровий водяний знак;

встановлення того, чи включають алгоритми якісь додаткові процеси, щоб обійти компроміс між непомітністю та стійкістю.

Причиною вибору вищезазначених критеріїв є те, що вони є важливими факторами для оцінки ефективності алгоритму водяних знаків у практичному

застосуванні. Крім того, ці критерії корисні для визначення того, чи варто проводити подальші дослідження розглянутих алгоритмів.

1.2 Кількісна оцінка продуктивності методів нанесення цифрових водяних знаків на аудіофайли

Для того, щоб ЦВЗ ефективно виконували свою функцію, необхідно оцінювати їх ефективність за допомогою кількісних методів. Така оцінка включає аналіз різних аспектів, як-от непомітність водяного знака, його стійкість до атак, а також здатність до точного вилучення.

Розглянемо основні критерії кількісної оцінки ефективності аудіо водяних знаків. До цих критеріїв належать:

- непомітність (інтеграція): Водяний знак має бути непомітним для слухача, тобто він не повинен погіршувати якість аудіо [2];
- робастність: Водяний знак повинен залишатися незмінним і розпізнаваним навіть після застосування різних видів обробки, таких як стиснення, редагування або шумові атаки [3-4];
- здатність до вилучення: Важливо, щоб водяний знак можна було надійно вилучити і розпізнати за допомогою спеціальних алгоритмів [5];
- обчислювальна ефективність: Процес вбудовування та вилучення водяного знака повинен бути ефективним з точки зору обчислювальних ресурсів і часу [6].

Розглянуті критерії є ключовими факторами, що впливають на оцінку алгоритмів водяних знаків та визначають їх практичну цінність. Проаналізуємо кожен з цих аспектів, представимо методи їх кількісної оцінки та обговоримо результати тестування різних алгоритмів водяних знаків на прикладі реальних аудіофайлів.

Метою є надання всебічний огляд методів кількісної оцінки ефективності аудіо водяних знаків та провести всебічну оцінку методів нанесення ЦВЗ, що сприятиме подальшому розвитку технологій захисту

авторських прав і підвищенню їхньої надійності у цифровому середовищі.

Загалом, існує три підходи до оцінки якості сприйняття аудіо [7]:

- суб'єктивна оцінка за допомогою тесту на слух людиною;
- об'єктивна оцінка за допомогою сигнально-орієнтованих методів, таких як відношення сигнал/шум (SNR);
- об'єктивна оцінка, яка включає одну з моделей слухової системи людини (HAS), наприклад, перцептивну оцінку якості звуку (PEAQ) [8].

Суб'єктивна оцінка може бути проведена кількома способами. Один з підходів полягає у використанні ABX-тесту [9]. Кожен тест складається з оригінального аудіофайлу А, аудіофайлу з водяним знаком В і невизначеного аудіофайлу Х, який може бути як А, так і В. Слухача просять визначити, чи є Х А або В. Високий рівень правильної ідентифікації свідчить про те, що водяний знак помітний, тоді як приблизно 50% правильної ідентифікації означає, що водяний знак непомітний, оскільки ідентифікація схожа на випадкове вгадування.

Крім того, для оцінки суб'єктивної якості прослуховування контенту з водяними знаками можна використовувати показник середньої оцінки (MOS) [10]. Шкала оцінки MOS наведено в таблиці 1 [10].

Таблиця 1.1– Систематизація атак на водяні знаки

MOS	Опис
5	Чудово.
4	Добре
3	Задовільно
2	Погано
1	Дуже погано.

Однак суб'єктивне оцінювання, засноване на слухових тестах людиною, займає багато часу, а результати можуть бути непослідовними для різних слухачів. Причиною такої неузгодженості є те, що слухові здібності різних слухачів варіюються залежно від їхнього віку, впливу гучних звуків протягом життя і навіть особистих музичних уподобань, не кажучи вже про те, що деякі слухачі можуть бути підготовленими експертами-слухачами. Тому іноді буває важко справедливо порівнювати різні суб'єктивні результати оцінювання, і бажано мати більш об'єктивну оцінку, засновану на певних характеристиках сигналу.

Співвідношення сигнал/шум (SNR) широко використовується як об'єктивна міра якості звуку. Він простий для інтерпретації, зрозумілий у застосуванні та орієнтований на сигнал [11]. Згідно з рекомендації міжнародної федерації фонографічної індустрії IFPI, коли SNR перевищує 20 децибел (дБ), звукові водяні знаки будуть вважатися непомітними. SNR можна сформулювати наступним чином:

$$SNR = 10 \log_{10} \frac{\sum_n s^2(n)}{\sum_n [s(n) - s'(n)]^2} \quad (1.1)$$

де $s(n)$ - часова область оригінального сигналу, а $s'(n)$ - часова область сигналу з водяним знаком.

Оскільки рівняння (1.1) однаково зважає всі помилки часової області, не беручи до уваги енергію, що змінюється в часі, і спотворення, що змінюються в часі, можна отримати покращену оцінку, якщо обчислити SNR для коротких кадрів і усереднити результати. Кадрова міра, яка називається "Сегментне відношення сигнал/шум" (SNR_{seg}), визначається наступним чином [12]:

$$SNR_{seg} = \frac{1}{M} \sum_{j=1}^M 10 \log_{10} \left[\sum_{n=N*(j-1)+1}^{N*j} \frac{s^2(n)}{[s(n) - s'(n)]^2} \right] \quad (1.2)$$

де M - кількість кадрів, а N - розмір кадру.

Проблеми з SNRseg виникають при включенні кадрів без звуку, оскільки вони можуть призвести до великих від'ємних значень SNRseg. Цю проблему можна вирішити, встановивши низький поріг і замінивши всі кадри з SNRseg нижче цього порогу на рівні порогу (розумним є поріг 0 дБ) [12]. З іншого боку, кадри з SNRseg понад 35 дБ не сприймаються слухачами як такі, що суттєво відрізняються, але впливають на результуючий SNRseg. Верхній поріг (зазвичай 35 дБ) можна використовувати для скидання будь-яких незвично високих значень SNRseg до цього верхнього порогу .

Низький SNR або SNRseg чітко вказує на те, що спотворення, внесені водяними знаками, чутні, але високий SNR або SNRseg не є достатнім для того, щоб стверджувати, що водяний знак є прозорим для сприйняття, оскільки ця міра не враховує жодної моделі HAS. На основі багатьох проведених експериментів поведінка HAS була детально досліджена багатьма науковцями [13, 14]. Дослідження досягли значного прогресу у визначенні характеристик HAS. Були запропоновані деякі терміни, такі як "абсолютний поріг чутності", "одночасне маскування" і "тимчасове маскування" [15]. "Абсолютний поріг чутності" характеризує кількість енергії, необхідної для чистого тону, щоб його міг розпізнати слухач у безшумному середовищі. "Маскування" - це явище, коли один звук стає нечутним через присутність іншого звуку. Це явище може відбуватися в частотній області, що називається "одночасним маскуванням", або в часовій області, що називається "тимчасовим маскуванням". Для того, щоб точніше відобразити людське сприйняття, краще мати об'єктивну оцінку, яка включає одну з моделей HAS.

PEAQ є одним з таких об'єктивних методів оцінки. Він був визначений як рекомендаційний в стандартах у BS.1387 [16]. Результатом PEAQ є оцінка об'єктивних відмінностей (ODG). Він класифікує відмінності у сприйнятті між оригінальним аудіосигналом і аудіосигналом з водяними знаками. Значення ODG знаходиться в діапазоні $[-4, 0]$, як показано в таблиці 2, де 0

означає, що обидва сигнали сприймаються ідентично, а -4 означає, що відмінності між ними "дуже дратують". Таким чином, чим ближче значення ODG до нуля, тим більша ймовірність того, що сигнали сприймаються як ідентичні.

Таблиця 1.2 – Опис показника ODG

ODG	Опис
0	Непомітний
-1	Відчутний, але не дратівливий
-2	Злегка дратує
-3	Дратує
-4	Дуже дратує

Було досліджено кореляцію між PEAQ та суб'єктивним тестом слуху. Було виявлено, що коефіцієнт кореляції між ними становить 0,837 та 0,851 для базової та розширеної версії PEAQ відповідно. Безумовно, PEAQ не може бути повною заміною суб'єктивного тесту на прослуховування, але він є загальноприйнятим об'єктивним показником якості звуку в індустрії і широко використовується для оцінки непомітності алгоритмів водяних знаків.

Точність алгоритму нанесення водяних знаків визначається як точність виявлення водяного знаку без впливу будь-якої атаки. Її можна виміряти коефіцієнтом бітових помилок (BER) [17], який визначається у формулах:

$$BER(W_1, W_2) = \frac{\sum_{i=1}^N W_1(i) \oplus W_2(i)}{N} \quad (1.3)$$

де W_1, W_2 позначає оригінальну послідовність бітів водяного знаку та виявлену послідовність бітів водяного знаку відповідно, N позначає кількість бітів та i позначає номер біта. У цій статті для оцінки продуктивності використовується "точність", оскільки вона є більш прямолінійною. Вона визначається наступним чином:

$$\begin{aligned} Precision(W_1, W_2) &= \frac{N - \sum_{i=1}^N W_1(i) \oplus W_2(i)}{N} = \\ &= 1 - BER(W_1, W_2) \end{aligned} \quad (1.4)$$

Значення кожної змінної таке ж, як i в рівнянні (1.3). Якщо в експерименті використовується N аудіосигналів, середня точність, позначена як $Precision_{mean}$, обчислюється за формулою (1.5), де i - номер сигналу.

$$Precision_{mean} = \frac{\sum_{i=1}^N Precision_i}{N} \quad (1.5)$$

Стійкість алгоритму нанесення водяних знаків визначається як точність виявлення водяних знаків після атак. Її також можна виміряти за допомогою BER [17]. Аналогічно, точність використовується для оцінки стійкості.

Як правило, для підвищення стійкості в схему включається процес "повторення". При цьому повторюється та сама послідовність бітів водяного знаку. На стороні виявлення для ідентифікації послідовності бітів водяного знаку використовується операція "mode". У статистиці операція "mode" використовується для пошуку даних, які найчастіше зустрічаються у визначеному наборі даних [18]. Наприклад, у наборі даних $\{0, 1, 0, 1, 1\}$ мода - це "1", оскільки вона зустрічається на один раз частіше, ніж "0". Процедуру включення процесу "повторення" у схему водяних знаків можна формалізувати наступним чином:

1 згенерувати бітову послідовність водяного знаку B_w для сигналу, що підлягає нанесенню водяного знаку;

2 на етапі вбудовування вставити V_w у сигнал, скажімо, d разів;

3 на етапі виявлення, виявлена послідовність бітів V_e розбивається на d груп: $V_{e1}, V_{e2}, \dots, V_{ed}$. i -й біт виявленої бітової послідовності водяного знаку V_w визначається як мода бітового набору $\{V_{e1,i}, V_{e2,i}, \dots, V_{ed,i}\}$.

Ємність водяного знаку можна виміряти як кількість біт на секунду (біт/с). Припустимо, що тривалість аудіозапису становить k секунд, а кількість бітів вбудованого водяного знаку - n , тоді пропускна здатність дорівнює біт/с.

Обчислювальну ефективність можна оцінити як процесорний час, необхідний для вбудовування та виявлення водяних знаків. Це залежить від платформи реалізації.

1.3 Огляд алгоритмів нанесення водяних знаків на аудіо

Різноманітність доступних алгоритмів можна розділити на різні категорії відповідно до методології, яку вони використовують. Переважна більшість алгоритмів водяних знаків для аудіо поділяються на три категорії:

- 1 алгоритми на основі часової області;
- 2 алгоритми на основі перетворень;
- 3 гібридні алгоритми.

Алгоритми водяних знаків, що належать до кожної з цих трьох категорій, будуть детально розглянуті далі.

Алгоритми, засновані на часовій області, буквально вбудовують водяний знак у часову область. Вони прості в реалізації. Було розроблено багато алгоритмів на основі часової області [19-21]. Однак алгоритми, що належать до цієї категорії, менш стійкі до атак, і для підвищення стійкості часто використовуються статистичні методи [22]. Розглянемо два основні алгоритми, що належать до цієї категорії, а саме: алгоритм на основі найменш значущого біта (LSB) та алгоритм на основі приховування ехосигналів (Echo hiding).

LSB - одна з найперших методик для нанесення водяних знаків на аудіо, а також для інших типів цифрового контенту. Стандартний підхід полягає у вбудові бітів водяного знаку шляхом зміни значень певних семплів у цифровому аудіо. Біти водяного знаку виявляються шляхом порівняння змінених значень семплів з оригінальними значеннями семплів.

Основною перевагою цього алгоритму є те, що він може досягти надзвичайно високої пропускнуєї здатності. Основним недоліком є його надзвичайно низька стійкість, оскільки випадкові зміни сигналу можуть знищити водяний знак. Дуже мало ймовірно, що вбудовані біти водяного знаку витримають ЦАП і наступні АЦП перетворення. Крім того, зміна значень дискретизації вносить малопотужний адитивний білий гаусівський шум, який робить цей алгоритм менш прозорим для сприйняття, оскільки слухачі дуже чутливі до цього шуму.

Суттєве покращення стандартного алгоритму LSB було запропоновано в статті [22]. Основна ідея полягає в тому, що після вбудовування бітів водяного знаку шляхом маніпуляції з одним бітом 16-бітового зразка wav, всі інші 15 бітів зразка також можуть бути змінені таким чином, щоб різниця між початковим значенням зразка і значенням зразка, яким маніпулювали, була мінімальною. В результаті, це призводить до менших спотворень. Наприклад, якщо оригінальне 16-бітне значення зразка має вигляд "0000000000001000" у двійковому форматі, а біт водяного знаку, який потрібно вбудувати, дорівнює "0". Припустимо, що біт водяного знаку буде вбудовано в 4-й останній біт зразка, і замість того, щоб створити значення "0000000000000000" у двійковому вигляді, як це передбачено стандартним алгоритмом, покращений алгоритм також перевертає перші три біти зразка, створюючи значення "000000000000111" у двійковому вигляді. В результаті різниця між оригінальним зразком і зразком, що піддався маніпуляціям, становить лише 1 у десятковій системі числення, що є найбільш близьким до оригінального значення зразка. Таким чином, внесені спотворення зведено до мінімуму.

Приховування ехосигналів вбудовує біт водяного знаку шляхом введення "ехосигналу". Ехо - це відображення звуку, що доходить до слухача через деякий час після прямого звуку [24]. Використовуються чотири параметри ехо: початкова амплітуда, швидкість спаду амплітуди ехосигналу, "одичний" зсув (час затримки до початкового сигналу) і "нульвий" зсув. Коли зсув між оригіналом і ехо-сигналом зменшується, обидва сигнали зливаються. У певний момент людське вухо чує не оригінальний сигнал і відлуння, а один змішаний сигнал. Точно визначити момент, коли це відбувається, важко. Це залежить від якості оригінального запису, типу звуку, що відлунюється, і слухача. Алгоритм використовує два різних ядра: "одичне" ядро, яке використовується для генерування ехосигналу з "одичним" зсувом, що відповідає двійковій "1", і "нульве" ядро, яке використовується для генерування ехосигналу з "нульвим" зсувом, що відповідає двійковому "0" [25].

Алгоритми на основі перетворень зазвичай вбудовують біти водяних знаків, використовуючи властивості даних у представленні після перетворень. Популярними перетвореннями є швидке перетворення Фур'є (ШПФ), дискретне косинусне перетворення (ДКП) і дискретне вейвлет-перетворення (DWT) [26-27]. Деякі методи, такі як модуляція індексу квантування (QIM), розкладання за сингулярним значенням (SVD) та інтерполяція, часто використовуються для маніпулювання даними з метою вбудовування бітів водяного знаку в представлення після перетворень. Багато алгоритмів накладання водяних знаків належать до цієї категорії, оскільки вбудовані біти водяних знаків є більш стійкими до атак.

Модель HAS, як правило, використовується для мінімізації введених спотворень сприйняття. Однак тут існує певний компроміс, оскільки вбудовування бітів водяного знаку в значущі для сприйняття компоненти є більш стійким, але менш прозорим для сприйняття. З іншого боку, вбудовування бітів водяних знаків у менш значущі для сприйняття компоненти є менш стійким, але більш прозорим для сприйняття. Крім того,

використання HAS-моделі збільшує час обчислень, що обмежує використання цих алгоритмів у критичних до часу додатках. Далі буде розглянуто типовий алгоритми цієї категорії.

ШПФ було розроблено як швидку версію дискретного перетворення Фур'є (ДПФ). ДПФ є добре відомим і потужним обчислювальним інструментом для виконання частотного аналізу дискретних часових сигналів. Він бере дискретний сигнал в часовій області і перетворює цей сигнал в дискретну частотну область. Було запропоновано безліч алгоритмів нанесення водяних знаків, які базуються на маніпулюванні компонентами, що містяться в спектрі ШПФ. Більшість алгоритмів маніпулюють величиною компонентів ШПФ і підвищують стійкість до типових систем стиснення звуку шляхом включення моделі HAS.

Схема, запропонована в [27], вибирає набір частот, порівнюючи спектр ШПФ вихідного сигналу зі спектром відповідного стисненого декомпресованого сигналу. Біти водяного знаку вбудовуються на тих частотах, які мають схожі величини в обох спектрах. Однак такий вибір призводить до збурення вихідного сигналу на найбільш значущих частотах, що є небажаним з точки зору прозорості сприйняття.

Схема, запропонована в [28], вносить деяку випадковість у процес вибору частот, що дозволяє поліпшити прозорість схеми ціною втрати деякої стійкості. Всі ці схеми не є сліпими, тобто для виявлення вбудованих бітів водяного знаку потрібен спектр вихідного сигналу.

Алгоритм, запропонований в [29], вбудовує біти водяного знаку на основі сплайн-інтерполяції даних, отриманих в результаті перетворення ШПФ. Процес вбудовування зображено на рисунку 1.

Як видно з рисунка, ШПФ-аналіз застосовується до кожного кадру (тобто короткого сегменту) вихідного сигналу для отримання величин непарних бітів. Потім інтерпольовані значення парних сегментів отримуються шляхом сплайн-інтерполяції значень непарних сегментів. Біти водяного знаку вбудовуються шляхом маніпулювання цими сплайн-

інтерпольованими значеннями парних сегментів. Нарешті, сигнал з водяним знаком відновлюється за допомогою зворотного БПФ.

Як видно, з рисунка 1.1, до сигналу з водяним знаком застосовується ШПФ-аналіз для отримання значень непарних і парних сегментів на основі кожного кадру. Потім використовується сплайн-інтерполяція для отримання інтерпольованих значень парних сегментів. Ці інтерпольовані значення парних бітів порівнюються з отриманими за допомогою ШПФ значеннями парних бітів для виявлення бітів водяного знаку.



Рисунок 1.1 – Процес вбудовування алгоритму, запропонованого в статті [29]

Процес виявлення ЦВЗ представлено на рисунку 1.2.

Цей алгоритм досяг високої пропускної здатності близько 3000 біт/с і є стійким до більшості атак. Середній показник ODG становить -0,5, що є прийнятним. Обчислювальна ефективність цього алгоритму висока, оскільки задіяні лише процеси інтерполяції, ШПФ та зворотного ШПФ. Недоліком цього алгоритму є те, що вбудовані біти водяного знаку легко видалити, оскільки позиція вбудовування відома. Крім того, оскільки цей алгоритм базується лише на порівнянні величин, які легко порушити, алгоритм буде

вразливим до певних атак. Нарешті, оскільки тест базувався лише на п'яти піснях з одного альбому, оцінка була обмеженою.

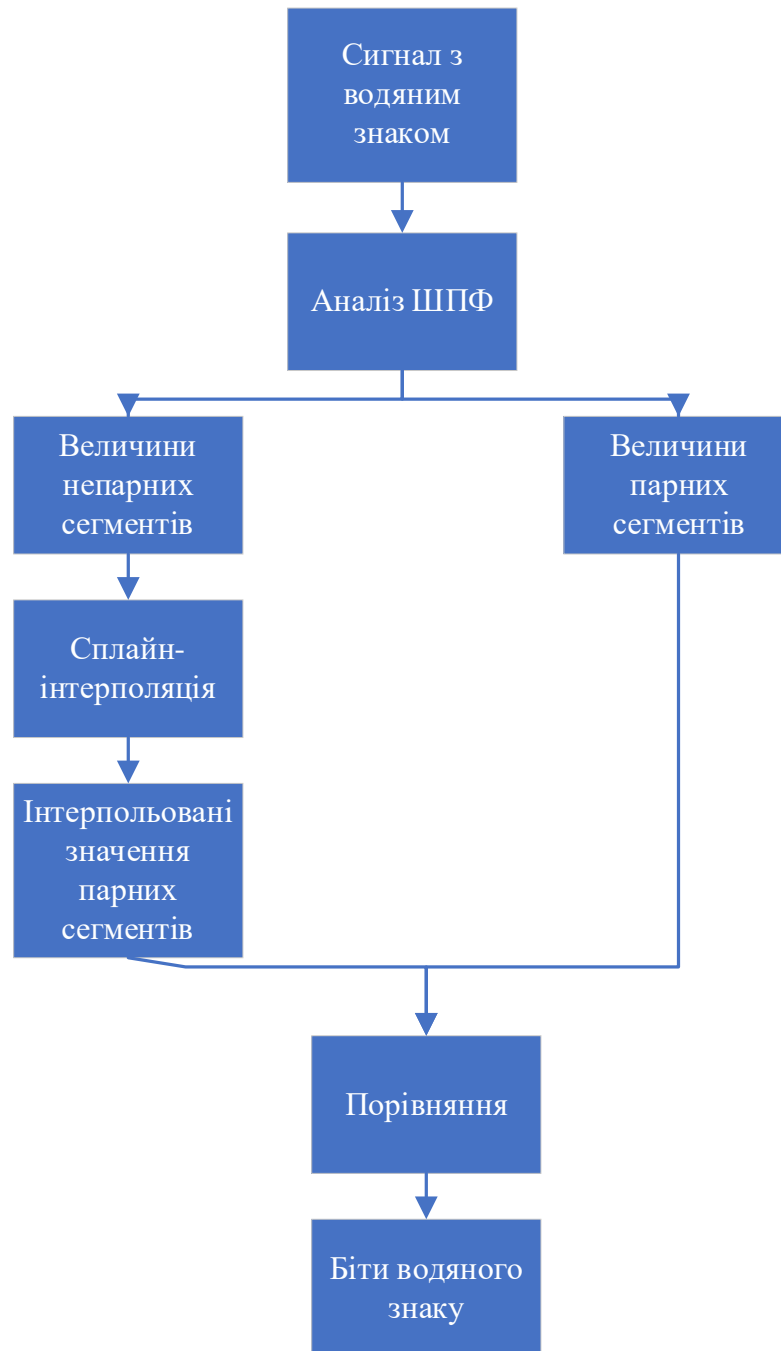


Рисунок 1.2 – Процес виявлення водяного знаку, запропонованого в статті [29]

Гібридні алгоритми - це такі нові алгоритми, як алгоритм на основі Chirp-кодування [30], алгоритм "клаптикового шифрування" [31] та алгоритм

на основі SVD [32], які не можуть бути просто віднесені до жодної з двох вищезгаданих категорій. Особлива причина виділення цієї категорії полягає в тому, щоб підкреслити їхню новизну. Приклад одного з таких алгоритмів буде детально розглянуті далі.

У роботі [33] було запропоновано алгоритм крихких водяних знаків. Процес вбудовування можна описати наступним чином: спочатку виконується 7-рівневе вейвлет-розкладання сигналу для отримання 7 рівнів коефіцієнтів "деталізації". Для того, щоб виміряти глобальний ефект від нанесення водяного знаку в сигнал, використовуються також коефіцієнти "апроксимації" на 7-му рівні. Таким чином, всього генерується 8 векторів розкладу. Причина використання коефіцієнтів "деталізації" полягає в тому, що вони дуже чутливі до таких атак, як стиснення з втратами та нарізка аудіо. Потім обчислюється відсоток енергії кожного вектора від загальної енергії 8 векторів. Ці відсотки округлюються до найближчого цілого числа і перетворюються в двійковий потік, який буде використано як послідовність бітів водяного знаку. Потім створюється функція Chirp. Потім ця функція множиться на новий сигнал, який утворюється на основі двійкової послідовності, і масштабується заздалегідь визначеним масштабним коефіцієнтом для отримання Chirp-коду. Цей Chirp-код додається до вихідного сигналу для створення сигналу з водяним знаком. Для того, щоб зробити водяний знак нечутним, згенерований Chirp-код має дуже низьку частоту та амплітуду.

На стороні виявлення застосовується та сама Chirp-функція, що й на стороні вбудовування, яка співвідноситься з сигналом водяного знаку. Таким чином, біти водяного знаку можуть бути відновлені. Потім, чи був сигнал підроблений чи ні, можна перевірити, порівнюючи відновлену послідовність бітів водяного знаку з можливо підробленим двійковим потоком, який можна згенерувати безпосередньо з сигналу з водяним знаком. Що стосується цього алгоритму, то вбудовані водяні знаки важко видалити з хоста, оскільки початкова і кінцева частота Chirp-функції задаються на розсуд користувача, а

її позиція в потоці даних може змінюватися за допомогою зсуву, причому всі ці параметри в сукупності утворюють приватний ключ [32]. Тести прослуховування показали, що немає ніякої перцептивної різниці між оригінальним сигналом і сигналом з водяними знаками. Відновлення Chirp - коду є унікально стійким у випадку дуже низького SNR, тому цей алгоритм легко адаптувати як стійкий алгоритм нанесення водяних знаків.

Отже було розглянуто різноманітні популярні алгоритми, розроблені для нанесення водяних знаків на аудіофайли. Для того, щоб порівняти всі ці алгоритми, непомітність оцінюється однаково за показником MOS, а показник ODG можна безпосередньо зіставити з показником MOS, як показано в таблиці 3.

В таблицях 1.3-1.5 наведено для узагальнення результатів роботи типових алгоритмів, які були розглянуті. Алгоритм аудіо водяних знаків на основі кодування Chirp не вказано в таблицях [32], оскільки він спочатку був розроблений як крихкий водяний знак. Однак він має великий потенціал для розвитку в якості надійного алгоритму водяного маркування. У таблиці 4 наведено чотири основні характеристики кожного алгоритму водяних знаків: непомітність, оцінена за MOS, стійкість, пропускну здатність та обчислювальна ефективність. У таблиці 1.5 наведено деякі інші характеристики алгоритмів.

Таблиця 1.3 – Співвідношення між MOS та ODG

ODG	MOS
0	5
-1	4
-2	3
-3	2

Таблиця 1.4 – Чотири основні характеристики кожного типового алгоритму нанесення водяних знаків на аудіо

	LSB	Echo hiding	FFT
Непомітність	5	5	4.5
Надійність	низька	низька	висока
Ємність	44100	n/a	3000
Ефективність	висока	n/a	висока

Таблиця 1.5 – Інші характеристики кожного алгоритмів накладання водяних знаків на аудіо

	LSB	Echo hiding	FFT
Невидимість.	+	-	+
Додаткова обробка	-	-	-
Можливість зняття	легко	важко	легко
Надійність	-	+	+

З таблиць 1.4 та 1.5 видно, що різні алгоритми мають різні сильні та слабкі сторони.

У розділі було розглянуто різні підходи до захисту аудіофайлів за допомогою цифрових водяних знаків. Існує велика кількість методів нанесення цифрових водяних знаків, які можна класифікувати за різними критеріями: за типом сигналу, методом вбудовування, місцем вбудовування, стійкістю до атак та іншими параметрами. Найбільш важливими характеристиками водяних знаків є їх стійкість до атак (зокрема, до змін та маніпуляцій з аудіофайлами) та непомітність для слухача. Це означає, що водяний знак не повинен впливати на якість звуку, але при цьому повинен бути досить міцним, щоб витримати різні види обробки аудіо.

Методи, що використовують частотну або часово-частотну область для

нанесення водяних знаків, часто є більш стійкими до атак та змін у порівнянні з методами, що працюють у часовій області. Широко використовуються методи на основі перетворень, такі як дискретне косинусне перетворення (DCT) та інші. Ці методи дозволяють вбудовувати водяні знаки таким чином, що вони стають менш вразливими до атак. Адаптивні методи, що враховують властивості аудіофайлу при вбудовуванні водяного знака, можуть значно покращити як стійкість, так і непомітність водяних знаків.

Для оцінки ефективності методів нанесення водяних знаків використовуються різні метрики, такі як співвідношення сигнал/шум (SNR), ймовірність виявлення, та інші. Експериментальні дослідження показують, що комбіноване використання різних методів може забезпечити кращі результати.

Вибір конкретного методу залежить від конкретних вимог та умов використання, таких як необхідний рівень захисту, допустимі зміни якості звуку та очікувані типи атак.

2 СПЕКТРАЛЬНИЙ АНАЛІЗ СИГНАЛІВ

У цій главі описано алгоритм накладання водяних знаків на аудіо на основі CSPE, який є інструментом спектрального аналізу з надвисокою роздільною здатністю. Цей алгоритм накладання водяних знаків базується на перетворенні сигналу в частотній області. Спочатку пояснюється вдосконалення оригінального CSPE. Потім запропоновано алгоритм водяних знаків, заснований на цьому вдосконаленому CSPE, і проведено експерименти для демонстрації ефективності цього алгоритму водяних знаків на синтетичних сигналах. Потім запропонований алгоритм застосовано до реальних аудіосигналів, виявлено та вирішено проблеми, пов'язані з цим. Розділ завершується оцінкою перцептивної прозорості та робастності покращеного алгоритму накладання водяних знаків на реальні аудіосигнали.

2.1 Комплексна еволюція спектральної фази

Оцінка частот сигналу, який складається з синусоїдальних складових, часто виконується в частотній області з використанням пік-пкінгу зі спектра амплітуд сигналу, де спектр майже завжди обчислюється за допомогою добре відомого алгоритму ШПФ. Однак точність цього алгоритму сильно обмежена у випадках, коли частота компонента не є кратною довжині віконного сигналу, поділеній на частоту дискретизації. По суті, це означає, що тільки тоді, коли частота компонента точно збігається з частотами аналізу ШПФ, вона може бути точно виміряна. Враховуючи велику кількість частотних компонентів, які створюють реалістичні звуки, це є серйозним обмеженням при спробі точного аналізу сигналу. Коли частотний компонент не збігається точно з частотами аналізу ШПФ, загальним рішенням, яке використовується в алгоритмах синусоїдального моделювання, є застосування квадратичної інтерполяції до спектральних амплітуд компонентів безпосередньо по обидва

боки від істинної частоти для знаходження правильних значень частоти і амплітуди. Однак продуктивність цього методу сильно залежить від використовуваної віконної функції і довжини даних для аналізу. CSPE.

CSPE було представлено в роботі [20] як метод точної оцінки частоти компонентів, що існують у сигналі. Він також був розроблений, щоб бути обчислювально ефективним. У деяких аспектах він пов'язаний з технікою перехресних спектрограм і технікою перепризначеного спектра.

Процедуру CSPE можна зобразити на рисунку 2.1, де $x_0(n)$ - це вихідний сигнал, що аналізується, а $x_1(n)$ - це зсунута на один відлік версія $x_0(n)$. Як видно з рисунка 2.1, CSPE працює наступним чином: ШПФ-аналіз виконується двічі. Спочатку для сигналу $x_0(n)$, а потім для $x_1(n)$. Потім, шляхом множення спектру ШПФ $x_0(n)$ на комплексно спряжений спектр ШПФ $x_1(n)$, формується частотно-залежна функція, з якої можна визначити точні значення частотних компонентів, що містяться в ній, шляхом вилучення інформації про фазу кута. На графіку ця частотно-залежна функція має схожий на сходи вигляд, де плоскі частини графіка вказують на точні частоти компонентів.

2.2 Математичне представлення комплексної еволюція спектральної фази

Математично алгоритм можна описати наступним чином: розглянемо реальний синусоїдальний сигнал з частотою в Гц $\beta = q + \delta$, де q - ціле число, а δ - дробове, з амплітудою a і початковою фазою b . w - віконна функція, що використовується у ШПФ, N - довжина вікна, F_{wx0} - віконне перетворення Фур'є $x_0(n)$, а F_{wx1} - віконне перетворення Фур'є $x_1(n)$. Тоді, записавши спочатку

$$D = e^{\frac{j2\pi\beta}{N}} \quad (2.1)$$

Залежна від частоти функція CSPE згодом може бути записана як:

$$CSPE_w = F_{wx_0} F_{wx_1}^* = \left(\frac{a}{2}\right)^2 [D^* \|F_w(D^n)\|^2 + 2\text{Re}\{e^{j2b} DF_w(D^n) \odot F_w^*(D^{-n})\} + D \|F_w(D^{-n})\|^2] \quad (2.2)$$

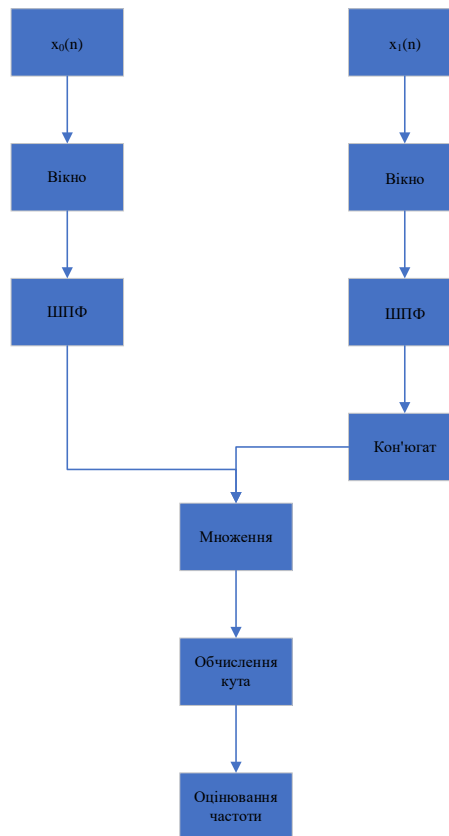


Рисунок 2.1 – Процедура CSPE

де \odot позначає поелементний добуток, а $*$ позначає операцію сполучення. Перший і третій доданки у квадратних дужках - це комплексна синусоїда додатної частоти і комплексна синусоїда від'ємної частоти відповідно. Середній доданок представляє доданки само-взаємодії від інтерференції між додатною та від'ємною частотними складовими.

Віконне перетворення Фур'є вимагає множення сигналу на вікно аналізу в часовій області, і, таким чином, результуюче перетворення є згортокою перетворення віконної функції, W , з перетворенням комплексної

синусоїди. Оскільки перетворення синусоїди є нічим іншим, як парою дельта-функцій в додатній і від'ємній частотних позиціях, результатом згортки є просто частотна копія W з центрами в точках $+\beta$ і $-\beta$. Отже, зі стандартною віконною функцією лише при $k \approx \beta$, де k - номер частотного відсіку, член $\|F_w(D^n)\|^2$ є значним, і він швидко спадає, коли k далекий від β . Тому вікно аналізу слід вибирати ретельно, щоб воно швидко спадало, щоб мінімізувати будь-який спектральний витік у сусідні відсіки. Якщо це так, то це зробить інтерференційні члени, тобто середній член, у рівнянні (2.2) несуттєвим. Таким чином, для додатніх частот CSPE залежить від частоти:

$$CSPE_w \approx \frac{a^2}{4} \|F_w(D^n)\|^2 D^{-1} \quad (2.3)$$

Знаходження інформації про кут у рівнянні (2.3) призводить до оцінки частоти

$$f = \frac{-N \angle(CSPE_w)}{2\pi} = \frac{-N(\frac{2\pi}{N} \beta)}{2\pi} \quad (2.4)$$

Таким чином, неціле значення частоти f отримуємо за допомогою рівняння (2.4). Оцінку фази та амплітуди можна отримати за допомогою рівнянь (2.5) та (2.6) відповідно.

$$\phi = \angle \frac{2F_{wx_0}(k)}{W_{(k-f)}} \quad (2.5)$$

$$m = \left\| \frac{2F_{wx_0}(k)}{W_{(k-f)}} \right\| \quad (2.6)$$

Оцінка частоти, фази і амплітуди за допомогою CSPE, як показано

вище, також може бути застосована до сигналів, що містять більше однієї частотної складової.

2.3 Демонстрація роботи CSPE

Приклад вихідного сигналу CSPE показано на рисунку 2.2, де вісь x позначає номер діапазону, а вісь y - оцінене значення частоти. Розглянемо приклад сигналу $x_a(n)$, який містить компоненти з частотами 17 Гц, 293,5 Гц, 313,9 Гц, 204,6 Гц, 153,7 Гц, 378 Гц і 423 Гц відповідно. Частота дискретизації становить 1024 Гц. Кадр довжиною 1024 відліки розбивається на вікна за допомогою вікна Блекмана і заповнюється 1024 нулями. Частота обчислюється за формулою (2.4). Як показано на рисунку 2.2, ці частоти можуть бути точно визначені і позначені стрілками на графіку на рисунку 2.2. Їхні частоти можна знайти, відкладаючи значення по осі абсцис. Найбільша похибка серед усіх оцінок частот компонентів становить лише приблизно 0,15 Гц.

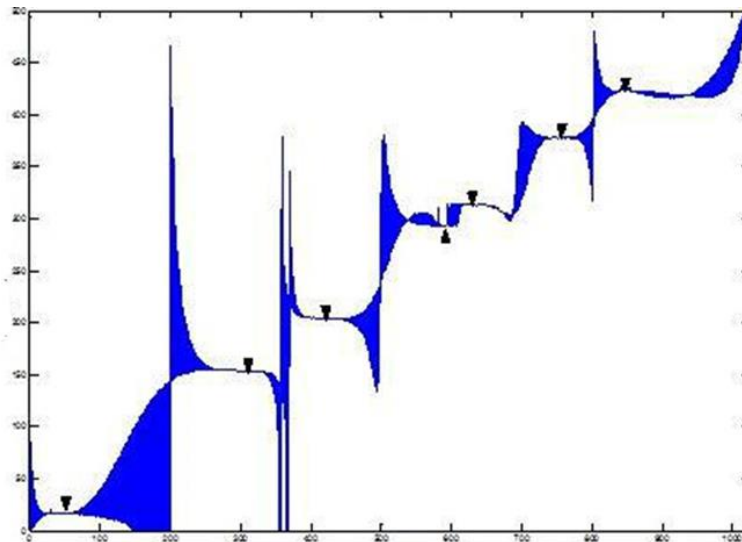


Рисунок 2.2 – Частотне оцінювання багатокompонентного сигналу алгоритмом CSPE

2.4 Швидке перетворення Фур'є

Як має бути представлене аудіо під час обробки? Сире аудіо семплюється з високою частотою, зазвичай 44100 Гц для музики. Це призводить до дуже високої розмірності даних і залежностей між відліками, що простягаються далеко один від одного.

Альтернативою для представлення аудіо є використання короткого часового перетворення Фур'є (Short Time Fourier Transform, STFT). Виділяються сегменти хвильового файлу, і для кожного сегмента обчислюється відповідне перетворення Фур'є. Сегменти об'єднуються у двовимірне представлення звуку з часом на одній осі та частотами на іншій осі. Комплексне представлення часто виражається як амплітуда і фаза, де амплітуда також називається спектрограмою звуку. Значення на спектрограмі представляють енергію сигналу у відповідному частотному діапазоні, в кожному сегменті. Розмір сегментів задається розміром вікна, а відстань між початком кожного сегмента - довжиною стрибка. Зазвичай довжина стрибка менша за розмір вікна, так що вікна перекривають одне одного. На Рисунку 2.3 відрізки відбираються з 50% перекриттям.

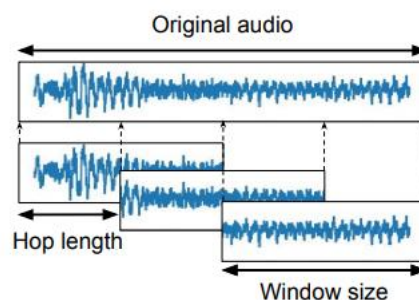


Рисунок 2.3 – Сегменти відбираються з форми сигналу з 50% перекриттям

З STFT можна без втрат відновити форму сигналу за допомогою зворотного STFT. Це дозволяє обробляти аудіо в STFT, використовуючи такі переваги, як спектральне представлення та зменшення часових залежностей,

а потім реконструювати його до форми, придатної для відтворення. Рисунок 2.4 ілюструє такий конвер.

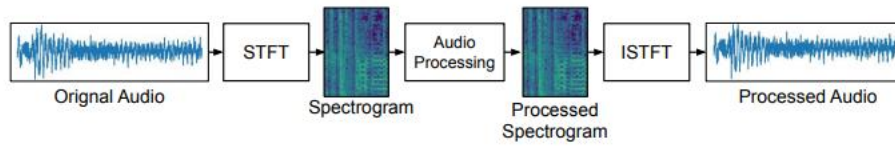


Рисунок 2.4 – Конвер для обробки звуку за допомогою STFT

3 НАНЕСЕННЯ ВОДЯНИХ ЗНАКІВ НА АУДІО

3.1 Запропонована схема нанесення водяних знаків

Запропонований підхід є комбінацією матричного кодування та LSB-відповідності (LSBM). Матричне кодування дозволить забезпечити включення секретного повідомлення з мінімальною кількістю змін рис. 3.1. LSB-відповідність пом'якшить ефект парних чисел (PoV) методу LSB-заміни (LSB). Запропоновані схеми приховування даних базуються на різниці значень коефіцієнтів. Значення коефіцієнтів порівнюються між собою відповідно до використовуваного варіанту. Водяний знак, що вставляється, являє собою послідовність 0 і 1; ці дані замінюються на LSB значень коефіцієнтів в залежності від використовуваного варіанту. Енергія сигналу сконцентрована в низькочастотних складових. Вставка мітки в цих частотах забезпечує хорошу робастність, але вносить видимі спотворення в просторово-часовій області. З іншого боку, вставка мітки у високочастотні компоненти не погіршує якість сигналу, але робить мітку вразливою до атак. Тому середньочастотний діапазон є найбільш придатним для інтеграції водяних знаків, оскільки він відповідає вимозі компромісу між стійкістю та непомітністю. Процес вбудовування рис. 3.1. складається з п'яти кроків:

- 1 оригінальний звук спочатку перетворюється БПФ;
- 2 виділяється дійсна частина високочастотних компонентів;
- 3 парність коефіцієнтів модулюється відповідно до бітів водяного знаку;
- 4 помічені реальні частини поєднуються з уявними частинами;
- 5 зворотне перетворення виконується для створення аудіофайлу з водяним знаком.

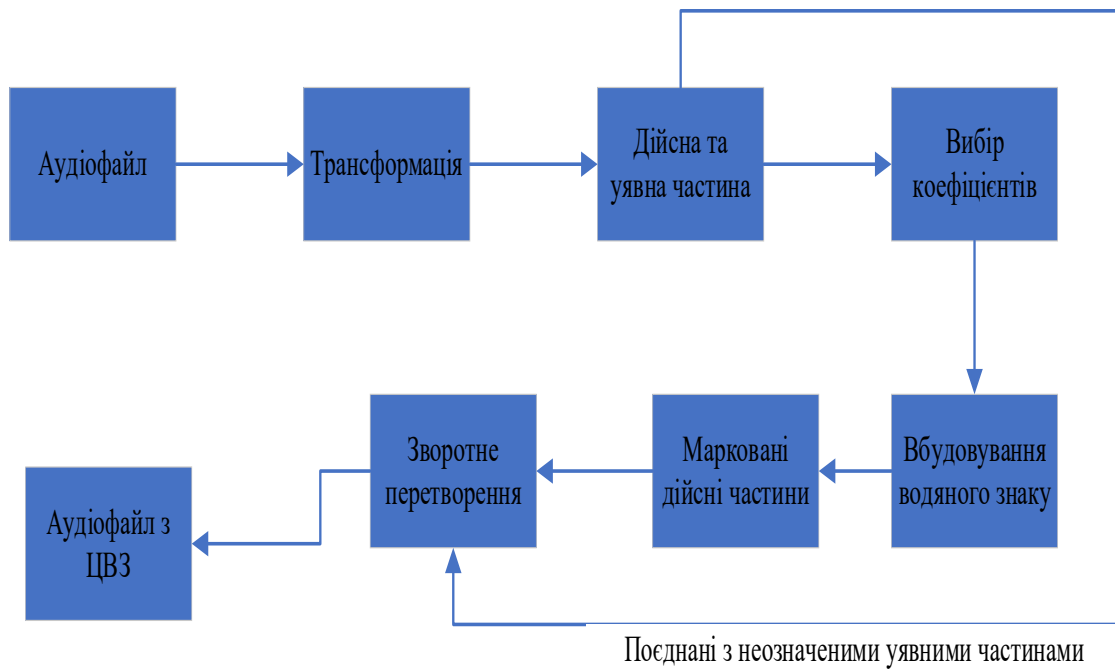


Рисунок 3.1 – Процес нанесення водяних знаків

Таким чином, водяний знак інтегрується в дійсну частину цих коефіцієнтів відповідно до правил таблиці 3.1 або таблиці 3.2, залежно від використовуваного варіанту. У першому варіанті для приховування двох бітів водяного знаку використовуються три коефіцієнти; обчислюється парність трьох послідовних коефіцієнтів, які потім порівнюються з бітами, що підлягають приховуванню. У випадку нерівності, парність одного з коефіцієнтів змінюється, щоб задовольнити рівність згідно з правилами, наведеними в таблиці 3.1. Розглянемо для цього $C1$, $C2$ і $C3$ - три послідовні коефіцієнти, а X , Y - два біти повідомлення, які потрібно вставити.

Таблиця 3.1 – Зміна правил для першого варіанту

Умова	Дія
$X = (C1 - C2) \% 2 \ \& \ Y = (C1 - C3) \% 2$	Зміни не потрібні
$X \neq (C1 - C2) \% 2 \ \& \ Y = (C1 - C3) \% 2$	Зміна $C2$
$X = (C1 - C2) \% 2 \ \& \ Y \neq (C1 - C3) \% 2$	Зміна $C3$
$X \neq (C1 - C2) \% 2 \ \& \ Y \neq (C1 - C3) \% 2$	Зміна $C1$

Таблиця 3.2 – Зміна правил для другого варіанту

Умова	Дія
$X = (C1 - C2) \% 2$	Зміни не потрібні
$X \neq (C1 - C2) \% 2$	Зміна C1

У другому запропонованому варіанті для приховування одного біта водяного знаку використовуються два коефіцієнти, обчислюється парність двох послідовних коефіцієнтів, а заміна одного біта виконується за одним з правил, наведених у таблиці 3.2. Вважатимемо, що C1 і C2 - це два послідовних коефіцієнти, а X - біт повідомлення, який потрібно вставити.

У першому варіанті інтегруюча здатність відповідатиме 2/3 коефіцієнтів, що вбудовуються, у другому варіанті - половині коефіцієнтів, що вбудовуються. Нарешті, до трьох компонентів буде застосовано обернене перетворення, щоб сформувати зображення з водяним знаком.

3.2 Експериментальні результати та обговорення

Для того, щоб оцінити процес інтеграції, застосовано різні перетворення Фур'є до аудіофайлів нашої бази даних, для кожного результату перетворення отримані коефіцієнти позначені двома запропонованими варіантами. Водяний знак, який використовувався в наших експериментах, є випадковою послідовністю з 4500 біт. У першому експерименті до аудіофайлу застосовується ДПФ, який генерує в середньому 13 503 коефіцієнти для вбудови. Перший варіант вимагає трьох коефіцієнтів для вбудовування двох бітів, тому можна вбудувати 9002 біти. Другий варіант вимагає двох коефіцієнтів для вбудовування одного біта, і матиме ємність 6751 біт. На рис. 3.2 показано амплітуди вихідного сигналу та сигналу з водяним знаком після модифікації коефіцієнтів ДПФ.

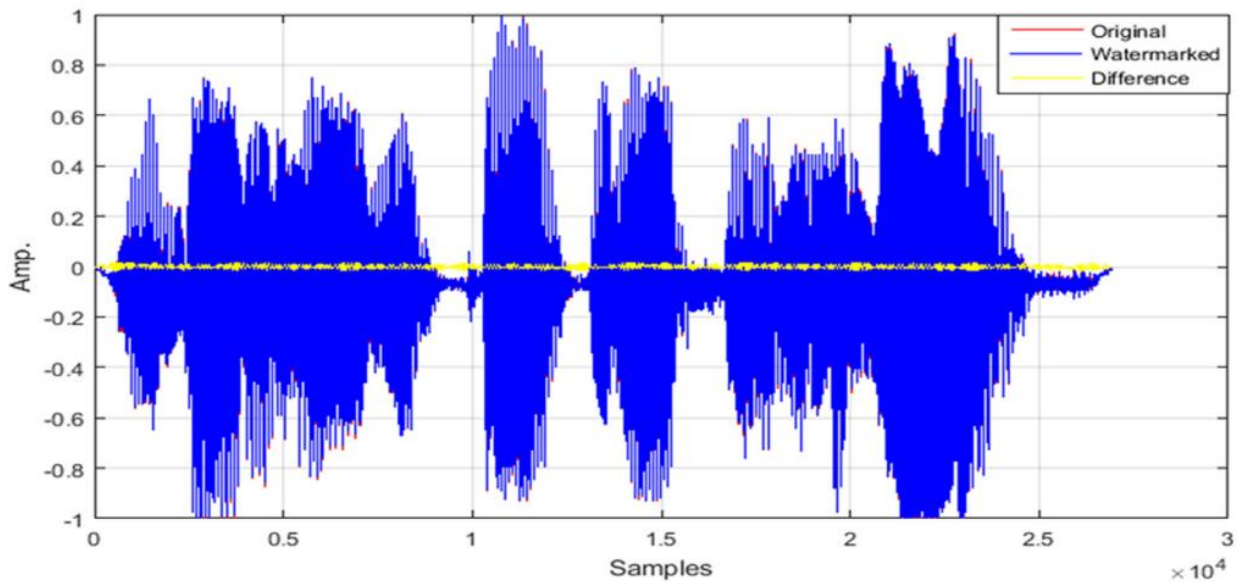


Рисунок 3.2 – Амплітудне представлення коефіцієнтів ДПФ з водяними знаками

Для нашого другого експерименту два варіанти запропонованого процесу нанесення водяних знаків застосовуються до коефіцієнта, отриманого в результаті застосування ШПФ до аудіофайлу. Це генерує в середньому 26 906 коефіцієнтів на площину, тому матимемо ємність 17 937 біт для першого варіанту і 13 453 біт для другого варіанту. На рис. 3.3 показано амплітуди вихідного сигналу та сигналу з водяним знаком, модифікованого коефіцієнтом ШПФ.

У третьому експерименті застосування QDFT до аудіофайлу дозволяє отримати 13 427 коефіцієнтів. Оскільки перший варіант вимагає трьох коефіцієнтів для вбудовування двох бітів, то таким чином можна буде інтегрувати 8951 біт. Другий варіант вимагає двох коефіцієнтів для вбудовування одного біта, і матиме ємність 6713 біт. На рис. 3.4 показано амплітуди вихідного сигналу та сигналу з водяним знаком після модифікації коефіцієнтів QDFT.

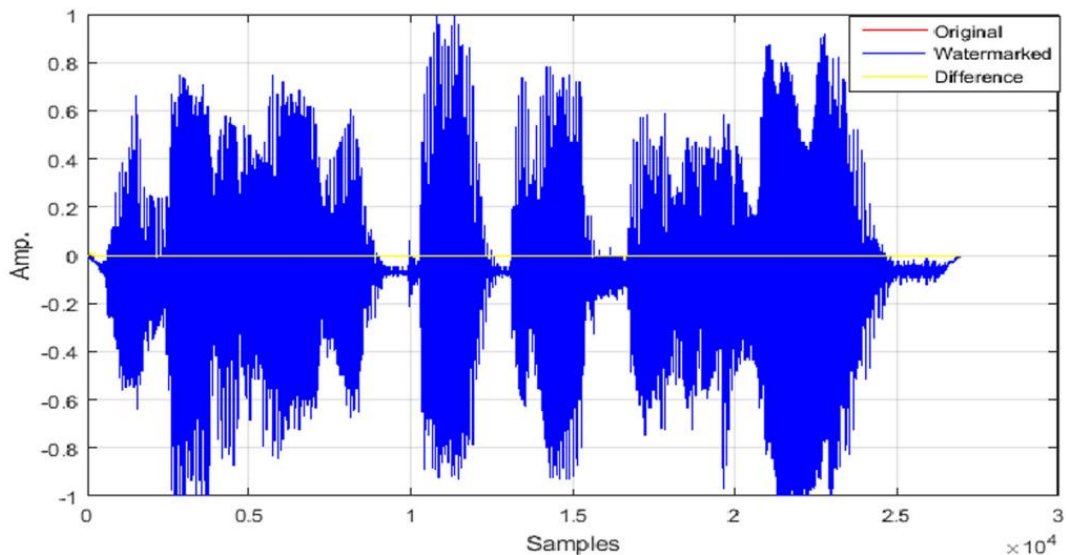


Рисунок 3.3 – Амплітудне представлення коефіцієнтів ШПФ з водяними знаками

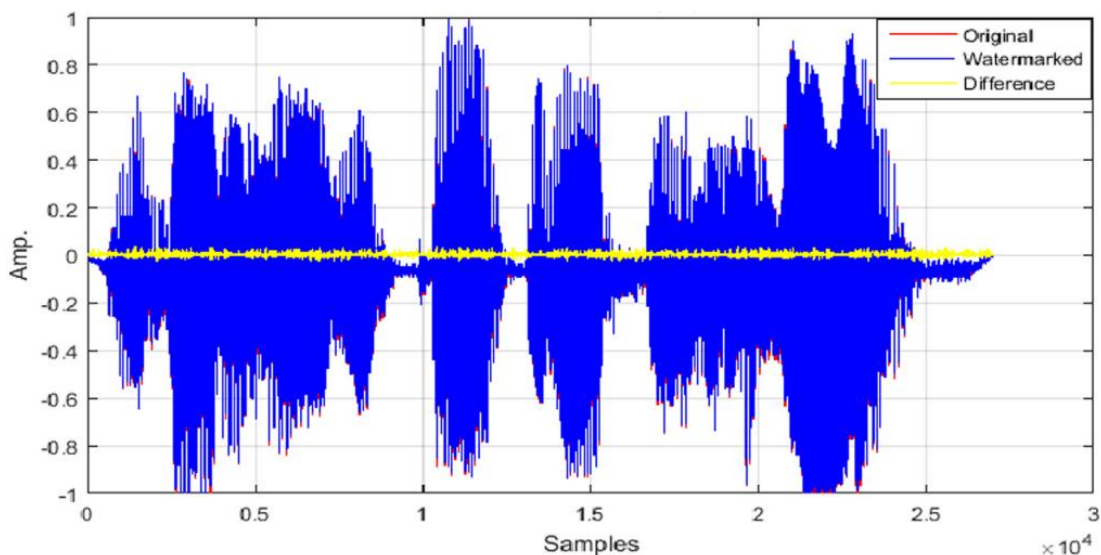


Рисунок 3.4 – Амплітудне представлення коефіцієнтів QDFT з водяними знаками

Як видно з рис. 3.2-3.5, отримані результати не дозволяють візуально виявити спотворення, спричинені процесом приховування. З цієї причини вимірювання схожості та спотворень були розраховані шляхом порівняння оригінального аудіофайлу та аудіофайлу з водяними знаками (табл. 3.3). Для вимірювання та тестування ефективності запропонованих підходів використовується показник SNR.

SNR - це метрика, яка вимірює кількість шуму, доданого до аудіофайлу

в процесі накладання водяного знаку [33]. Високе значення SNR є бажаною характеристикою і означає, що рівень спотворень є низьким; SNR є найбільш часто використовуваною оцінкою. Для кращої оцінки кожного алгоритму також розраховували середню квадратичну похибку (MSE), коефіцієнт кореляції (CORR), дисперсію (VAR) та середньоквадратичне відхилення (STD). У цій таблиці DFTV1 і DFT-V2 представляють результати нанесення водяних знаків, отримані шляхом інтегрування в коефіцієнти дискретного перетворення Фур'є, FRFT-V1 і FRFT-V2 представляють результати, отримані шляхом застосування дробового перетворення Фур'є. QDFT-V1 і QDFT-V2 відповідають результатам нанесення водяних знаків із застосуванням кватерніонного дискретного перетворення Фур'є (V1 і V2 відповідають першому і другому варіанту відповідно).

Таблиця 3.3 – Результати непомітності

	SNR	MSE	SSIM	CORR	VAR	STD
DFT-V1	35,78	6,66	0.7308	0,9996	0,571	0,742
DFT-V2	37,97	9,6	0.8235	0,9998	0,784	0,886
FRFT-V1	35,79	6,67	0.7829	0,9999	0,572	0,743
FRFT-V2	37,98	9,61	0.8369	0,9999	0,785	0,886
QDFT-V1	37,95	9,54	0.8242	0,9995	0,783	0,885
QDFT-V2	35,93	6,68	0.8038	0,9997	0,579	0,748

Для оцінки стійкості запропонованих схем (для кожного перетворення) було виконано декілька атак на зразки з водяними знаками, а отримані результати після вилучення водяного знаку порівняно з відповідними роботами. Нормована кореляція (НК) застосовується для оцінки стійкості схеми водяного знаку шляхом порівняння оригіналу та вилученого водяного знаку після атаки на зразки з водяними знаками. Зазвичай, $NC > 0.85$ означає значну схожість між оригінальним та вилученим водяним знаком.

Як видно з таблиці 3.4, запропоновані схеми генерують досить стійкі

до різних атак зразки з високоякісним водяним знаком. Зазвичай, у просторовій області, якщо LSB модифікуються, вилучення водяного знаку стає неможливим. Однак у запропонованих схемах ми можемо відновити вбудований водяний знак навіть після спотворення бітів LSB. Це пояснюється незначною модифікацією, що виконується, у запропонованих підходах для інтеграції водяного знаку комбінується парність послідовних елементів, що зменшує ймовірність модифікації. Таким чином, з отриманих результатів можна зробити висновок, що наші підходи є стійкими до декількох зловмисних атак видалення водяних знаків, гарантуючи при цьому прийнятне значення SNR.

Таблиця 3.3 – Порівняння надійності

	DFT- V1	DFT- V2	FRFT- V1	FRFT- V2	QDFT- V1	QDFT- V2
Гауссів шум	0.91800	0.90543	0.90833	0.90542	0.92871	0.9109
Низькочастотна фільтрація	0.92237	0.91735	0.96345	0.90213	0.93057	0.9446
Re-sampling	0.96838	0.96082	0.98265	0.98814	0.94325	0.9873
Переквантифікація	0.97752	0.97639	0.96131	0.97513	0.98521	0.9614

ВИСНОВКИ

У роботі запропоновано два нових підходи для нанесення водяних знаків на аудіо в частотній області: після обчислення ДПФ виконується заміна бітів, що підлягають інтегруванню, шляхом модифікації LSB квантувальних коефіцієнтів, порівнюючи парність послідовних коефіцієнтів. Комбінація між коефіцієнтами дозволила нам приховати два біти, змінивши лише одне значення. Інтегрування водяного знаку в старші коефіцієнти може призвести до сильного спотворення зображення, тоді як інтегрування в молодші коефіцієнти робить водяний знак стійким до стиснення та фільтрації. Для цього водяний знак був інтегрований з використанням середньосмугових коефіцієнтів (змішані часові та частотні компоненти сигналу), що дозволило нам гарантувати непомітність і одночасно стійкість до атак стиснення та фільтрації.

Використання перетворень робить повідомлення більш стійким до стиснення, оскільки воно використовує той самий простір, що і для кодування. Крім того, ці схеми можуть забезпечити більшу стійкість до атак фільтрації, оскільки коефіцієнти водяних знаків розподілені по всьому аудіофайлу. Модифікація якомога меншої кількості значень дозволила нам згенерувати меншу кількість спотворень, що дає нам прийнятні результати SNR (більше 36 дБ).

Порівнявши підходи з нещодавніми роботами в трьох областях вставки, можна зробити висновок, що наш другий варіант пропонує хороші результати непомітності для вставки в коефіцієнти DFT і FrFT. А перший підхід більше підходить для інтеграції в коефіцієнти QDFT. Отримані результати залишаються задовільними порівняно з деякими нещодавніми роботами, в яких SNR в середньому перевищує 36,90 дБ.

Однак, використовуючи невеликі аудіофайли (що зменшує кількість коефіцієнтів) і оскільки наш підхід вимагає трьох коефіцієнтів для вставки

двох бітів, пропускна здатність наших підходів дуже обмежена в частотній області (між 9002 і 17 937 бітами на файл).

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Corporate Kaijus Clash: 15 Famous Copyright Infringement Cases [Електронний ресурс]. – Режим доступу: <https://www.abounaja.com/blogs/copyright-infringement-cases>
2. Zhou, N.R., Hou, W.M.X., Wen, R.H. et al. Imperceptible digital watermarking scheme in multiple transform domains. *Multimed Tools Appl* 77, 30251–30267 (2018). <https://doi.org/10.1007/s11042-018-6128-9>
3. Hosny, Khalid M., and Mohamed M. Darwish. "Invariant image watermarking using accurate polar harmonic transforms." *Computers & Electrical Engineering* 62 (2017): 429-447. <https://doi.org/10.1016/j.compeleceng.2017.05.015>
4. Tao, H., Chongmin, L., Zain, J. M., & Abdalla, A. N. (2014). Robust image watermarking theories and techniques: A review. *Journal of applied research and technology*, 12(1), 122-138.
5. Yuan, X. C., & Li, M. (2018). Local multi-watermarking method based on robust and adaptive feature extraction. *Signal Processing*, 149, 103-117. <https://doi.org/10.1016/j.sigpro.2018.03.007>
6. Yamni, M., Karmouni, H., Sayyouri, M., & Qjidaa, H. (2022). Efficient watermarking algorithm for digital audio/speech signal. *Digital Signal Processing*, 120, 103251. <https://doi.org/10.1016/j.dsp.2021.103251>
7. M. Torcoli, T. Kastner and J. Herre, "Objective Measures of Perceptual Audio Quality Reviewed: An Evaluation of Their Application Domain Dependence," in *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 1530-1541, 2021, doi: 10.1109/TASLP.2021.3069302.
8. P. M. Delgado and J. Herre, "A Data-Driven Cognitive Saliency Model for Objective Perceptual Audio Quality Assessment," *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Singapore, Singapore, 2022, pp. 986-990, doi:

10.1109/ICASSP43922.2022.9747064.

9. Greenspun, Philip, and Leigh Klotz. "Audio analysis VI: testing audio cables." *Computer Music Journal* 12.1 (1988): 58-64.
10. Streijl, R.C., Winkler, S. & Hands, D.S. Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives. *Multimedia Systems* 22, 213–227 (2016). <https://doi.org/10.1007/s00530-014-0446-1>
11. J. Zhang, K. Tan, J. Zhao, H. Wu and Y. Zhang, "A Practical SNR-Guided Rate Adaptation," *IEEE INFOCOM 2008 - The 27th Conference on Computer Communications*, Phoenix, AZ, USA, 2008, pp. 2083-2091, doi: 10.1109/INFOCOM.2008.274.
12. Deller Jr, John R., John G. Proakis, and John H. Hansen. *Discrete time processing of speech signals*. Prentice Hall PTR, 1993.
13. Moore, J. K., & Linthicum, F. H. (2007). The human auditory system: A timeline of development. *International Journal of Audiology*, 46(9), 460–478. <https://doi.org/10.1080/14992020701383019>
14. Vecchi, A. O., Varnet, L., Carney, L. H., Dau, T., Bruce, I. C., Verhulst, S., & Majdak, P. (2022). A comparative study of eight human auditory models of monaural processing. *Acta Acustica*, 6, 17. <https://doi.org/10.1051/aacus/2022008>
15. Becerra Martinez, H.; Hines, A.; Farias, M.C.Q. Perceptual Quality of Audio-Visual Content with Common Video and Audio Degradations. *Appl. Sci.* 2021, 11, 5813. <https://doi.org/10.3390/app11135813>
16. Kabal, Peter. "An examination and interpretation of ITU-R BS. 1387: Perceptual evaluation of audio quality." TSP Lab Technical Report, Dept. Electrical & Computer Engineering, McGill University (2002): 1-89.
17. JULIY, BOIKO; ANDRIY, MOKRYTSKY; ILYA, PYATIN. ДОСЛІДЖЕННЯ КІЛ СИНХРОНІЗАЦІЇ ЦИФРОВИХ СИСТЕМ ЗВ'ЯЗКУ. Хмельницького національного університету, 2022, 113. DOI 10.31891/2307-5732-2022-313-5-113-121
18. Griffiths, Dawn. *Head first statistics*. O'Reilly Germany, 2008.

19. Y. Xiong and Z. X. Ming, "Covert Communication Audio Watermarking Algorithm Based on LSB," International Conference on Communication Technology, ICCT 06, pp. 1-4, 2006.
20. B. S. Ko, R. Nishimura, and Y. Suzuki, "Time-spread Echo Method for Digital Audio Watermarking," IEEE Transactions on Multimedia, vol.7(2), pp. 212-221, 2005.
21. H. O. Oh, J. W. Seok, J. W. Hong, and D. H. Youn, "New Echo Embedding Technique for Robust and Imperceptible Audio Watermarking," Proc. ICASSP 2001, pp. 1341-1344, 2001.
22. Wang, Xiang-Yang, Pan-Pan Niu, and Hong-Ying Yang. "A robust digital audio watermarking based on statistics characteristics." Pattern recognition 42.11 (2009): 3057-3064. <https://doi.org/10.1016/j.patcog.2009.01.015>
23. N. Cvejic and T. Seppänen, "Increasing robustness of LSB audio steganography by reduced distortion LSB coding," Journal of University Computer Science, vol. 11(1), pp. 56-65, 2005.
24. N. Rashmi, "Analysis of Audio Steganography combined with Cryptography for RC4 and 3DES Encryption," 2020 Fourth International Conference on Inventive Systems and Control (ICISC), Coimbatore, India, 2020, pp. 210-214, doi: 10.1109/ICISC47916.2020.9171142.
25. Gruhl, Daniel, Anthony Lu, and Walter Bender. "Echo hiding." Information Hiding: First International Workshop Cambridge, UK, May 30–June 1, 1996 Proceedings 1. Springer Berlin Heidelberg, 1996.
26. Huang, HN., Chen, ST., Lin, MS. et al. Optimization-Based Embedding for Wavelet-Domain Audio Watermarking. J Sign Process Syst 80, 197–208 (2015). <https://doi.org/10.1007/s11265-013-0863-y>
27. Masmoudi, S., Charfeddine, M. & Ben Amar, C. A Semi-Fragile Digital Audio Watermarking Scheme for MP3-Encoded Signals Using Huffman Data. Circuits Syst Signal Process 39, 3019–3034 (2020). <https://doi.org/10.1007/s00034-019-01299-4>
28. D. Megías, J. H. Joancomartí, and J. Minguillón, "Total Disclosure of

the Embedding and Detection Algorithms for a Secure Digital Watermarking Scheme for Audio,” ICICS 2005, pp. 427-440, 2005.

29. M. Fallahpour and D. Megas, “High capacity audio watermarking using FFT amplitude interpolation,” IEICE Electronics Express, vol. 6 (14), pp. 1057–1063, 2009.

30. C. Cai, Z. Chen, J. Luo, H. Pu, M. Hu and R. Zheng, "Boosting Chirp Signal Based Aerial Acoustic Communication Under Dynamic Channel Conditions," in IEEE Transactions on Mobile Computing, vol. 21, no. 9, pp. 3110-3121, 1 Sept. 2022, doi: 10.1109/TMC.2021.3051665

31. He, J., Liu, Z., Lin, K. et al. A novel audio watermarking algorithm robust against recapturing attacks. *Multimed Tools Appl* 82, 18599–18616 (2023). <https://doi.org/10.1007/s11042-022-14197-w>

32. Abdelwahab, Khaled M., et al. "Efficient SVD-based audio watermarking technique in FRT domain." *Multimedia Tools and Applications* 79 (2020): 5617-5648.

33. J.Blackledge, “Digital Watermarking and Self-Authentication using Chirp Coding,” ISAST Transactions on Electronics and Signal Processing, ISSN 1797-2329, vol. 1 (1), pp. 61 – 71, 2007.

34. Р. І. Растегаєв, В. О. Мартовицький, Н. М. Бологова, Б. В. Філоненко, О. В. Чечуй ОГЛЯД МЕТОДІВ НАНЕСЕННЯ ЦИФРОВИХ ВОДЯНИХ ЗНАКІВ ДЛЯ ЗАХИСТУ АУДІОФАЙЛІВ Системи управління, навігації та зв'язку. 2024. № 4 с. 69-76