

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМ. В. ДАЛЯ
ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ ТА ЕЛЕКТРОНІКИ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК ТА ІНЖЕНЕРІЇ

До захисту допускається
т.в.о. завідувача кафедри
_____ Сафонова С.О.
« ____ » _____ 2020 р.

МАГІСТЕРСЬКА РОБОТА

НА ТЕМУ:

Дослідження алгоритму розпізнавання музики і обробки аудіо сигналів

Освітньо-кваліфікаційний рівень “Магістр”
Спеціальність 122 – “Комп’ютерні науки”

Науковий керівник роботи:

_____ (підпис)

Щербакова М.Є.

(ініціали, прізвище)

Консультант з охорони праці:

_____ (підпис)

Критська Я.О.

(ініціали, прізвище)

Студент:

_____ (підпис)

Хоткін І.О.

(ініціали, прізвище)

Група:

КН-18дм

Севєродонецьк 2020

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
СХІДНОУКРАЇНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ВОЛОДИМИРА ДАЛЯ

Факультет Інформаційних технологій та електроніки
Кафедра Комп'ютерних наук та інженерії
Освітньо-кваліфікаційний
рівень Магістр
Напрямок підготовки _____
(шифр і назва)
Спеціальність 122 – “Комп'ютерні науки”
(шифр і назва)

ЗАТВЕРДЖУЮ:

Т.в.о. завідувача кафедри

_____ С.О. Сафонова

« _____ » _____ 2020 р.

**З А В Д А Н Н Я
НА МАГІСТЕРСЬКУ РОБОТУ СТУДЕНТУ**

Хоткіну Іллі Олександровичу

(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження алгоритму розпізнавання музики і обробки аудіо сигналів

керівник проекту (роботи) к.т.н., доц. Щербакова М.Є.

(прізвище, ім'я, по батькові, науковий ступінь, вчене звання)

затверджені наказом вищого навчального закладу від " 11" 10 2019 р. № 35/15.15

2. Строк подання студентом роботи _____

3. Вихідні дані до роботи Матеріали переддипломної практики

4. Зміст розрахунково-пояснювальної записки (перелік питань, які потрібно розробити) 1. Кодування звукової інформації

2. Розвиток алгоритмів ШПФ і їх застосування

3. Проектування підсистеми розпізнавання музики

4. Охорона праці в галузі

5. Перелік графічного матеріалу (з точним зазначенням обов'язкових креслень)
Електронні плакати

6. Консультанти розділів проекту (роботи)

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
Охорона праці	ст. викл. Критська Я.О.		

7. Дата видачі завдання _____

Керівник

_____ (підпис)

Завдання прийняв до виконання

_____ (підпис)

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту (роботи)	Строк виконання етапів проекту (роботи)	Примітка
1	Отримання завдання, збір матеріалів	18.10.19- 24.10.19	
2	Огляд літератури й обґрунтування необхідності дослідження	25.10.19 –23.10.19	
3	Дослідження методів розпізнавання акустичних сигналів	24.10.19 – 24.11.19	
4	Аналіз технологій вилучення інформації з аудіозапису	11.11.19 –20.11.19	
5	Розроблення і тестування моделей розпізнавання акустичних сигналів	03.12.20 – 27.12.20	
6	Оформлення пояснювальної записки	28.12.20 – 08.01.20	
7	Підготовка та подання магістерської роботи до захисту	09.01.20 – 10.01.20	

Студент

_____ (підпис)

Хоткін І.О.

_____ (прізвище та ініціали)

Науковий керівник

_____ (підпис)

Щербакова М.Є.

_____ (прізвище та ініціали)

АНОТАЦІЯ

Хоткін І.О. Дослідження алгоритму розпізнавання музики і обробки аудіо сигналів.

Розглянуті труднощі, пов'язані з розпізнаванням аудіо. Розроблено підсистему розпізнавання аудіозаписів. Підсистема аналізує аудіозаписи і після розкладення їх на хеші, вона ідентифікує відтворені користувачем невідомі аудіо.

Результатом виконаної роботи є робочий прототип підсистеми, розробленої на мові Python та СУБД MySQL.

Ключові слова: аудіовідбиток, хеш, ШПФ, автоматизована підсистема, Python, MySQL.

АННОТАЦИЯ

Хоткин И.А. Исследование алгоритма распознавания музыки и обработки аудио сигналов.

Рассмотрены трудности, связанные с распознаванием аудио. Разработана подсистема распознавания аудиозаписей. Подсистема анализирует аудиозаписи и, после разложения их на хэши, она идентифицирует воспроизведены пользователем неизвестные аудио.

Результатом проделанной работы является рабочий прототип подсистемы, разработанной на языке Python и СУБД MySQL.

Ключевые слова: отпечаток аудио, хэш, БПФ, автоматизированная подсистема, Python, MySQL.

ABSTRACT

Khotkin I.A. Research of music recognition algorithm and audio signal processing.

The difficulties associated with audio recognition are considered. A subsystem for recognizing audio recordings has been developed. The subsystem analyzes audio recordings and, after decomposing them into hashes, it identifies unknown audio reproduced by the user.

The result of this work is a working prototype of a subsystem developed in Python and MySQL.

Key words: audio bitstream, hash, FFT, automated subsystem, Python, MySQL.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧОК І СКОРОЧЕНЬ	8
ВСТУП	9
РОЗДІЛ 1. КОДУВАННЯ ЗВУКОВОЇ ІНФОРМАЦІЇ	11
1.1 Загальні відомості про звук	11
1.2 Цифрові аудіоформати	14
1.2.1 Різновиди цифрових аудіоформатів	14
1.3 Аналоговий сигнал	15
1.4 Оцифровка звуку	17
1.4.1 Дискретизація звуку	19
1.4.2 Лінійне (однорідне) квантування	21
1.5 Складнощі, пов'язані з оцифруванням	22
1.6 Технічне завдання на розробку	24
1.6.1 Ціль підсистеми	24
1.6.2 Загальні вимоги до вилучення відбитків	24
1.6.3 Користувачі підсистеми	25
1.6.4 Технічні вимоги	25
1.6.5 Складнощі у реалізації підсистеми	25
1.7 Висновки до розділу 1	26
РОЗДІЛ 2. РОЗВИТОК АЛГОРИТМІВ ШПФ І ЇХ ЗАСТОСУВАННЯ	27
2.1 Загальні відомості про ДПФ і їх застосування	27
2.2 Історична перспектива	29
2.2.1 Дискретне перетворення Фур'є	29
2.2.2 Значення ДПФ	30
2.2.3 Обчислювальна складність ДПФ	32
2.2.4 Історична перспектива швидких обчислень ДПФ	33
2.3 Просування в алгоритмах ШПФ	34
2.3.1 Комплексні алгоритми ШПФ	34
2.3.1.1 Radix-2 ШПФ алгоритми	34
2.3.1.2 Radix-4 ШПФ алгоритм	37
2.3.1.3 Алгоритми Radix-2 та вищий Radix ШПФ	39
2.3.1.4 Алгоритм split-radix ШПФ	42

2.3.1.5	Складність обчислень для ціннісних алгоритмів ШПФ	44
2.3.2	Алгоритми ШПФ реальної цінності	44
2.3.2.1	Алгоритми ШПФ реальної цінності.....	44
2.3.2.2	ШПФ реальних даних.....	46
2.3.2.3	Швидкий алгоритм на основі трансформації Хартлі	47
2.3.2.4	Швидке дискретне перетворення Фур'є.....	48
2.3.2.5	Складність обчислень для алгоритмів реальної цінності ШПФ	49
2.3.3	Особливі випадки алгоритмів ШПФ.....	50
2.3.3.1	Обрізання ШПФ.....	50
2.3.3.2	Алгоритм Герцеля	52
2.3.3.3	Швидке перетворення Фур'є з рідким введенням.....	52
2.3.3.4	Алгоритми безсилості двох довжин.....	53
2.3.3.5	Масштабовані ДПФ.	54
2.3.3.6	Багатовимірні ШПФ.	54
2.3.3.7	Квантове перетворення Фур'є.....	55
2.4	Застосування ШПФ	55
2.5	Висновки до розділу 2.....	59
РОЗДІЛ 3. ПРОЕКТУВАННЯ ПІДСИСТЕМИ РОЗПІЗНАВАННЯ МУЗИКИ		60
3.1	Структура підсистеми.....	60
3.2	Музика як сигнал.....	61
3.3	Вибірки (Sampling).....	61
3.4	Спектрограма	62
3.5	Знаходження піків	64
3.6	Хешування відбитків.....	65
3.7	Послідовність роботи програми.....	68
3.8	Вирівнювання відбитків.....	74
3.9	Використання системи за призначенням	75
3.10	Продуктивність зберігання даних.....	76
3.11	Висновки до розділу 4.....	77
РОЗДІЛ 4. ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ		78
4.1	Аналіз стану умов праці	78
4.2	Виробнича санітарія	82
4.2.1	Аналіз небезпечних та шкідливих факторів при виробництві (експлуатації) виробу ..	82
4.2.2	Пожежна безпека	83

4.2.3 Електробезпека.....	83
4.3 Гігієнічні вимоги до параметрів виробничого середовища	84
4.4 Рекомендації із пожежної профілактики.....	85
4.5 Охорона навколишнього природного середовища	87
4.6 Мікроклімат	88
4.6 Висновки до розділу 4.....	89
ВИСНОВКИ.....	90
ПЕРЕЛІК ДЖЕРЕЛ І ПОСИЛАНЬ.....	91
ДОДАТОК А. ЕЛЕКТРОННІ ПЛАКАТИ	94

ПЕРЕЛІК УМОВНИХ ПОЗНАЧОК І СКОРОЧЕНЬ

АЦП – аналого-цифровий перетворювач

АЧХ – амплітудно-частотна характеристика

АШПФВ - алгоритм ШПФ Виноград

ДД – динамічний діапазон

ДКП - дискретне косинусне перетворення

ДПФ – дискретне перетворення Фур'є

ДПХ - дискретне перетворення Хартлі

ДСП - дискретне синусне перетворення

ЗШПФ - зворотне ШПФ

ІКМ – імпульсно-кодова модуляція

КПФ – квантове перетворення Фур'є

МОЧР - мультиплексування з ортогональним частотним розділенням

ПЗЧ ШПФ – алгоритм ШПФ з проріджуванням по частоті

РоШПФ - розріджене ШПФ

СГС – одна з систем одиниць вимірювання

ШПФ – швидке перетворення Фур'є

ШПХ - швидке перетворення Хартлі

ВСТУП

Аудіо-відбиток відомий завдяки його здатності пов'язувати невизначений звук до відповідних метаданих, незалежно від формату звуку. Хоча для аудіо дактилоскопії є більше застосувань, таких як наступні: перевірка цілісності на основі вмісту або підтримка водяних знаків. Аудіо-відбитки на основі вмісту аудіо-ідентифікації витягують перцептивний довідник частини аудіоконтенту, тобто відбитка і зберігають його в базі даних. Коли він подається з маркованим аудіо, його відбиток обчислюється і порівнюється з тими, що зберігаються в базі даних. Використовуючи відбитки та відповідні алгоритми, спотворені версії запису можуть бути ідентифіковані як однаковий звуковий вміст.

Джерело труднощів при автоматичному визначенні аудіоконтенту впливає з його високої розмірності та значної відмінності аудіоданих щодо сприйнятого подібного вмісту. Найпростіший підхід, про який можна думати - пряме порівняння оцифрованої форми хвилі. Ефективна реалізація цього підходу може використовувати хеш-метод, такий як SHA1 або CRC (циклічна перевірка надмірності) для отримання компактного представлення двійкового файлу. У цій установці можна порівнювати хеш-значення замість цілих файлів. Однак значення хешу неміцні, одного перевернутого біта достатньо, щоб хеш повністю змінився. Звичайно, ця установка не є надійною для стиснення або мінімальних деформацій будь-якого виду, і насправді вона не може розглядатися як ідентифікація на основі контенту, оскільки вона не розглядає вміст, зрозумілий як інформацію, лише біти.

Більшість методів стиснення звуку (AAC, MP3, WMA, Vorbis) внесуть радикальні зміни у двійкове кодування аудіофайлу, не впливаючи радикально на те, як воно сприймається людським вухом. Міцний акустичний відбиток дозволить ідентифікувати запис після такого стиснення, навіть якщо якість звуку значно знизилася.

Мета і завдання дослідження.

Метою магістерської роботи є розробка і дослідження математичних моделей, методів, алгоритмів, програмних модулів і комплексів для вирішення завдання автоматизації ідентифікації аудіо.

Метою дослідження є підвищення якості та зручності зіставлення і знаходження аудіовідбитків часткової автоматизації даного процесу з використанням розробленого алгоритму.

Об'єкт дослідження – це процеси автоматизації ідентифікації аудіо-записів.

Предмет дослідження – моделі, методи й інструментальні засоби для ідентифікації аудіо.

Методи дослідження. Для вирішення поставлених в роботі завдань використані методи математичного моделювання, структури і бази даних, теорія реляційних відносин.

Наукова новизна отриманих результатів.

Удосконалено метод автоматизованого зняття відбитків з аудіо шляхом розробки алгоритму на основі лінійного програмування, що дозволяє підвищити якість та зручність у знаходженні аудіо.

Практичне значення отриманих результатів полягає в тому, що основні наукові положення магістерської дисертації реалізовані у вигляді програмних засобів, які утворюють прикладну технологію для забезпечення автоматизації ідентифікації пісень.

Розроблена підсистема призначена для використання користувачем, який намагається встановити невідому для нього пісню.

Особистий внесок здобувача полягає у розробленні нових методів та інструментальних засобів, що дозволяють вирішити поставлені завдання.

У роботі виконано наступне:

- розроблено підсистему, яка проводить аналіз аудіо і записує отримані після аналізу дані до бази даних;
- здійснено опис розробленої підсистеми визначення аудіо.

Апробація результатів магістерської роботи. Результати досліджень оприлюднено на V Молодіжному форумі «ІТ-Ідея» (м. Сєвєродонецьк, 6 грудня 2019 року) і Всеукраїнській науково-практичній конференції з міжнародною участю «Майбутній науковець-2019» (м. Сєвєродонецьк, 12 грудня 2019 р.).

Публікації. За темою магістерської роботи з викладенням її основних результатів опубліковано 3 наукові праці, серед яких 1 стаття у науковому фаховому виданні і 2 тез доповідей.

Структура та обсяг магістерської роботи. Магістерська дисертація складається із вступу, чотирьох розділів, висновків, списку використаних джерел. Загальний обсяг дисертації складає 99 сторінки. Робота містить 27 джерел, 8 таблиць та 49 рисунків, 1 додаток.

РОЗДІЛ 1

КОДУВАННЯ ЗВУКОВОЇ ІНФОРМАЦІЇ

1.1 Загальні відомості про звук

Звук - фізичне явище, що представляє собою поширення у вигляді пружних хвиль механічних коливань у твердій, рідкому або газоподібному середовищі. У вузькому сенсі під звуком мають на увазі ці коливання, що розглядаються в зв'язку з тим, як вони сприймаються органами чуття тварин.

Як і будь-яка хвиля, звук характеризується амплітудою і частотою. Амплітуда характеризує гучність звуку. Частота визначає тон, висоту (см. Висота звуку). Звичайна людина здатна чути звукові коливання в діапазоні частот від 16-20 Гц до 15-20 кГц. Звук нижче діапазону чутності людини називають інфразвуком; вище: до 1 ГГц, - ультразвуком, від 1 ГГц - гіперзвуком. Гучність звуку складним чином залежить від ефективного звукового тиску, частоти і форми коливань, а висота звуку - не тільки від частоти, а й від величини звукового тиску.

Звукові хвилі можуть служити прикладом коливального процесу. Будь-яке коливання пов'язане з порушенням рівноважного стану системи і виражається у відхиленні її характеристик від рівноважних значень з наступним поверненням до початкового значення. Для звукових коливань такою характеристикою є тиск в точці середовища, а її відхилення - звуковим тиском.

Колівальна швидкість вимірюється в м/с або см/с. В енергетичному відношенні реальні коливальні системи характеризуються зміною енергії внаслідок часткової її витрати на роботу проти сил тертя і випромінювання в навколишній простір. У пружному середовищі коливання поступово загасають. Для характеристики затухаючих коливань використовуються коефіцієнт загасання (S), логарифмічний декремент (D) і добротність (Q).

Коефіцієнт загасання відображає швидкість спадання амплітуди з часом. Якщо позначити час, протягом якого амплітуда зменшується в $e = 2,718$ разів, через τ , то:

$$S = 1/\tau \quad (1.1)$$

Зменшення амплітуди за один цикл характеризується логарифмічним декрементом. Логарифмічний декремент дорівнює відношенню періоду коливань до часу загасання τ :

$$D = T/\tau \quad (1.2)$$

Якщо на коливальну систему з втратами діяти періодичною силою, то виникають вимушені коливання, характер яких в тій чи іншій мірі повторює зміни зовнішньої сили. Частота вимушених коливань не залежить від параметрів коливальної системи. Навпаки, амплітуда залежить від маси, механічного опору та гнучкості системи. Таке явище, коли амплітуда коливальної швидкості досягає максимального значення, називається механічним резонансом. При цьому частота вимушених коливань збігається з частотою власних незгасаючих коливань механічної системи.

При частотах впливу, значно менших резонансної, зовнішня гармонійна сила врівноважується практично тільки силою пружності. При частотах збудження, близьких до резонансної, головну роль відіграють сили тертя. За умови, коли частота зовнішнього впливу значно більше резонансної, поведінка коливальної системи залежить від сили інерції або маси.

Властивість середовища проводити акустичну енергію, в тому числі і ультразвукову, характеризується акустичним опором. Акустичний опір середовища виражається ставленням звуковий щільності до об'ємної швидкості ультразвукових хвиль. Питомий акустичний опір середовища встановлюється співвідношенням амплітуди звукового тиску в середовищі до амплітуді коливальної швидкості її часток. Чим більше акустичний опір, тим вище ступінь стиснення і розрідження середовища при даній амплітуді коливання частинок середовища. Чисельно, питомий акустичний опір середовища (Z) знаходиться як добуток щільності середовища (ρ) на швидкість (c) поширення в ній звукових хвиль.

$$Z = \rho c \quad (1.3)$$

Питомий акустичний опір вимірюється в паскаль-секундах на метр ($\text{Па} \cdot \text{с}/\text{м}$) або дин * $\text{с}/\text{см}^3$ (СГС); $1 \text{Па} \cdot \text{с}/\text{м} = 10^{-1} \text{дин} \cdot \text{с}/\text{см}^3$.

Значення питомої акустичного опору середовища часто виражається в $\text{г}/\text{с} \cdot \text{см}^2$, причому $1 \text{г}/\text{с} \cdot \text{см}^2 = 1 \text{дин} \cdot \text{с}/\text{см}^3$. Акустичний опір середовища визначається поглинанням, заломленням і відбиттям ультразвукових хвиль.

Звуковий, або акустичний, тиск в середовищі являє собою різницю між миттєвим значенням тиску в даній точці середовища за наявності звукових коливань і статичним тиском в тій же точці при їх відсутності. Іншими словами, звуковий тиск є змінним тиском в середовищі, обумовлений акустичними коливаннями. Максимальне значення змінного акустичного тиску (амплітуда тиску) може бути розраховане через амплітуду коливання частинок:

$$P = 2\pi f \rho c A \quad (1.4)$$

де P - максимальний акустичний тиск (амплітуда тиску);

f - частота;

c - швидкість поширення ультразвуку;

ρ - щільність середовища;

A - амплітуда коливання частинок середовища.

На відстані в половину довжини хвилі ($\lambda / 2$) значення звукового тиску з позитивного стає негативним. Різниця тисків у двох точках з максимальним і мінімальним його значенням (віддалених один від одного на $\lambda / 2$ вздовж напрямку поширення хвилі) дорівнює $2P$.

Для вираження звукового тиску в одиницях СІ використовується паскаль (Па), що дорівнює тиску в один ньютон на квадратний метр ($\text{Н} / \text{м}^2$). Звуковий тиск в системі СГС вимірюється в $\text{дин} / \text{см}^2$; $1 \text{ дин} / \text{см}^2 = 10^{-1} \text{Па} = 10^{-1} \text{Н} / \text{м}^2$. Поряд із зазначеними одиницями часто користуються позасистемна одиниця тиску - атмосфера (атм) і технічна атмосфера (ат), при цьому $1 \text{ ат} = 0,98 \cdot 10^6 \text{ дин} / \text{см}^2 = 0,98 \cdot 10^5 \text{ Н} / \text{м}^2$. Іноді застосовується одиниця, звана баром або міробар (акустичним баром); $1 \text{ бар} = 10^6 \text{ дин} / \text{см}^2$.

Тиск, який чиниться на частинки середовища при поширенні хвилі, є результатом дії пружних і інерційних сил. Останні викликаються прискореннями, величина яких також зростає протягом періоду від нуля до максимуму (амплітудне значення прискорення). Крім того, протягом періоду прискорення змінює свій знак.

Максимальні значення величин прискорення і тиску, що виникають в середовищі при проходженні в ній ультразвукових хвиль, для даної частки не збігаються в часі. У момент, коли перепад прискорення досягає свого максимуму, перепад тиску стає рівним нулю. Амплітудне значення прискорення (a) визначається виразом:

$$a = \omega^2 A = (2\pi f)^2 A \quad (1.5)$$

Якщо ультразвукові хвилі, які пробігають, нашкодуються на перешкоду, вона відчуває не тільки змінний тиск, а й постійний. Виникаючі при проходженні ультразвукових хвиль ділянки згущування і розрідження середовища створюють додаткові зміни тиску в середовищі по відношенню до навколишнього її зовнішньому тиску. Таке додаткове зовнішнє тиск носить назву тиску випромінювання (радіаційного тиску). Воно служить причиною того, що при переході ультразвукових хвиль через кордон рідини з повітрям утворюються фонтанчики рідини і відбувається відрив окремих крапельок від поверхні. Цей механізм знайшов застосування в утворенні аерозолів лікарських речовин. Радіаційний тиск часто

використовується при вимірюванні потужності ультразвукових коливань в спеціальних вимірниках - ультразвукових вагах.

1.2 Цифрові аудіоформати

Цифровий аудіоформат - формат подання звукових даних, використовуваний при цифровому звукозаписі, а також для подальшого зберігання записаного матеріалу на комп'ютері та інших електронних носіях інформації, так званих звукових носіях.

Аудіофайл (файл, який містить звукозапис) - комп'ютерний файл, що складається з інформації про амплітуду і частоту звуку, збережену для подальшого відтворення на комп'ютері або програвачі.

1.2.1 Різновиди цифрових аудіоформатів

Формат представлення звукових даних в цифровому вигляді залежить від способу квантування аналогово-цифровим перетворювачем (АЦП). У звукотехніці в даний час найбільш поширені два види квантування:

- імпульсно-кодова модуляція;
- сигма-дельта-модуляція.

При імпульсно-кодовій модуляції аналоговий сигнал, що передається перетворюється в цифрову форму за допомогою трьох операцій: дискретизації за часом, квантування по амплітуді і кодування.

Для перетворення аналогового сигналу в цифровий використовується аналого-цифровий перетворювач (АЦП). АЦП через рівні проміжки часу вимірює амплітуду аналогового сигналу - отримує миттєві значення або відліки сигналу, потім перетворює відліки в двійкові слова.

Миттєве вимірне значення аналогового сигналу квантується за рівнями (округляється від найближчого цілого). Число рівнів квантування, зазвичай, дорівнює або кратне цілій ступені числа 2, наприклад, $2^3 = 8$, $2^4 = 16$, $2^5 = 32$ і т. д.

Потім вихідні слова АЦП в паралельному коді піддаються кодуванню за допомогою передачі на регістр зсуву. На виході регістра зсуву формуються пачки кодованих імпульсів в послідовному коді. Потім пачки імпульсів передаються в канал зв'язку.

Частота відліків сигналу (або швидкість оцифровки, частота дискретизації) для виключення втрат інформації відповідно до теореми Котельникова повинна бути не менше подвоєної максимальної частоти в спектрі аналогового сигналу.

Сигма-дельта модулятор заснований на періодичному неповному зрівноважуванні заряду конденсатора інтегратора. Однорозрядний сигма-дельта модулятор першого порядку

працює наступним чином: на першому такті роботи інтегрується вхідний сигнал до тих пір, поки вихідний сигнал інтегратора не досягне порога перемикання синхронного компаратора. Вихідний сигнал компаратора змінюється тільки по зовнішньому сигналу тактування. Даний цифровий сигнал є виходом модулятора, він же надходить в негативний зворотний зв'язок, де за допомогою ЦАП формується аналоговий сигнал, який віднімається з вхідного аналогового сигналу, і тим самим, врівноважує інтегратор, змушуючи його вихід змінюватися в зворотному напрямку. Таким чином, інтегратор починає інтегрувати цю різницю і його вихідний сигнал змінюється в протилежну сторону до тих пір, поки компаратор не перебере в зворотну сторону. Далі ці цикли повторюються, формуючи на виході синхронного компаратора цифрову послідовність.

Найчастіше розрядність квантування і частоту дискретизації вказують для різних звукових пристроїв запису і відтворення як формат уявлення цифрового звуку (24 біта / 192 кГц; 16 біт / 48 кГц).

Формат файлу визначає структуру і особливості подання звукових даних при зберіганні на накопичувачі ПК. Для усунення надмірності аудіо використовуються аудіокодеки, за допомогою яких проводиться стиснення аудіоданих. Виділяють три групи звукових форматів файлів:

- аудіоформати без стиснення, такі як WAV, AIFF;
- аудіоформати зі стисненням без втрат (APE, FLAC);
- аудіоформати зі стисненням з втратами (MP3, Ogg).

Окремо стоять модульні музичні формати файлів. Створені синтетично або з семплів заздалегідь записаних живих інструментів, вони в основному служать для створення сучасної електронної музики (MOD). Також сюди можна віднести формат MIDI, який не є звукозаписом, але за допомогою секвенсора дозволяє записувати і відтворювати музику, використовуючи певний набір команд в текстовому вигляді.

Формати носіїв цифрового звуку застосовують як для масового поширення звукових записів (CD, SACD), так і в професійній звукозапису (DAT, минидиск).

1.3 Аналоговий сигнал

Аналоговим називають сигнал, неперервний у часі і значеннях. Такий сигнал описується неперервною або кусочно неперервною функцією $f(t)$ при цьому і аргумент і функція можуть приймати довільні значення (рис. 1.1).

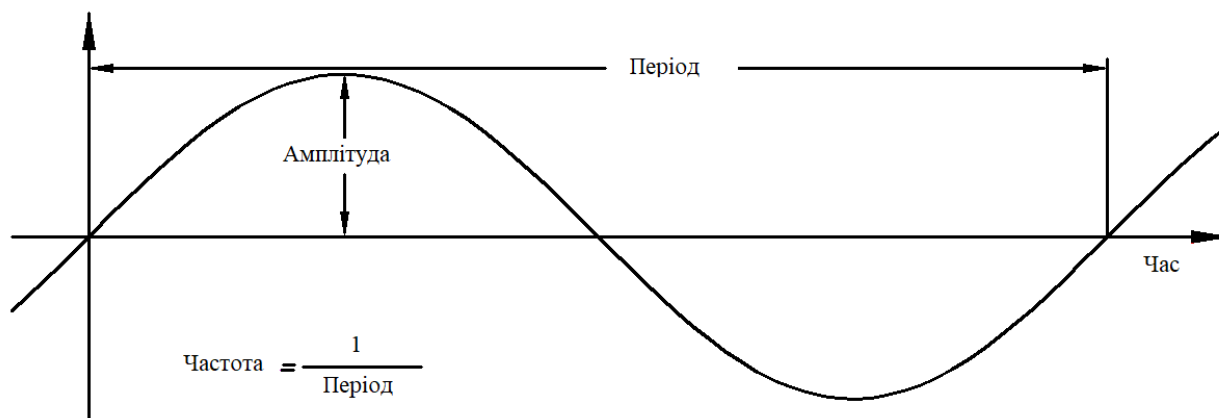


Рисунок 1.1 - Аналоговий сигнал

Аналоговий звук (запис) має безліч параметрів, за допомогою яких можна оцінити його якість. Найважливіших три: частотний діапазон, динамічний діапазон, спотворення.

Частотний діапазон - набір частот, що містяться в звуці. Прийнято вважати, що частотний діапазон людського слуху 20 ... 20.000 Гц (іноді вказується 16 - 22.000 Гц). Сам по собі частотний діапазон музики ніякого інтересу в плані оцінки якості не представляє (наприклад, частотний діапазон все того ж літака, що злітає буде дуже широкий, а вокальної партії тенора - набагато вужчі). Якісним параметром, навушників є потенційний частотний діапазон, а оцінюється він за допомогою амплітудно-частотної характеристики (АЧХ). Ідеальна АЧХ - пряма лінія на всьому діапазоні частот слуху - означає, що джерело звуку не посилює і не послаблює якісь окремі частоти, а значить який отримують звук збігається з оригіналом.

Динамічний діапазон - різниця між самим тихим і найгучнішим звуком. Вимірюється гучність в децибелах (дБ). Прийнято вважати, що максимальна гучність, що не завдає травм людині - це 130 дБ - звук злітаючого літака, а мінімальна чутна гучність - 5 ... 10 дБ - на рівні шелесту листя в маловітряну погоду. Природно, що шелест листя на тлі літака, який взлітає розібрати буде неможливо, та й слухати музику з рівнем 130 дБ вкрай

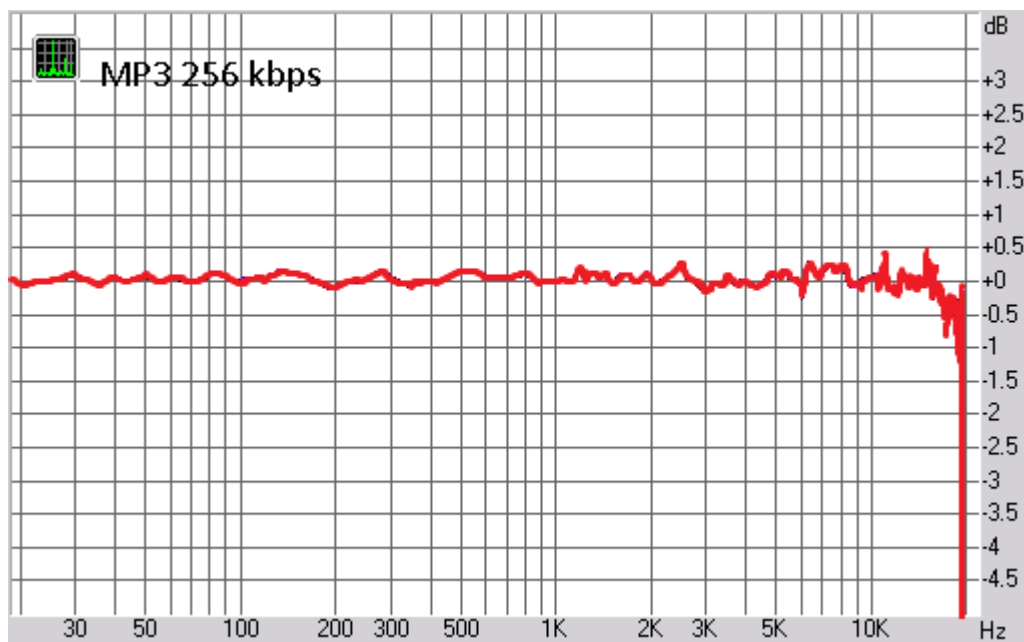


Рисунок 1.2 - АЧХ MP3 файла 256 kbps

неприємно. Тому прийнято вважати, що комфортний ДД для прослуховування музики - 80 ... 100 дБ.

Викривлення - не що інше, як відхилення сигналу від оригіналу.

1.4 Оцифровка звуку

Для того, щоб комп'ютер міг виконати обробку сигналу необхідно виконати перетворення сигналу з аналогової форми в цифрову.

Після обробки виконується зворотне перетворення, оскільки більшість побутових пристроїв управляються аналоговими сигналами.

Структурна схема цифрової обробки сигналу в загальному вигляді зображено на рисунку 1.3.

Цифровий звук - це аналоговий звуковий сигнал, представлений за допомогою дискретних чисельних значень його амплітуди.

Оцифровка звуку - технологія перетворення аналогового звукового сигналу в цифровий вигляд. Полягає в здійсненні вимірів амплітуди сигналу з певним часовим кроком і подальшого запису отриманих значень в чисельному вигляді. Інша назва оцифровки звуку - аналогово-цифрове перетворення звуку.

Звукова хвиля являє собою складну функцію залежності амплітуди хвилі від часу. Наприклад, такі звукові хвилі, неможливо описати за допомогою аналітичних функцій, але при цьому ми можемо в кожен момент часу виміряти значення амплітуди сигналу і записати у

вигляді чисел, і таким чином запам'ятати ці числа. Однак значення амплітуди сигналу ми не можемо записувати з нескінченної точністю, і тому змушені їх округляти. Таким чином, оцифровка сигналу включає в себе два процеси - процес дискретизації, коли ми вимірюємо величину сигналу на певних дискретних проміжках, і квантування, коли ці значення обмежуються певним набором рівнів.

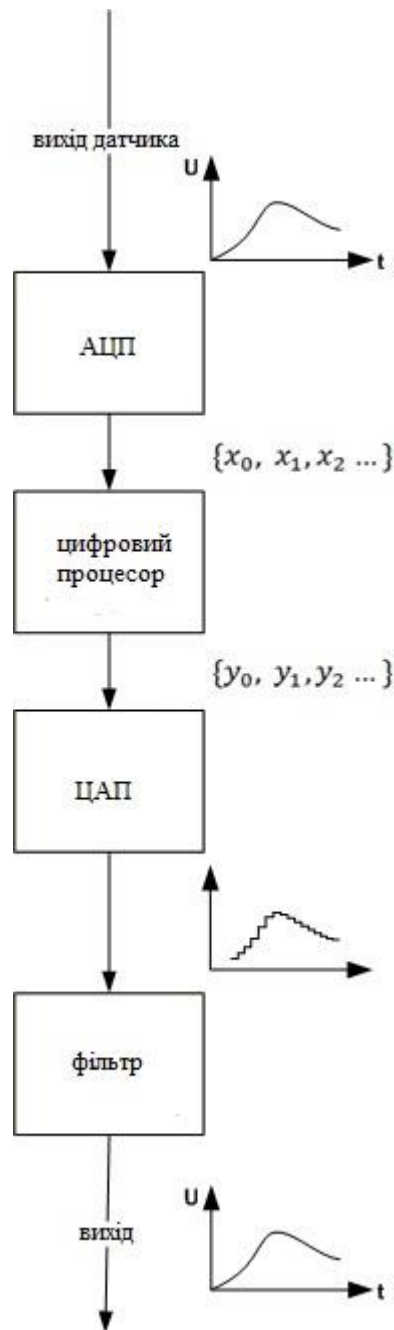


Рисунок 1.3 - Структурна схема цифрової обробки сигналу

1.4.1 Дискретизація звуку

Дискретизація за часом - це процес отримання миттєвих значень аналогового сигналу, який перетворюється з певним часовим кроком, званним кроком дискретизації (рис. 1.4).



Рисунок 1.4 - Дискретизація аналогових сигналів

Кількість здійснюваних в одну секунду замірів величини сигналу називають частотою дискретизації, або частотою вибірки, або частотою семплювання (sampling - вибірка). Очевидно, що, чим менший крок дискретизації, тим правдивіше більша частота дискретизації і значить, тим більш точне уявлення про сигнал можна отримати. Це міркування підтверджується доведеною теоремою Котельникова (вона ж теорема Шеннона). Відповідно до цієї теореми, аналоговий сигнал з обмеженим спектром може бути точно описаний дискретною послідовністю значень його амплітуди, якщо ці значення слідує з частотою, яка як мінімум удвічі перевищує найвищу частоту спектра. Інакше кажучи, аналоговий сигнал, в якому частота найвищої складової спектра дорівнює F_d , може бути точно описаний послідовністю дискретних значень амплітуди, якщо для частоти дискретизації F_d виконується умова $F_d > 2F_m$. На практиці це означає наступне: для того щоб оцифрований сигнал містив інформацію про весь діапазон чутних людиною частот вихідного аналогового сигналу (0-20 кГц), необхідно, щоб обрана частота дискретизації при оцифруванні сигналу була не менше 40 кГц.

Будь-який сигнал має своє спектральне подання. Будь-яке уявлення сигналу - це уявлення у вигляді суми (або інтеграла) гармонійних складових (синусоїд і косинусоїд), різних частот взятих з певними ваговими коефіцієнтами (мають певну амплітуду)

Для періодичних сигналів це сума, для неперіодична - інтеграл.

Перехід до спектру сигналу здійснюється за допомогою прямого перетворення Фур'є.

Перехід до спектрального поданням у вигляді періодичної функції виглядає наступним чином:

$$L(t) = L(t + KT) \quad (1.5)$$

Як відомо періодична функція задовольняє умові Дирихле може бути представлена рядом гармонійних функцій.

$$\sum_{-\infty}^{\infty} C_n e^{\frac{i2\pi n}{T}t} \quad (1.6)$$

За формулою Ейлера будь-який вираз можна представити у вигляді

$$e^{ix} = \cos x + i \sin x \quad (1.7)$$

$\frac{1}{T} = f_1$ - частота першої гармоніки

$\frac{n}{T} = nf_1 = f_n$ - частота n-ї гармоніки

$2\pi f_n = \omega_n$ - кругова частота n-ої гармоніки

$C_n = |C_n|e^{-i\phi_n}$ - комплексна амплітуда гармоніки, де ϕ_n - фазовий спектр.

Сукупність амплітуд гармонік ряду Фур'є називається амплітудним спектром, а сукупність їх фаз називається фазовим спектром.

$$C_n = \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} L(t) e^{-\frac{i2\pi n}{T}t} dt \quad (1.8)$$

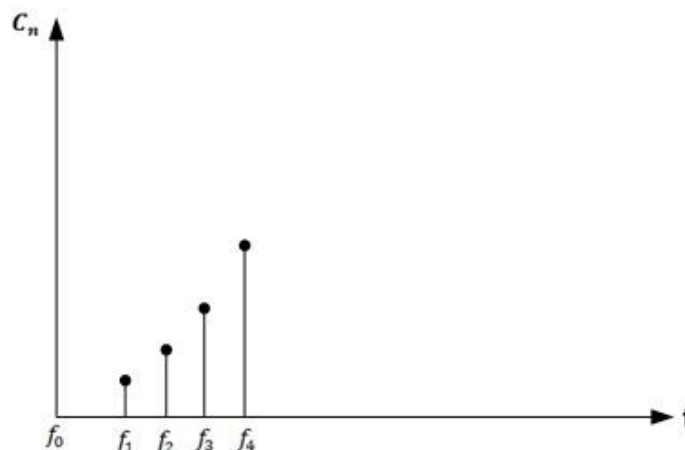


Рисунок 1.5 – Приклад спектру

Для неперіодична функції $T \rightarrow \infty$, а $f \rightarrow 0$ тоді $f_n = \frac{n}{T}$ замінюється безперервно змінюється частотою \Rightarrow сума замінюється інтегралом.

$$L(t) = \frac{1}{2\pi} M(\omega) e^{i\omega t} d\omega \quad (1.9)$$

1.4.2 Лінійне (однорідне) квантування

Припустимо, що для запису одного значення амплітуди сигналу в пам'яті комп'ютера ми відводимо N біт. Відповідно за допомогою одного N -бітного слова можна описати 2^N різних положень. Припустимо тепер, що амплітуда оцифрованого сигналу коливається в межах від -1 до 1 деяких умовних одиниць. Виміряні значення амплітуди можуть бути дробовими (наприклад, $-0,126$ або $0,997$). Уявімо цей діапазон зміни амплітуди - динамічний діапазон сигналу - у вигляді $2^N - 1$ рівних проміжків, розділивши його на 2^N рівнів - квантів (зробивши таким чином однорідне, лінійне розбиття амплітудної шкали). Тепер для запису кожного окремого значення амплітуди його необхідно округлити до найближчого рівня квантування. Цей процес називається квантуванням за амплітудою. Квантування по амплітуді - це процес заміни реальних (вимірних) значень амплітуди сигналу значеннями, наближеними з деякою точністю.

Кожен з 2^N можливих рівнів називається рівнем квантування, а відстань між двома найближчими рівнями називається кроком квантування. У разі лінійного розбиття амплітудної шкали на рівні квантування називають лінійним (однорідним).

Результатом такої оцифровки став ступінчастий сигнал (рис. 1.5), складений з прямокутників, кожен з яких має ширину рівну величині кроку дискретизації, і висоту рівну вимірному значенню амплітуди сигналу.

Точність округлення залежить від вибраної кількості (2^N) рівнів квантування, яке, в свою чергу, залежить від кількості біт (N), відведених для запису значення амплітуди. Чим більше рівнів квантування і чим ближче вони один до одного (для деякого фіксованого діапазону зміни амплітуди відстань між рівнями квантування обернено пропорційно їх кількості), тим на меншу величину доводиться округляти виміряні значення амплітуди і, таким чином, тим менше похибка квантування. Число N називають розрядністю квантування (маючи на увазі кількість розрядів, тобто біт, в кожному слові), а отримані в результаті округлення значень амплітуди числа - відліками або семплами (sample - завмер). Вважається, що похибка квантування з розрядністю 16 біт залишається для слухача майже непомітною (рис. 1.6).

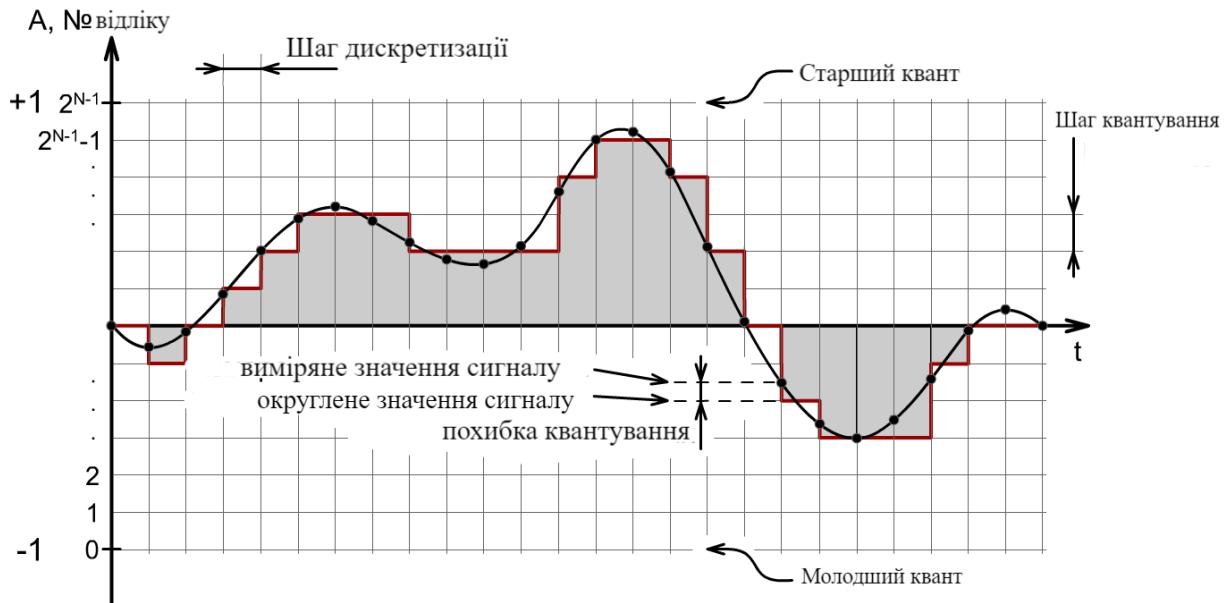


Рисунок 1.6 – Квантування сигналу

Описаний спосіб оцифровки сигналу - дискретизація сигналу в часі в сукупності з методом однорідного квантування - називається сигнал PCM, ІКМ (Pulse Code Modulation, PCM). Стандартний звуковий компакт-диск (CD-DA), що застосовується з початку 1980-х рр. XX століття, зберігає інформацію в форматі PCM з частотою дискретизації 44,1 кГц і розрядністю квантування 16 біт.

1.5 Складнощі, пов'язані з оцифруванням

Незважаючи на гадану простоту і інтуїтивність процесу оцифровки, ця процедура пов'язана з багатьма труднощами і проблемами.

По-перше, частота дискретизації, що встановлюється теоремою Котельникова, є мінімально необхідною, але не достатньою. Дискретизація вносить додаткові перешкоди в спектр вихідного сигналу, додаючи в нього його ж дзеркальну копію. Тому значення частоти дискретизації слід вибирати трохи більшим, ніж частота, що встановлюється теоремою Котельникова, щоб уможливити успішну фільтрацію зайвих спектральних складових.

По-друге, квантування значень сигналу привносить в спектр сигналу додаткову перешкоду, яка називається шумом квантування або шумом дроблення. Шумом (помилкою) квантування називають сигнал, що становить різницю між відновленим цифровим і вихідним аудіо сигналами. Ця різниця утворюється в результаті округлення вимірних значень сигналу. При цьому виконується наступна закономірність: чим вище розрядність квантування, тим

нижче рівень шуму квантування (оскільки тим на менше значення потрібно округляти кожне виміряне значення сигналу).

Природа шуму квантування така, що ширина спектральної області, в якій він простягається, пропорційна значенню частоти дискретизації. При цьому при фіксованій розрядності квантування, загальна енергія шуму квантування залишається постійною. А це значить, що чим більша їх кількість, тим в більш широкій спектральній області простягається шум квантування і, відповідно, тим нижче його потужність в деякій фіксованій смузі спектра, наприклад, в смузі чутних частот. Цей факт має велике практичне значення.

Треба сказати, що рівень шуму квантування також залежить і від форми самого сигналу. В ідеальному випадку, помилка при округленні значень сигналу є випадковою і, отже, спектр шуму квантування виявляється рівномірним. У реальності, однак, цього не відбувається. Форма реальних звукових сигналів є певною мірою не випадковою, а, значить, і помилка квантування теж є не випадковою. В цьому випадку спектр шуму квантування виявляється не рівномірним і концентрується в якійсь певній області, що негативно позначається на звучанні цифрового сигналу.

Існує кілька способів боротьби з цим явищем. Так, небажану концентрацію шуму квантування в деякій частотній області можна нейтралізувати шляхом підмішування до вихідного аналогового сигналу деякого слабкого по потужності псевдовипадкового шуму. Цей шум мінімізує ступінь залежності помилок округлення від форми сигналу, який перетворюється, що добре розсіює спектр шуму квантування і робить його рівномірним. Таким чином, описаний прийом як би підміняє небажаний шум квантування штучно підмішаним псевдовипадковим шумом. При цьому підмісив псевдовипадковий шум виявляється менш помітним на слух, ніж той шум квантування, який вийшов би без застосування тремтіння.

Ще один спосіб боротьби з рівнем шуму квантування називається формуванням шуму (від англ. "Noise shaping"). Ідея прийому полягає в навмисному зміні форми вихідного аналогового сигналу спеціальним чином, щоб наступне квантування призвело до появи шуму квантування, основна енергія якого розташувалася б у найменш помітних на слух частотних областях. Формування шуму з таким спектром досягають шляхом використання фільтра, що моделює криву рівної гучності (про цю кривої ми говорили вище).

Імпульсний сигнал, що отримується в результаті аналогово-цифрового перетворення, через недосконалості перетворюючих пристроїв має деякі вади. Ці вади виражаються в випадкових відхиленнях тривалостей прямокутних імпульсів від номінальної величини кроку дискретизації, а також в неабсолютності крутизни фронтів імпульсів. Інакше кажучи, здійснення вибірки сигналу при оцифрування відбувається не через абсолютно рівні проміжки часу, а з

деякими випадковими відхиленнями від номіналу, при цьому результуючий імпульсний сигнал до того ж має неідеальну прямокутну форму.

Якщо, наприклад, дискретизація проводиться з частотою 44.1 кГц, то відліки беруться не точно кожні секунди, а то трохи раніше, то трохи пізніше. А так як вхідний сигнал постійно змінюється, то така похибка призводить до реєстрації не зовсім вірного рівня сигналу. Описаний неприємний ефект називається тремтінням і по суті є виключно результатом не абсолютної стабільності апаратури (АЦП). На слух джиттер сприймається як деяке тремтіння сигналу на високих частотах, при цьому на низьких частотах джиттер виражається в деяке «розмазування» спектра сигналу. Для боротьби з джиттером застосовують високостабільні тактові генератори. Слід зауважити, що причиною появи джиттера може бути не тільки аналогово-цифрове перетворення, але і передача імпульсного сигналу по цифровому каналу від одного пристрою до іншого. У цьому випадку поява джиттера є результатом недосконалої комутації / синхронізації пристроїв і може бути усунуто тільки шляхом використання апаратури, яка регенерує цифровий сигнал.

Ще один неприємний ефект оцифровки, званий гранулярним шумом. Гранулярним шумом називають ефект нестабільності округлення в процесі квантування. Якщо величина сигналу незначно змінюється при певній величині, що є кордоном між двома сусідніми рівнями квантування, навіть найменші коливання величини сигналу навколо цієї межі можуть викликати помітні зміни результатів округлення при квантуванні значень амплітуди. Це пов'язано з тим, що квантувач в цьому випадку округлює виміряне значення сигналу то до величини одного, то до величини другого з сусідніх рівнів квантування.

1.6 Технічне завдання на розробку

1.6.1 Ціль підсистеми

Дати можливість користувачу ідентифікувати невідоме аудіо.

1.6.2 Загальні вимоги до вилучення відбитків

- Перцептивний довідник запису. Відбиток повинен зберігати максимум акустично-відповідної інформації.

- Незмінний до викривлень. Це впливає з вимоги надійності. Програми цілісності вмісту, однак, знімають це обмеження для збереження вмісту деформацій, щоб виявити навмисні маніпуляції.
- Компактний. Представлення невеликого розміру цікаве за складністю, оскільки потрібно зберігати та порівнювати велику кількість (можливо мільйони) відбитків. Однак занадто коротке представлення може бути недостатньо для дискримінації між записами, що впливає, таким чином, на точність та надійність.
- Легко обчислюється. Зі складних причин вилучення відбитків не повинно займати багато часу.

1.6.3 Користувачі підсистеми

- Люди будь-якого віку.
- Будь-яка кількість чоловіків і жінок.

1.6.4 Технічні вимоги

- Підсистема повинна вміти проводити аналіз аудіо і записувати отримані після аналізу дані до бази даних.
- Повинна бути можливість аналізувати одразу декілька аудіо в черзі.
- Підсистема повинна стискати дані для їх компактного зберігання.
- Підсистема повинна ідентифікувати невідоме (чи відоме) для користувача аудіо, яке вже міститься у БД.

1.6.5 Складнощі у реалізації підсистеми

Складність стосується обчислювальних накладних витрат, пов'язаних з вилученням різних відбитків, які включають:

- розмір відбитка;
- складність використовуваного алгоритму пошуку;
- складність порівняння відбитків та їх відповідність;
- витрати на додавання нових елементів до база даних.

Це надзвичайно важливі параметри для систем, які потребують роботи і розгортання в режимі реального часу, а також для систем з обмеженими обчислювальними ресурсами або пропускнуою здатністю.

1.7 Висновки до розділу 1

У першому розділі магістерської роботи розглянуто і проаналізовано таке фізичне явище як звук. Розглянуто процес перетворення аудіо сигналів у набір даних, з яким за допомогою мов програмування можна виконувати деякі маніпуляції.

Під час досліджень, проведених в рамках даної роботи, були виділені певні складнощі, які пов'язані з оцифруванням аудіо.

РОЗДІЛ 2

РОЗВИТОК АЛГОРИТМІВ ШПФ І ЇХ ЗАСТОСУВАННЯ

2.1 Загальні відомості про ДПФ і їх застосування

В математиці, аналіз Фур'є це наука, що вивчає яким чином загальні математичні функції можуть бути представлені або апроксимовані через суму більш простих тригонометричних функцій. Аналіз Фур'є виник із вивчення властивостей ряду Фур'є, і названий в честь Джозефа Фур'є, який показав, що представлення функції у вигляді суми тригонометричних функцій значно спрощує вивчення процесу теплообміну.

Сьогодні, предметом аналізу Фур'є є широкий спектр математичних задач. В науці і техніці, процес декомпозиції функції на коливальні компоненти часто називають аналізом Фур'є, хоча оперування і відновлення функцій із таких частин відомо як синтез Фур'є. Наприклад, при визначенні які саме компоненти частот присутні в музичній ноті, застосовують розрахунки перетворення Фур'є вибраної музичної ноти. Після чого можна ре-синтезувати той самий звук використовуючи ті частотні компоненти, які виявив аналіз Фур'є. В математиці, термін аналіз Фур'є часто відноситься для вивчення обох цих операцій.

Аналіз Фур'є має багато застосувань в науці – в фізиці, диференційних рівняннях з частинними похідними, теорії чисел, комбінаториці, обробці сигналів, обробці цифрових зображень, теорії ймовірності, статистиці, експертизі, криптографії, чисельному аналізі, акустиці, океанографії, сонарах, оптиці, дифракції, геометрії, структурному аналізі білків, та інших областях.

Така широка застосованість зумовлена багатьма корисними властивостями перетворення:

- Перетворення є лінійними відображенні і, при відповідній нормалізації, є так само унітарними .
- Перетворення як правило є оберненими.
- Показникові функції є власними функціями для диференціювання, це означає що таке представлення перетворює лінійні диференційні рівняння із сталими коефіцієнти в звичайні алгебраїчні рівняння. Таким чином, можна аналізувати поведінку лінійних стаціонарних систем незалежно для кожної частоти.
- Завдяки теоремі про згортки, перетворення Фур'є перетворює складну операцію згортки у просте множення, що означає, що такі перетворення дозволяють

робити розрахунки із операціями на основі згорток, такими як множення многочленів і множення великих чисел, ефективнішим способом.

- Дискретна версія перетворення Фур'є може швидко розраховуватися комп'ютерами із використанням алгоритмів швидкого перетворення Фур'є (ШПФ).

При експертизі, при використанні лабораторних інфрачервоних спектрофотометрів застосовують аналіз перетворення Фур'є для вимірювання довжини хвилі світла при якій матеріал буде поглинати інфрачервоний спектр. Метод перетворення Фур'є використовується для декодування вимірних сигналів і запису даних про довжину хвилі. А при використанні комп'ютера, такі обчислення використовуються швидко, тому такий комп'ютерно-керований пристрій може видати спектр поглинання інфрачервоного випромінення за лічені секунди.

Перетворення Фур'є також використовують для компактного представлення сигналу. Наприклад, алгоритм стиснення JPEG використовує модифікацію перетворення Фур'є (дискретне косинусне перетворення) для невеликих квадратних фрагментів цифрового зображення. Компоненти Фур'є кожного квадрату округлюються до меншої арифметичної точності, а не значними компонентами нехтують, тому компоненти, що залишилися можна зберігати дуже компактно. При реконструкції зображення, кожен квадрат відновлюється із збережених наближених компонентів перетворення Фур'є, які потім зворотно перетворюються для наближеного відновлення початкового зображення.

Дискретне перетворення Фур'є (ДПФ) є найбільш широко використовуваним інструментом у системах цифрової обробки сигналів (DSP). Воно відіграє незамінну роль у багатьох програмах, таких як обробка аудіо та зображень, аналіз сигналів, системи зв'язку та багато інших. Воно відображає послідовність часових областей до послідовності частотної області однакової довжини, тоді як зворотне дискретне перетворення Фур'є (ЗДПФ) здійснює протилежне. Обчислення грубої сили ДПФ довжини N вимагає множення $O(N^2)$, і потрібно використовувати ДПФ довжиною 128 або більше для багатьох практичних застосувань. Обчислювальна вимога в цьому випадку стає більше 10^4 . Через такі високі обчислювальні вимоги для обчислення ДПФ методом «грубої сили» було неможливо використовувати його для додатків DSP в реальному часі і в режимі онлайн до 1965 року, коли Кулі і Тьюкі розробили знаменитий алгоритм швидкого перетворення Фур'є (БПФ). Можна було б зменшити кількість операцій ДПФ з $O(N^2)$ до $(N \log_2 N)$ для ДПФ довжиною N .

2.2 Історична перспектива

У цьому розділі описано значення ДПФ та обчислювальну складність (множення та доповнення), що беруть участь у обчисленні грубої сили ДПФ. Наприкінці цього розділу описано підхід «розділяй і володарюй», який може бути використаний для зменшення обчислювальної складності ДПФ.

2.2.1 Дискретне перетворення Фур'є

N -точка ДПФ / ЗДПФ, відповідно, обчислюється як

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk}, k = 0, 1, 2, \dots, N - 1 \quad (2.1)$$

і

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k)W_N^{-nk}, k = 0, 1, 2, \dots, N - 1 \quad (2.2)$$

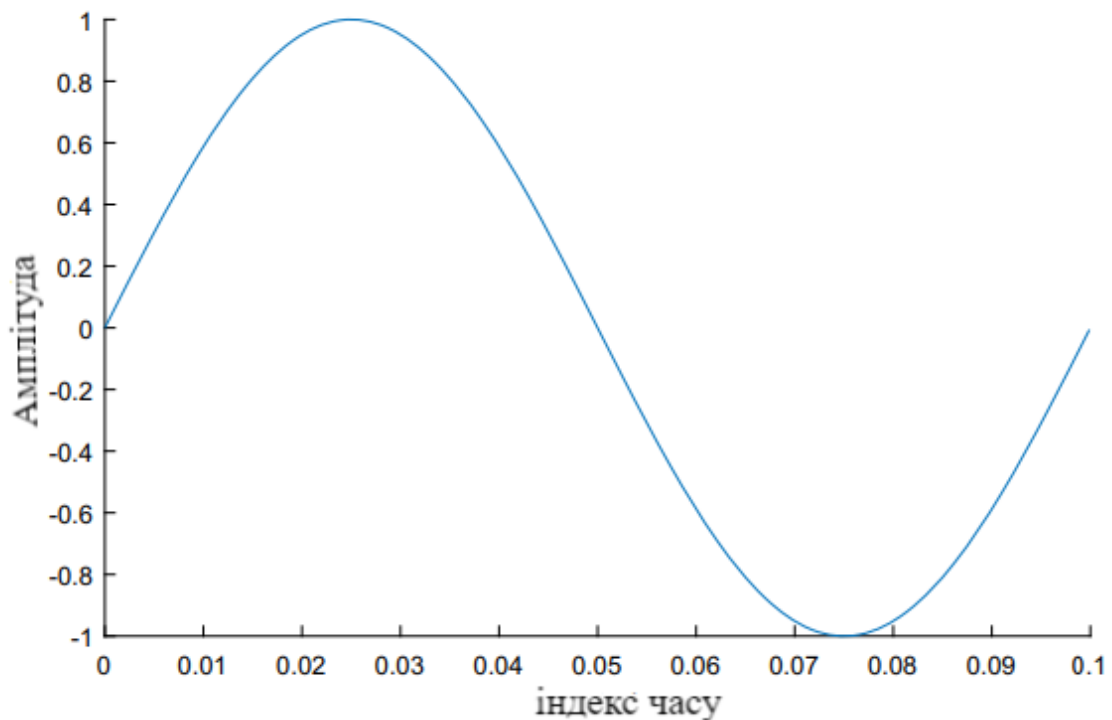


Рисунок 2.1 – Синусоїда

где n - временной индекс, а k - частотный индекс. Коэффициент поворота W_N^{nk} можно представить як:

$$W_N^{nk} = e^{-j2\pi nk/N} = \cos\left(\frac{2\pi nk}{N}\right) - j \sin\left(\frac{2\pi nk}{N}\right) \quad (2.3)$$

У рівняннях (1) та (2) послідовність даних $x(n)$ може бути складною, а k -та спектральна складова $X(k)$ завжди складна. Ці два рівняння відрізняються лише знаком (-) показника коефіцієнта провертання WN та коефіцієнта масштабу $1/N$. Тому алгоритми ефективного обчислення ДПФ можуть бути застосовані для ефективного обчислення ЗДПФ простими та прямими модифікаціями.

2.2.2 Значення ДПФ

Щоб проілюструвати значення ДПФ, необхідно розглянути 4-точкове ДПФ зразків синусоїдального сигналу 10 Гц, який виражається як:

$$x(t) = \sin(2\pi \times 10t) \quad (2.4)$$

Для цієї синусоїди основний період $T_0 = 0,1$ с, як показано на рисунку 2.1. Візьмемо швидкість вибірки $f_s = 40$ Гц, тобто вхід відбирається на кожні $1/f_s = T = 0,025$ с.

Оскільки $N = 4$, нам потрібні 4 вхідні значення вибірки, які можна отримати наступним чином:

$$x(n) = x(nT) = \sin(2\pi \times 10 \times nT) = \sin\left(\frac{n\pi}{2}\right)$$

$$\text{при } n = 0, x(0) = \sin(0) = 0$$

$$\text{при } n = 1, x(1) = \sin\left(\frac{\pi}{2}\right) = 1$$

$$\text{при } n = 2, x(2) = \sin(\pi) = 0$$

$$\text{при } n = 3, x(3) = \sin\left(\frac{3\pi}{2}\right) = -1$$

$$W_4^{nk} = \cos\left(\frac{2\pi nk}{4}\right) - j \sin\left(\frac{2\pi nk}{4}\right) \quad (2.5)$$

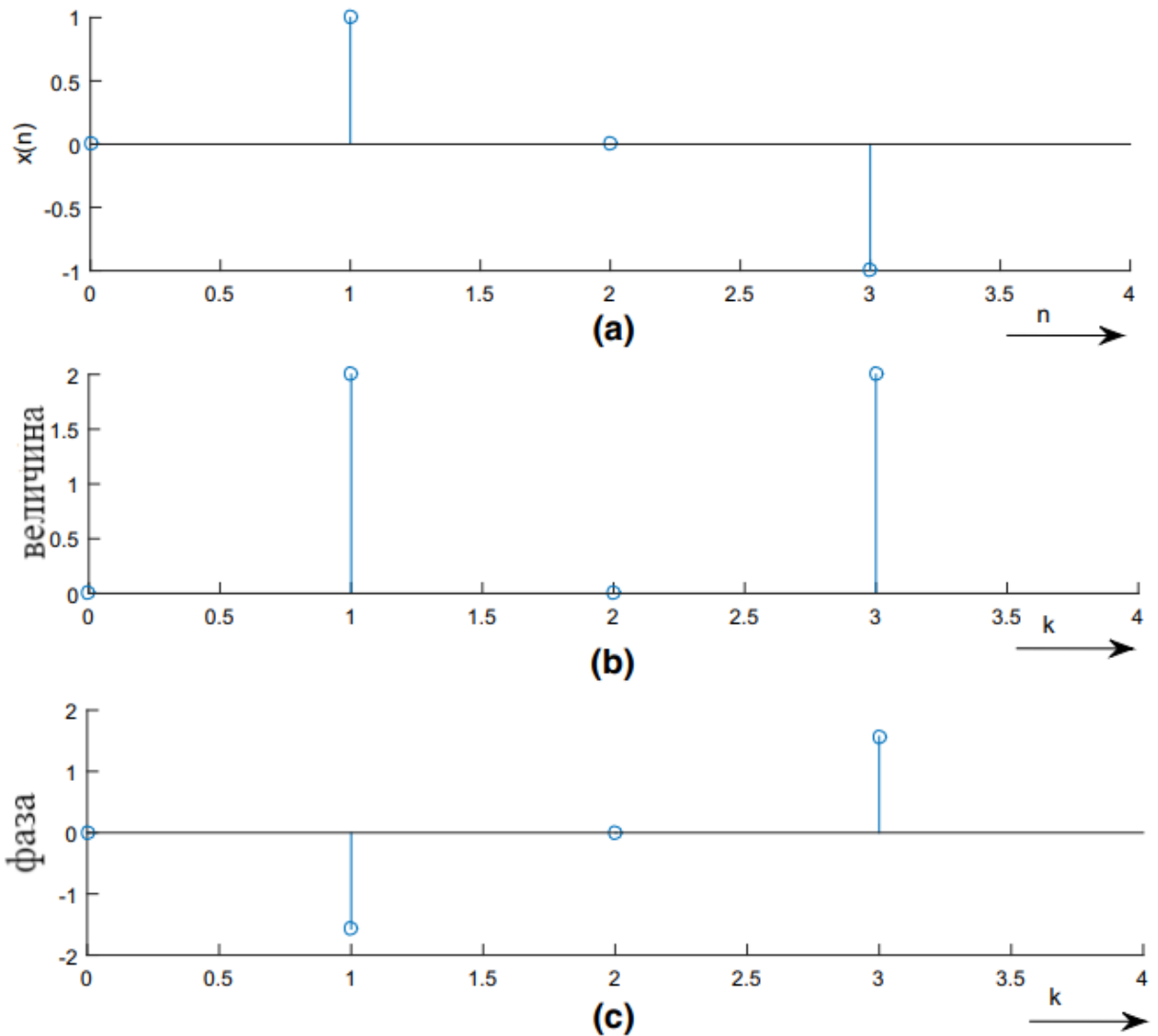


Рисунок 2.2 - Ілюстрація ДПФ для $N = 4$ синусоїди. а - послідовність кінцевої довжини $x(n)$, б величина ДПФ. с - фаза ДПФ

де $nk = 0$ до $N - 1$, тобто 0 до 3. Від рівняння (5), значення W_4^{nk} : $W_4^0 = 1$, $W_4^1 = -j$, $W_4^2 = -1$, $W_4^3 = j$.

Загальне рівняння для 4-точкового ДПФ можна записати як

$$X(k) = \sum_{n=0}^3 x(n)W_4^{kn} = x(0)W_4^{(k)(0)} + x(1)W_4^{(k)(1)} + x(2)W_4^{(k)(2)} + x(3)W_4^{(k)(3)} \quad 0 \leq k \leq 3 \quad (2.6)$$

Вихідні значення ДПФ отримують для $k = 0, 1, 2, 3$ як

$$X(k) = [0 \quad -2j \quad 0 \quad 2j]$$

З вихідних значень ДПФ може бути побудований синусоїдальний сигнал з точки зору його величини та фази, як показано на рисунку 2b, с відповідно. Значення $X(k)$ надає інформацію про k -ту частоту.

Роздільну здатність частоти можна отримати у вигляді:

$$\Delta f = \frac{1}{T_0} = \frac{1}{NT} = \frac{f_s}{N} \quad (2.7)$$

Оскільки основний період синусоїдального сигналу становить 0,1 с, то роздільна здатність частоти дорівнює 10 Гц. З рівняння (7) можна помітити, що для збільшення роздільної здатності частоти кількість точок даних N необхідно збільшити.

2.2.3 Обчислювальна складність ДПФ

Для розрахунку кожного компонента ДПФ безпосередньо відповідно до (1) потрібно N складних множень та $(N - 1)$ складних доповнень. Тому для обчислення всіх N значень ДПФ потрібно загальна кількість N^2 комплексних множень та $N(N - 1)$ складних доповнень. ДПФ N -точкової комплексної вхідної послідовності, $x(n)$, може бути виражена як

$$\begin{aligned} X(k) &= X_R(k) + jX_I(k) = \\ &= \sum_{n=0}^{N-1} [x_R(n) + jx_I(n)] [W_{RN}^{kn} + jW_{IN}^{kn}] = \\ &= \sum_{n=0}^{N-1} [(x_R(n)W_{RN}^{kn} - jx_I(n)W_{IN}^{kn}) + j(x_R(n)W_{IN}^{kn} + x_I(n)W_{RN}^{kn})] \end{aligned} \quad (2.8)$$

де $k = 0, 1, 2, \dots, N - 1$. Припустимо, що кожне комплексне множення в рівнянні (8) реалізується чотирма реальними множеннями та двома реальними додаваннями, тоді як кожне складне додавання реалізується двома реальними додаваннями, прямим обчисленням рівняння (8) вимагає $4N^2$ та $2N(2N - 1)$ кількості реальних множень та дійсних доданків. Більше того, для обчислення ДПФ також потрібен ряд операцій індексації та адресації для отримання вхідних значень, проміжних результатів та складних коефіцієнтів W_N^{kn} та для збереження кінцевих результатів. Для великих значень N арифметична складність ДПФ дуже висока. Тому запропоновані різні алгоритми для зменшення арифметичної складності для швидкого та ефективного обчислення ДПФ.

2.2.4 Історична перспектива швидких обчислень ДПФ

Обчислювальна складність ДПФ істотно зменшується за допомогою наступної тригонометричної симетрії та періодичності коефіцієнта повертання W_N^{kn} :

$$W_N^{k+\frac{N}{2}} = -W_N^k \text{ (Властивість симетрії)} \quad (2.9)$$

$$W_N^{k+N} = W_N^k \text{ (Періодичність власності)} \quad (2.10)$$

Ці властивості були відомі давно ще до початку цифрових обчислень. Гайдеман та ін. простежив першу появу ШПФ повертаючись до Гаусса в 1805 році. Гаус розробив алгоритм для обчислення ДПФ, який еквівалентний одному з алгоритму Кулі – Тукі. Однак Гаусс ніколи не публікував свій алгоритм поза своїми зібраними творами. Попередня робота Даніельсона та Ланцоса [8] посилалася на Рунге [9] щодо їх алгоритму подвоєння у проблемах розсіювання рентгенівських променів. Їх алгоритм показав, як зменшити коефіцієнт ДПФ у $2N$ точках до двох ДПФ у N точках лише трохи більше N операцій. Складність цих алгоритмів була набагато меншою, ніж N^2 , але більше $N \log_2 N$.

Ранні відкриття ШПФ не помічені до опублікування статті Кулі і Тьюки в 1965 році [10]. У цій статті представлений ефективний алгоритм, заснований на підході розділяй і володарюй для обчислення ДПФ. Підхід розділення та володарювання застосовувався до ДПФ рекурсивно, таким чином, що ДПФ будь-якого розміру $N = N_1 N_2$ обчислювався у вигляді менших ДПФ розмірів N_1 та N_2 . Якщо N можна перерахувати на $N = N_1 N_2$, індекси n та k в (1) для N -точки ДПФ можна переписати як:

$$n = N_2 n_1 + n_2 \begin{cases} 0 \leq n_1 \leq N_1 - 1 \\ 0 \leq n_2 \leq N_2 - 1 \end{cases} \quad (2.11a)$$

$$k = N_1 k_2 + k_1 \begin{cases} 0 \leq k_1 \leq N_1 - 1 \\ 0 \leq k_2 \leq N_2 - 1 \end{cases} \quad (2.11b)$$

Представлення індексу (11) може використовуватися в (1) для запису $X(k)$ як:

$$X(k) = \underbrace{\sum_{n_2=0}^{N_2-1} \left(\underbrace{\sum_{n_1=0}^{N_1-1} x(N_2 n_1 + n_2) W_{N_1}^{n_1 k_1}}_{\text{N1-точкове ДПФ}} \underbrace{W_N^{n_2 k_1}}_{\text{Коефіцієнт повертання}} \right) W_{N_2}^{n_2 k_2}}_{\text{N2-точкове ДПФ}} \quad (2.12)$$

де $0 \leq k_1 \leq N_1 - 1$ і $0 \leq k_2 \leq N_2 - 1$.

Обчислення $X(k)$ відповідно до (12) можна здійснити в три етапи:

- 1) обчислити N_1 -точкове ДПФ,
- 2) помножити на подвійні коефіцієнти,
- 3) остаточно обчислити N_2 -точкове ДПФ.

Вищенаведена тришагова процедура може здійснюватися послідовно до $N_1 = 2$. Обчислювальна складність ДПФ за допомогою такого рекурсивного підходу ділення та підкори знижується з $O(N^2)$ до $O(N \log_2 N)$. Це стало головним переломним фактором для DSP-програм у реальному часі ДПФ.

2.3 Просування в алгоритмах ШПФ

Основний принцип підходу «розділяй і володарюй» веде до різноманітних ефективних алгоритмів. Оскільки ці алгоритми покращують продуктивність з точки зору часу обчислень, вони відомі як швидкі алгоритми або алгоритми швидкого перетворення. Далі буде описано три основні класи алгоритмів ШФТ. Перший клас алгоритмів - це комплексні значення ШФТ, де вхід послідовності є комплексним. У другому класі алгоритмів послідовність введення є реальною величиною, відомою як ШФТ з реальною оцінкою. Наприкінці цього розділу описано особливі випадки ШФТ.

2.3.1 Комплексні алгоритми ШПФ

У цьому підрозділі описано популярний алгоритм ШПФ з наступними деякими алгоритмами, які можуть покращити обчислювальну швидкість та зменшити складність обладнання.

2.3.1.1 Radix-2 ШПФ алгоритми

Основними алгоритмами ШПФ є алгоритми децимації в часі (DIT) і алгоритми децимація частоти введення (DIF). Ці алгоритми застосовні для обчислення ДПФ цілочисельної потужності двох довжин

- 1) Алгоритм ШПФ Radix-2 за часом.

Цей алгоритм розкладає послідовність часових областей $\{x(n)\}$ на послідовно меншу послідовність. Тому його називають алгоритмом децимації в часі. Принцип роботи алгоритму

ШПФ radix-2 DIT проілюстровано нижче, враховуючи $N = 2^M$, де $M = 1, 2, 3, \dots$. Оскільки N є парним цілим числом, вхідні дані N -точок можна розділити на дві $(N/2)$ -потоківі підрядки $\{x_1(n)\}$ та $\{x_2(n)\}$, які відповідають парними та непарними індексованими зразками входу $\{x(n)\}$ відповідно, тобто

$$\begin{aligned} x_1(n) &= x(2n) \\ x_2(n) &= x(2n + 1), n = 0, 1, 2, \dots, N/2 - 1 \end{aligned} \quad (2.13)$$

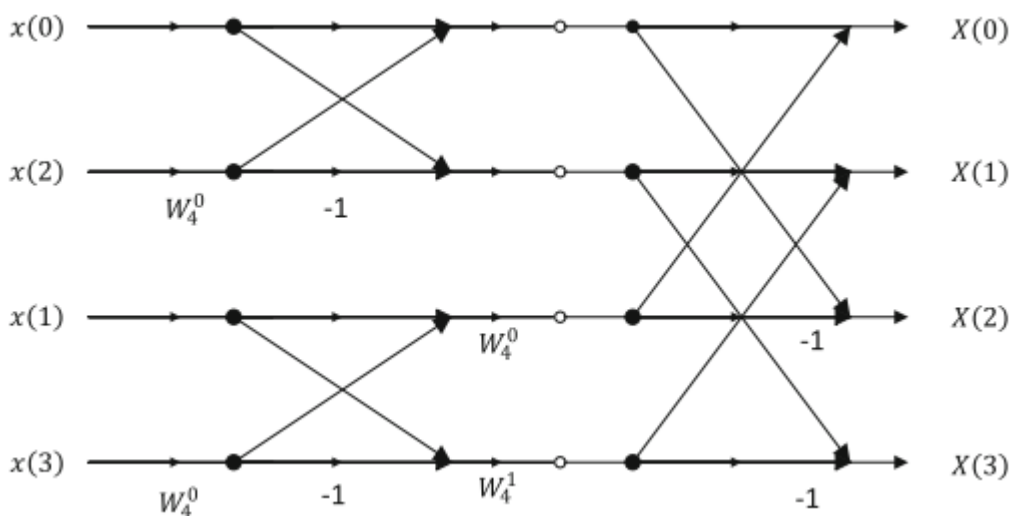


Рисунок. 2.3 - DIT radix-2 ШПФ, довжина – 4

Тепер N -точкове ДПФ може бути отримане з двох половини довжини ДПФ методом децимації в часу, як описано нижче:

$$\begin{aligned} X(k) &= \sum_{n=0}^{N-1} x(n)W_N^{nk} = \sum_{n=0}^{\frac{N}{2}-1} x(2n)W_N^{2nk} + \sum_{n=0}^{\frac{N}{2}-1} x(2n+1)W_N^{(2n+1)k} = \\ &= \sum_{n=0}^{\frac{N}{2}-1} x(2n)W_{\frac{N}{2}}^{nk} + W_N^k \sum_{n=0}^{\frac{N}{2}-1} x(2n+1)W_N^{(2n+1)k} \end{aligned} \quad (2.14)$$

Аналогічно, ДПФ-точки $(N/2)$ -точки можуть бути обчислені з пари $(N/4)$ точок ДПФ. Процес децимації продовжується до тих пір, поки він не містить лише двохточкових ДПФ. Для послідовних потужностей на дві довжини розкладання N -точкової ДПФ на 2-х точкових ДПФ може бути завершено на $M = \log_2 N$ кроків децимації. На малюнку 3 показано розкладання 4-токової radix-2 DIT ШПФ з використанням спрощених метеликів, що включає два етапи, кожен

з яких по два метелика на етап. Вхідні дані до цього знаходяться у бітно-зворотному порядку, а вихід ДПФ - у звичайному порядку.

2) ШПФ алгоритм Radix-2 - зрізання частоти.

Цей алгоритм заснований на обчисленні ДПФ шляхом декомпозиції послідовності коефіцієнтів $X(k)$ ДПФ на більш дрібні послідовності, тому називається алгоритмом Зрізання по частоті.

У разі ШПФ radix-2, обчислення ДПФ розбивається на дві частини так, що перша частина включає в себе перші точки даних $N / 2$, а друга частина включає наступні точки даних $N / 2$ таким чином:

$$X(k) = \sum_{n=0}^{\frac{N}{2}-1} x(n)W_N^{nk} + \sum_{n=\frac{N}{2}}^{N-1} x(n)W_N^{nk} \quad (2.15)$$

Оскільки $W_N^{nk} = e^{-j2\pi k n/N}$ і $W_N^{kN/2} = (-1)^k$, (15) спрощується як:

$$X(k) = \sum_{n=0}^{\frac{N}{2}-1} (x(n) + (-1)^k \times x(n + \frac{N}{2})) \times W_N^{nk} \quad (2.16)$$

Алгоритм radix-2-DIF перебудовується (16) у парні частоти з індексованими та непарними частотами як

$$X(2k) = \sum_{n=0}^{\frac{N}{2}-1} (x(n) + x(n + N/2)) \times W_{N/2}^{nk} \quad (2.17)$$

$$X(2k + 1) = \sum_{n=0}^{\frac{N}{2}-1} ((x(n) - x(n + N/2)) \times W_{N/2}^{nk}) \times W_{N/2}^{nk} \quad (2.18)$$

Згідно з (17) та (18), парно-індексовані та непарно-індексовані частотні виходи $X(k)$ можуть бути обчислені парою $N/2$ -довжина ДПФ. Весь процес включає $M = \log_2 N$ стадій децимації, де кожна стадія включає $N / 2$ метеликів. На малюнку 4 показаний графік потоку розкладання DIF radix-2 4-точкове обчислення ДПФ. У цьому графіку потоку введення знаходиться в нормальному порядку, а вихід ДПФ - у бітно-зворотному порядку. Для обчислення 4-точкове ДПФ йому потрібно чотири складних множення та вісім складних доповнень.

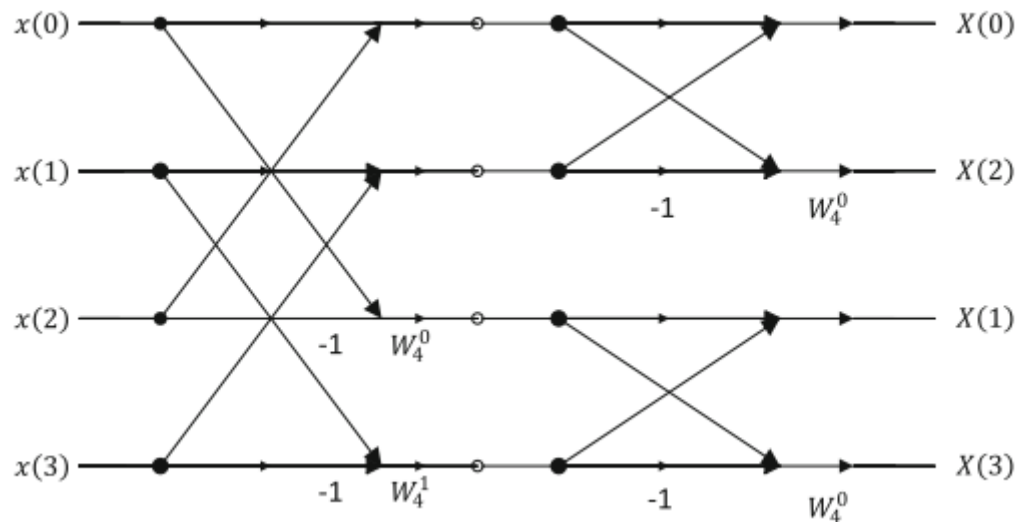


Рисунок. 2.4 - DIF radix-2 ШПФ, довжина-4

Для обчислення N -точкового ДПФ через алгоритми DIF або DIT ШПФ потрібні $(N / 2) \log_2 N$ і $N \log_2 N$ складних множень і складних доповнень відповідно. Для алгоритму radix-2 кількість операцій можна додатково зменшити, реалізуючи кожне комплексне множення на три реальних множення та три реальних додавання. Коли алгоритм використовується для складних множень, арифметична складність ШПФ radix-2 може бути задана:

$$R_M = \frac{3N}{2} \log_2 N - 5N + 8 \quad (2.19)$$

$$R_A = \frac{7N}{2} \log_2 N - 5N + 8 \quad (2.20)$$

де R_M і R_A - це дійсні множення і дійсні додавання для обчислення відповідно N -точкового ДПФ.

2.3.1.2 Radix-4 ШПФ алгоритм

Він може бути використаний, коли довжина ДПФ N дорівнює чотирьом (тобто $N = 4M$). На відміну від алгоритму ШПФ radix-2 в алгоритмі radix-4 під час кожного кроку, децимація проводиться у коефіцієнті 4.

На першому етапі ШПФ radix-4, вхідні дані N -точки поділяються на чотири послідовності як $x(4n)$, $x(4n + 1)$, $x(4n + 2)$, і $x(4n + 3)$, де $n = 0, 1, \dots, N/4 - 1$

Потім

$$\begin{aligned}
X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk} &= \sum_{n=0}^{\frac{N}{4}-1} x(4n)W_{\frac{N}{4}}^{nk} + W_N^k \sum_{n=0}^{\frac{N}{4}-1} x(4n+1)W_{\frac{N}{4}}^{nk} + \\
&+ W_N^{2k} \sum_{n=0}^{\frac{N}{4}-1} x(4n+2)W_{\frac{N}{4}}^{nk} + W_N^{3k} \sum_{n=0}^{\frac{N}{4}-1} x(4n+3)W_{\frac{N}{4}}^{nk}
\end{aligned} \quad (2.21)$$

Оскільки довжина ШПФ в radix-4 дорівнює $N = 4^M$, йому потрібно $M = \log_4 N = \log_2 N / 2$ стадії децимації, де кожна стадія включає $N / 4$ метеликів.

Структура метеликів radix-4 показана на рисунку 2.5.

Процес децимації кожного етапу аналогічний алгоритму radix-2. Оскільки $W_N^0 = 1$, кожен метелик radix-4 передбачає три складних множення та вісім складних додавань. Тому кількість складних множень становить $\frac{3N}{4} \log_4 N$. Якщо порівнювати з підходом radix-2, для цього потрібна менша кількість складних множень, хоча використовується однакова кількість складних додавань. Загальна кількість операцій для N -точкових ШПФ radix-4 становить:

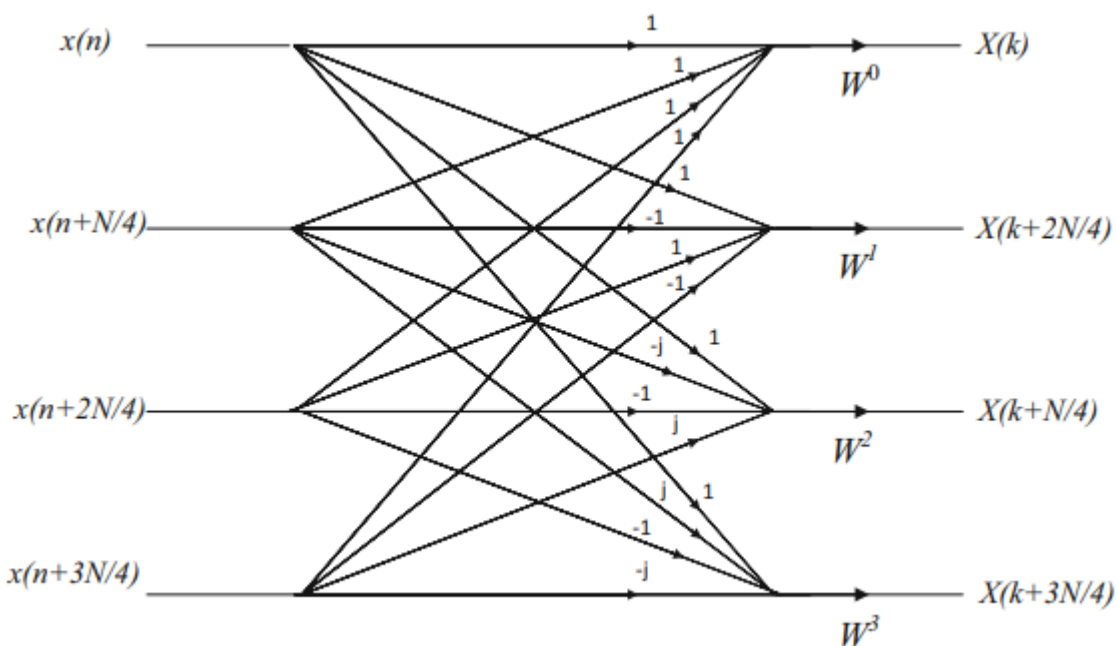


Рисунок. 2.5 - Метелик ШПФ Radix-4-4 ШПФ

$$R_M = \frac{9N}{8} \log_2 N - \frac{43N}{12} + \frac{16}{3} \quad (2.22)$$

$$R_A = \frac{25N}{8} \log_2 N - \frac{43N}{12} + \frac{16}{3} \quad (2.23)$$

2.3.1.3 Алгоритми Radix-2 та вищий Radix ШПФ

Мультиплікаційну складність коефіцієнта повертання можна зменшити, використовуючи більш високі корені, такі як radix-8 або radix-16. Але складність реалізації зростає, коли корень стає вище. Алгоритми radix-2² та radix-2³ мають таку ж кількість нетривіальних множень, як алгоритми radix-4 та radix-8 відповідно. Однак ці алгоритми різняться у коефіцієнтах повертання на різних етапах ШПФ, але підтримують однакову структуру метеликів алгоритму radix-2. Слідом за цими алгоритмами розроблено декілька алгоритмів radix-2ⁱ для вищих коренів, які включають в себе алгоритми radix-2⁴, модифікований radix-2⁴, radix-2⁵ та модифікований алгоритми радіакс-2⁵. Ідея цих алгоритмів radix-2ⁱ полягає в тому, щоб отримати простішу структуру метеликів з меншою мультиплікативною складністю. Наступний підрозділ пояснює походження алгоритму radix-2², який може бути розширений для більш високих коренів.

1) Алгоритм Radix-22.

Нижче представлено алгоритм радікс-22, який використовує техніку розкладання індексу. Для ілюстрації виведення цього алгоритму показники часу та частоти для $i = 2$ розкладаються так:

$$\begin{aligned} n &= \frac{N}{2}n_1 + \frac{N}{4}n_2 + n_3 \{n_1, n_2 = 0, 1, n_3 = 0 \sim \frac{N}{4} - 1\} \\ k &= k_1 + 2k_2 + 4k_3 \{k_1, k_2 = 0, 1, k_3 = 0 \sim \frac{N}{4} - 1\} \end{aligned} \quad (2.24)$$

Підставивши (24) в (1), ми можемо отримати такий вираз:

$$\begin{aligned} X(k_1 + 2k_2 + 4k_3) &= \sum_{n_3=0}^{\frac{N}{4}-1} \sum_{n_2=0}^1 \sum_{n_1=0}^1 x\left(\frac{N}{2}n_1 + \frac{N}{4}n_2 + n_3\right) W_N^{\left(\frac{N}{2}n_1 + \frac{N}{4}n_2 + n_3\right)(k_1 + 2k_2 + 4k_3)} = \\ &= \sum_{n_3=0}^{\frac{N}{4}-1} \sum_{n_2=0}^1 \left\{ B_N^{k_1} \left(\frac{N}{4}n_2 + n_3\right) \right\} W_N^{\left(\frac{N}{4}n_2 + n_3\right)(k_1 + 2k_2 + 4k_3)} \end{aligned} \quad (2.25)$$

де

$$B_N^{k_1} \left(\frac{N}{4}n_2 + n_3\right) = x\left(\frac{N}{4}n_2 + n_3\right) + (-1)^{k_1} x\left(\frac{N}{4}n_2 + n_3 + \frac{N}{2}\right) \quad (2.26)$$

Розкладання загального коефіцієнта повертання в (25) є ключовою відмінністю від розкладання алгоритму radix-2, яка може бути виражена як

$$W_N^{(\frac{N}{4}n_2+n_3)(k_1+2k_2+4k_3)} = (-j)^{n_2(k_1+2k_2)} W_N^{n_3(k_1+2k_2)} W_{N/4}^{n_3k_3} \quad (2.27)$$

Підставляючи (27) в (25), компоненти N-точкового ДПФ можна було отримати з чотирьох ДПФ довжиною N / 4 таким чином:

$$X(k_1 + 2k_2 + 4k_3) = \sum_{n_3=0}^{\frac{N}{4}-1} \left\{ B_{\frac{N}{4}}^{k_1k_2}(n_3) W_N^{n_3(k_1+2k_2)} \right\} W_{N/4}^{n_3k_3} \quad (2.28)$$

де

$$B_{\frac{N}{4}}^{k_1k_2}(n_3) = B_{\frac{N}{2}}^{k_1}(n_3) + (-1)^{k_2} (-j)^{k_1} B_{\frac{N}{2}}^{k_1}(n_3 + \frac{N}{4}) \quad (2.29)$$

N-точкове ДПФ тепер розкладається на чотири ДПФ з довжиною (N / 4), відповідно до (28). Кожне ДПФ довжиною N / 4 можна додатково розкласти таким же чином, поки ДПФ не досягне довжини 2 або 4.

На малюнку 6 показаний графік потоку 16-точкового radix-2² ПЗЧ ШПФ. Це вимагає тривіального множення на $W_{16}^4 = -j$ на першому та третьому етапах, тоді як воно вимагає нетривіального множення на другому етапі. Цей графік потоку відрізняється від алгоритму radix-2, в якому потрібні нетривіальні коефіцієнти повертання на виходах кожного етапу (крім останнього). Цей алгоритм має велику структурну перевагу в порівнянні з іншими алгоритмами (radix-2 і radix-4), коли вони реалізовані в конвеєрних архітектурах.

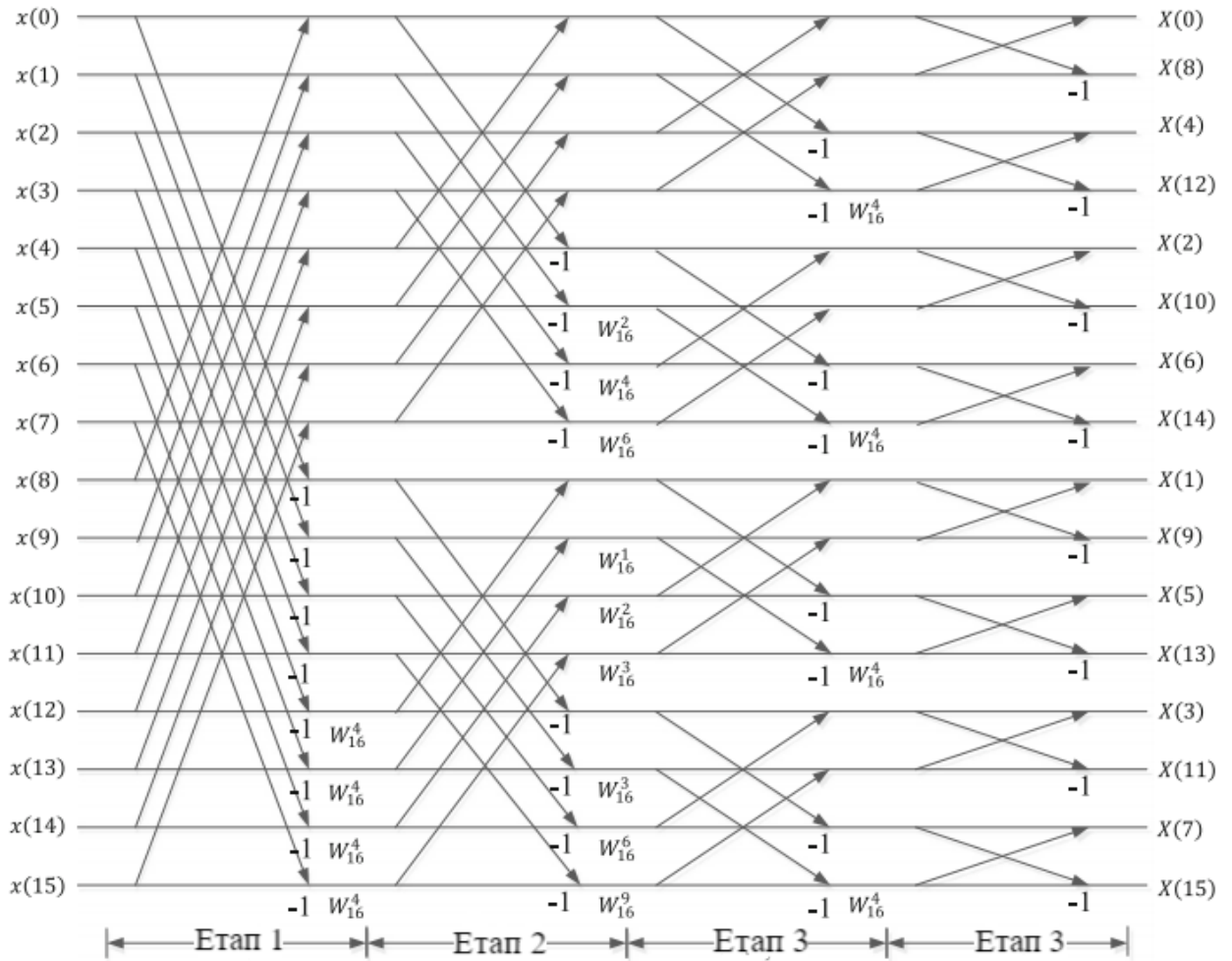


Рисунок 2.6 - Графік потоку сигналу 16-точкового radix-2² ПЗЧ ШПФ

2) Вищі алгоритми Radix.

Лінійна схема декомпозиції алгоритму radix-2² може поширюватися на більш високі корені, наприклад, radix-2³, radix-2⁴, модифікований radix-2⁴ (Radix-M.2⁴), radix-2⁵ та модифікований radix-2⁵ (Radix-M.2⁵). Обчислення ШПФ N-точок з алгоритмом radix-2ⁱ включає в себе log₂ N етапів. У таблиці 1 показано коефіцієнт повертання на кожному етапі для обчислення N точки ШПФ для різних алгоритмів radix-2ⁱ. Ці алгоритми мають однакову структуру метеликів, але структура множення коефіцієнту прокручення змінюється залежно від показника i. Подвійне множення множників класифікується на тривіальне (W₄, яке є множенням на -j), а інші множення нетривіальне. Коефіцієнти повертання множення класифікуються на тривіальні, а інші множення – нетривіальні.

Таблиця 2.1 – Коефіцієнт повертання різних етапів для діаграми потоку N- точкового ШПФ

Алгоритм	Етап							
	1	2	3	4	5	6	7	8
Radix-2 ²	W ₄	W _N	W ₄	W _{N/4}	W ₄	W _{N/16}	W ₄	W _{N/64}
Radix-2 ³	W ₄	W ₈	W _N	W ₄	W ₈	W _{N/8}	W ₄	W ₈
Radix-2 ⁴	W ₄	W ₈	W ₁₆	W _N	W ₄	W ₈	W ₁₆	W _{N/16}
Radix-M .2 ⁴	W ₄	W ₁₆	W ₄	W _N	W ₄	W ₁₆	W ₄	W _{N/16}
Radix-2 ⁵	W ₄	W ₈	W ₁₆	W ₃₂	W _N	W ₄	W ₈	W ₁₆
Radix-M .2 ⁵	W ₄	W ₈	W ₃₂	W ₄	W _N	W ₄	W ₁₆	W ₄

З таблиці 1 видно, що алгоритм radix-2³ вимагає тривіального множення на першому етапі і нетривіального множення на другому та третьому етапах. Цей тип послідовності повторюється кожні три етапи з метою отримання алгоритму radix-2³. Radix-2⁴ включає тривіальне множення на першому етапі та нетривіальне множення на наступних трьох етапах. В роботі [11] запропоновано модифікований алгоритм radix-2⁴, який вимагає меншої кількості множень. У модифікованому алгоритмі коефіцієнт повертання третього етапу W₁₆ переноситься на другий етап.

2.3.1.4 Алгоритм split-radix ШПФ

Split-radix алгоритм ШПФ розбиває N-точкове ДПФ на одну N / 2-точкове ДПФ та два N / 4-точкових ДПФ у вигляді:

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk} = \sum_{n=0}^{\frac{N}{2}-1} x(2n)W_{\frac{N}{2}}^{nk} + W_N^k \sum_{n=0}^{\frac{N}{4}-1} x(4n+1)W_{\frac{N}{4}}^{nk} + W_N^{3k} \sum_{n=0}^{\frac{N}{4}-1} x(4n+3)W_{\frac{N}{4}}^{nk}$$

Цей алгоритм використовує одночасно і поведінку radix-2, і radix -4 одночасно на верхній і нижній половині графіка потоку сигналу, як показано на рисунку 2.7. Арифметична складність алгоритму задається:

$$R_M = N \log_2 N - 3N + 4 \quad (2.30)$$

$$R_A = 3N \log_2 N - 3N + 4 \quad (2.31)$$

Split-radix алгоритм вимагає меншої кількості множень та додавань порівняно з алгоритмами radix-2 та radix-4. Було запропоновано багато алгоритмів radix для подальшого зменшення кількості складних множень та доповнень над алгоритмами на основі radix-2, radix-4 або будь-яких вищих алгоритмів на основі radix.

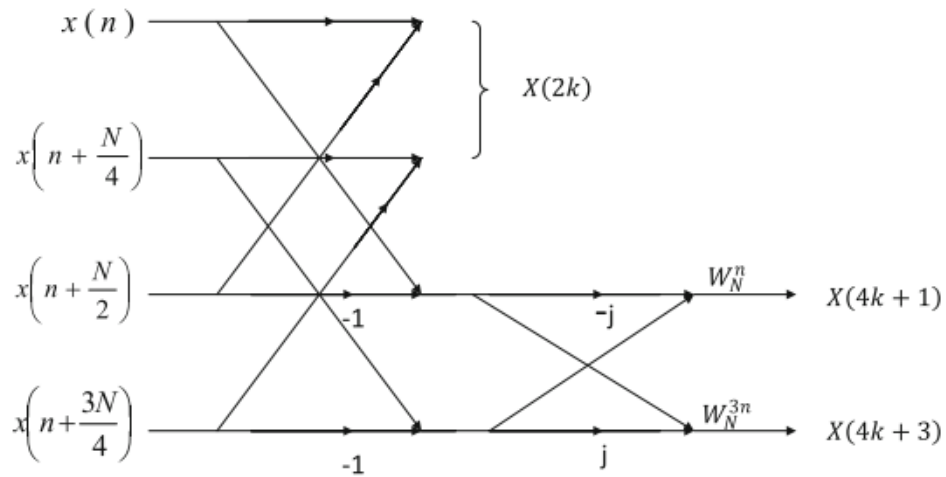


Рисунок 2.7 - Розділений алгоритм radix ШПФ

Таблиця 2.2 - Кількість реальних множень для обчислення комплексу ДПФ довжиною-N

N	Radix-2	Radix-4/Radix 2^2	Split-radix
16	24	20	20
32	88	-	68
64	264	208	196
128	712	-	516
256	1800	1392	1284
512	4360	-	3076
1024	10248	7856	7172

Таблиця 2.3 - Кількість реальних додавань для обчислення комплексу ДПФ довжиною-N

N	Radix-2	Radix-4/Radix 2^2	Split-radix
16	152	148	148
32	408	-	388
64	1032	976	964
128	2504	-	2308
256	5896	5488	5380

Продовження таблиці 2.3

N	Radix-2	Radix-4/Radix22	Split-radix
512	13566	-	12292
1024	30728	28336	27652

2.3.1.5 Складність обчислень для ціннісних алгоритмів ШПФ

У таблицях 2 та 3 показано порівняння кількості реальних множень та дійсних додавань для обчислення N-точкових ДПФ. З цих таблиць можна помітити, що split-radix ШПФ вимагає меншої кількості арифметичних операцій порівняно з іншими алгоритмами. Однак графік потоку цього алгоритму призводить до неправильної структури внаслідок змішування ШПФ різної довжини в різних частинах.

2.3.2 Алгоритми ШПФ реальної цінності

Коли послідовність введення $\{x(n)\}$ реально оцінена, компоненти ДПФ виявляють сполучену симетрію, тобто $X(k) = X^*(N - k)$. Тому потрібно обчислити лише половину кількості компонентів ДПФ в цьому випадку. Але алгоритми ШПФ для обчислення складних вхідних даних не можуть використовуватися безпосередньо для зменшення складності обчислень до половини, коли необхідно обчислити ДПФ реального значення вхідних даних. ШПФ за даними реальної цінності та ШПФ з комплексних даних зазвичай називають ШПФ (РШПФ) реальною цінності та складнозначні ШПФ (СШПФ) відповідно.

Крім того, ефективна реалізація РШПФ отримала велику увагу завдяки декільком важливим та новим програмам у галузі біомедичної інженерії та охорони здоров'я, аудіо- та відеообробці, аналізу часових рядів та багатьох інших. Тому для обчислення РШПФ пропонується кілька алгоритмів. ШПФ реальної цінності забезпечують поліпшення площі та швидкості порівняно із СШПФ. Алгоритми РШПФ, як правило, розроблені для даних з реальною вартістю, використовуючи тригонометричні симетрії та періодичності.

2.3.2.1 Алгоритми ШПФ реальної цінності

Найпростіший спосіб використання алгоритму СШПФ для обчислення РШПФ полягає в встановленні реально оціненої послідовності в реальну частину комплексного введення та встановлення уявної частини вхідних значень на нуль. Такий підхід не забезпечує значної

економії обчислень над СШПФ, оскільки проміжні результати набувають складного оцінювання відразу після першого етапу, коли складні коефіцієнти повертання множать. Тому алгоритм подвоєння та алгоритм упаковки пропонуються для обчислення РШПФ.

1) Алгоритм подвоєння

У цьому алгоритмі одночасно використовується пара реальних значень вхідних послідовностей. Перша послідовність даних з реальною цінністю використовується як реальна частина, а друга реальна послідовність даних як уявна частина складнозначної вхідної послідовності СШПФ. Отримане таким чином складне вхідне значення виражається як:

$$x(n) = p(n) + j \cdot q(n) \quad (2.32)$$

де $p(n)$ і $q(n)$ - елементи двох послідовностей даних, що реально оцінюються. Потім отримують N -точкове СШПФ складного входу $\{x(n)\}$ у вигляді:

$$X(K) = P(K) + j \cdot Q(K) \quad (2.33)$$

Оскільки $p(n)$ та $q(n)$ є реальними даними, має місце наступна симетрія

$$\begin{aligned} P * (N - k) &= P(K) \\ Q * (N - k) &= Q(K) \end{aligned} \quad (2.34)$$

звідси впливає послідовність виведення:

$$X * (N - k) = P(K) - j \cdot Q(K) \quad (2.35)$$

Використовуючи рівняння (33) і (35), $P(k)$ і $Q(k)$ можна отримати у вигляді:

$$\begin{aligned} P(k) &= \frac{1}{2}(X(k) + X * (N - k)) \\ Q(k) &= \frac{j}{2}(X * (N - k) - X(k)) \end{aligned} \quad (2.36)$$

Для відокремлення $P(k)$ і $Q(k)$ відповідно до (36), $2(N - 1)$ необхідні додаткові доповнення над звичайним складним ШПФ. Використовуючи алгоритм 3/3 для складного множення, РШПФ вимагає $\frac{1}{2}MN - \frac{3}{2}N + 2$ множень і $\frac{3}{2}MN - \frac{1}{2}N$ додавань. Цей алгоритм вимагає майже половини арифметичної складності алгоритму СШПФ.

2) Алгоритм упаковки

Це ще один підхід до обчислення N -точкового ШПФ реального значення вхідних даних за допомогою $N/2$ -точкового СШПФ. У ньому використовуються непарні і парні зразки індексованих N -точкових послідовностей введення з реальною вартістю для формування комплексних $(N/2)$ -точкових даних. Це називається алгоритмом упаковки, оскільки він пакує послідовність дійсних значень N -точок у комплексну послідовність $(N/2)$ -точок. Речові дані можуть бути представлені у вигляді складних даних у вигляді:

$$x(n) = x(2n) + j \cdot x(2n + 1) \quad (2.37)$$

де $n=0, 1, 2, \dots, N-1$.

Нехай $p(n) = x(2n)$ і $q(n) = x(2n + 1)$, тоді вихід ДПФ $X(K)$ можна отримати, використовуючи СШПФ, як у алгоритмі подвоєння. Тому для цього також потрібні $2(N - 1)$ додаткові доповнення для розділення виходів СШПФ, як у випадку алгоритму подвоєння. Більше того, він вимагає додаткового етапу для обчислення результатів РШПФ. Відповідне РШПФ вимагає $\frac{1}{2}MN - \frac{5}{4}N$ множень і $\frac{3}{2}MN - \frac{1}{4}N - 4$ додавань за допомогою алгоритму 3/3 для складних множень.

2.3.2.2 ШПФ реальних даних

Зменшення арифметичної складності можна отримати, використовуючи для обчислення РШПФ конкретні алгоритми, такі як алгоритм DIT ШПФ. Цього можна досягти, застосувавши сполучену симетричну властивість та обчисливши лише половину проміжних виходів на кожному етапі, оскільки інші можуть бути отримані сполучанням цих проміжних значень. Це призводить до меншої арифметичної складності для алгоритму radix-2 DIT ШПФ. Беручи до уваги алгоритм 3/3, потрібен ШПФ radix-2 для дійсної послідовності $\frac{3}{4}MN - \frac{5}{2}N + 4$ множень і $\frac{7}{4}MN - \frac{7}{2}N + 6$ додавань.

Алгоритми radix-4 та вищих radix алгоритми для вхідних даних з реальною величиною можна отримати аналогічно алгоритму radix -2 DIT ШПФ. Алгоритм split radix є більш ефективним з точки зору арифметичної складності, ніж вищі radix алгоритми. Він вимагає лише $\frac{1}{2}MN - \frac{3}{2}N + 2$ множень і $\frac{3}{2}MN - \frac{5}{2}N + 4$ додавань. Однак ці алгоритми не є дійсними для розкладання DIF ШПФ, оскільки неможливо застосувати сполучену симетрію на кожному етапі.

2.3.2.3 Швидкий алгоритм на основі трансформації Хартлі

ДПФ реальних цінностей може бути обчислене з дискретної трансформації Хартлі (ДТХ) тих самих даних. ДПХ реальної значущої вхідної послідовності визначається як:

$$X(k) = \sum_{n=0}^{N-1} x(n) \left[\cos\left(\frac{2\pi kn}{N}\right) + \sin\left(\frac{2\pi kn}{N}\right) \right] \quad (2.38)$$

для $k=0, 1, 2, \dots, N-1$.

На відміну від ДПФ, ДПХ приймає реальну оцінку та забезпечує реальну оцінку. Відсутність складної арифметики робить ДПХ швидше, ніж ДПФ. Алгоритми, схожі на ШПФ на основі radix алгоритмів, також можуть бути застосовані до обчислень ДПХ, які називаються алгоритмом швидкого перетворення Хартлі (ШПХ). Як правило, алгоритми ШПХ передбачають однакові множення і $(N-2)$ більше складання порівняно з відповідним алгоритмом ШПФ. Алгоритм split radix ШПХ вимагає $\frac{2N}{3} \log_2 N - \frac{19N}{9} + 3 + \frac{(-1)^M}{9}$ множень і $\frac{4N}{3} \log_2 N - \frac{14N}{9} + 3 + (-1)^M \frac{5}{9}$ додавань. N -точкове ДПФ реальних цінностей може бути обчислене з ДПХ тих самих даних, як:

$$\begin{aligned} \operatorname{Re}(DFT(k)) &= \frac{DHT(k) + DHT(N-k)}{2} \\ \operatorname{Im}(DFT(k)) &= \frac{DHT(k) - DHT(N-k)}{2} \end{aligned} \quad (2.39)$$

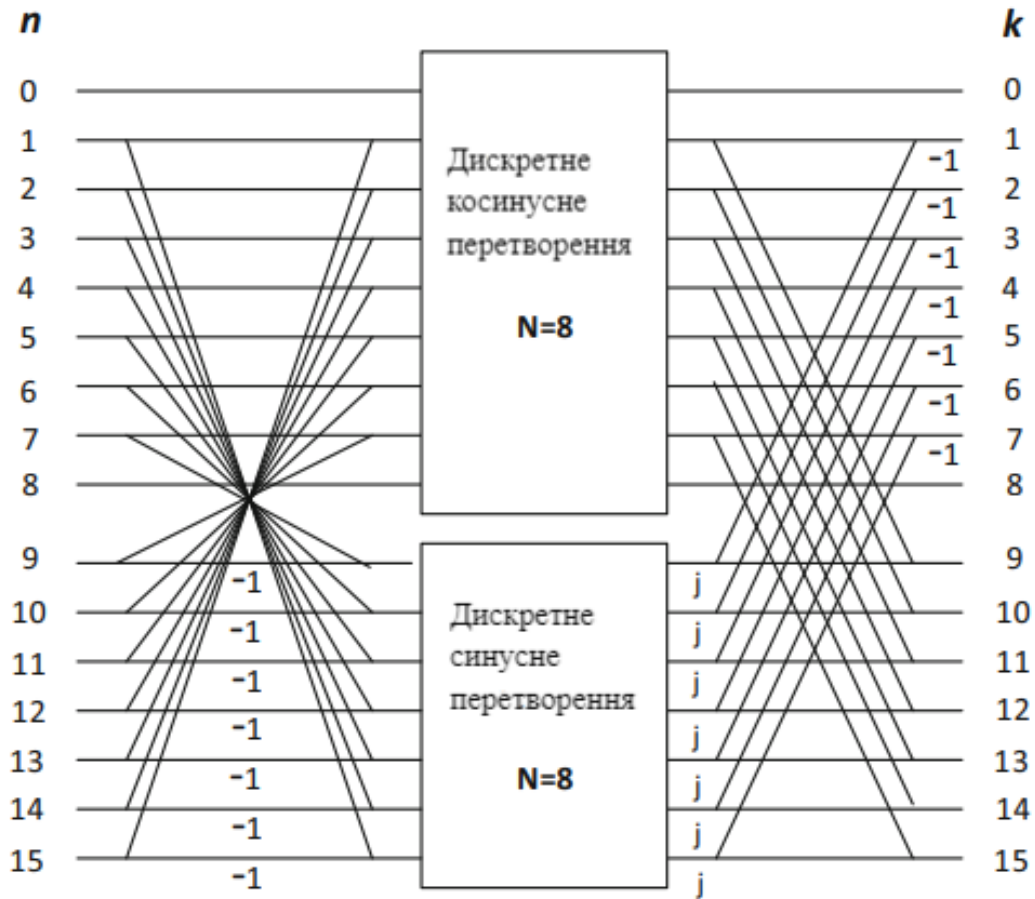


Рисунок 2.8 - Швидке дискретне перетворення Фур'є

2.3.2.4 Швидке дискретне перетворення Фур'є

Цей алгоритм обчислює ДПФ за допомогою дискретного косинусного перетворення (ДКП) та дискретної синусоїди (ДСП). Він розкладає N -точкове ДПФ на $(N / 2 + 1)$ -точкове ДКП і $(N / 2 - 1)$ -точкове ДСП. Швидкий показник ДПФ для 16-точкових даних показаний на рисунку 2.8. Це обчислює ДКП і ДСП окремо, приймаючи складні операції на останньому етапі. Арифметичні операції, необхідні цим алгоритмом для обчислення N -точкового ДПФ, є такими:

$$R_M = \frac{N}{2} \log_2 N - \frac{11}{8} N + 1 \quad (2.40)$$

$$R_A = \frac{7}{4} N \log_2 N - 3N + 2 \quad (2.41)$$

2.3.2.5 Складність обчислень для алгоритмів реальної цінності ШПФ

Хоча більшість алгоритмів ШПФ розроблені для складних вхідних даних, завдяки перевагам надмірностей і тригонометричних симетрій обчислювальна складність знижується у всіх цих алгоритмах РШПФ. Кількість реальних множень та реальних додавань, необхідних для функціонування реальної величини, наведено у Таблицях 4 та 5 відповідно. Якщо СШПФ використовується безпосередньо для реальних входів, це вимагає більшої арифметичної складності. Алгоритми упаковки та подвоєння передбачають більше доповнень, ніж алгоритм split-radix РШПФ для вхідних даних.

Split-radix ШПХ вимагає меншої кількості множень та додавань, ніж radix-2 РШПФ для N , більше ніж 16. Однак для нього потрібна більша кількість множень та додавань, ніж split radix РШПФ. Алгоритм швидкого ДПФ вимагає більш реального множення, ніж алгоритм подвоєння.

Таблиця 2.4 - Кількість реальних множень для обчислення ДПФ для вхідного значення

N	СШПФ пряме	СШПФ упаковка	СШПФ подвоєння	Radix-2 РШПФ	split-radix РШПФ	split-radix ШПХ	Швидке ДПФ
16	20	12	10	12	10	12	11
32	68	40	34	44	34	42	37
64	196	112	98	132	98	124	105
128	516	288	258	356	258	330	273
256	1284	704	642	900	642	828	673
512	3076	1664	1538	2180	1538	994	1601
1024	7172	3840	3586	5124	3586	4668	3713

Таблиця 2.5 - Кількість реальних додавань для обчислення ДПФ для вхідного значення

N	СШПФ пряме	СШПФ упаковка	СШПФ подвоєння	Radix-2 РШПФ	split-radix РШПФ	split-radix ШПХ	Швидке ДПФ
16	148	88	88	62	60	64	66
32	388	228	224	170	164	166	186
64	964	556	544	442	420	416	482
128	2308	1308	1280	1082	1028	998	1186
256	5308	3004	2944	2586	2436	2336	2818
512	12292	6780	6656	5978	5636	5350	6530

Продовження таблиці 2.5

N	СШПФ пряме	СШПФ упаковка	СШПФ подвоєння	Radix-2 РШПФ	split-radix РШПФ	split-radix ШПХ	Швидке ДПФ
1024	27652	15100	14848	13658	12804	12064	14580

2.3.3 Особливі випадки алгоритмів ШПФ

Алгоритм ШПФ може бути оптимізований для деяких особливих випадків, наприклад, коли потрібна лише частина виводу або коли на вході є велика кількість нулів або коли вхід не працює з двома або багатовимірними входами. У цьому підрозділі описано деякі особливі випадки алгоритмів ШПФ, які корисні для конкретних програм.

2.3.3.1 Обрізання ШПФ

Якщо послідовність даних містить 2^l ненульових точок даних з 2^m точок даних, де $m > l$, то відповідний ШПФ можна обчислити за допомогою обрізаного ШПФ, що забезпечує економію часу. Невелика модифікація алгоритму DIT-RX-2 дозволяє заощадити час приблизно $(m-l)/m$, де 2^m точки перетворюються, з яких лише 2^l – ненульові.

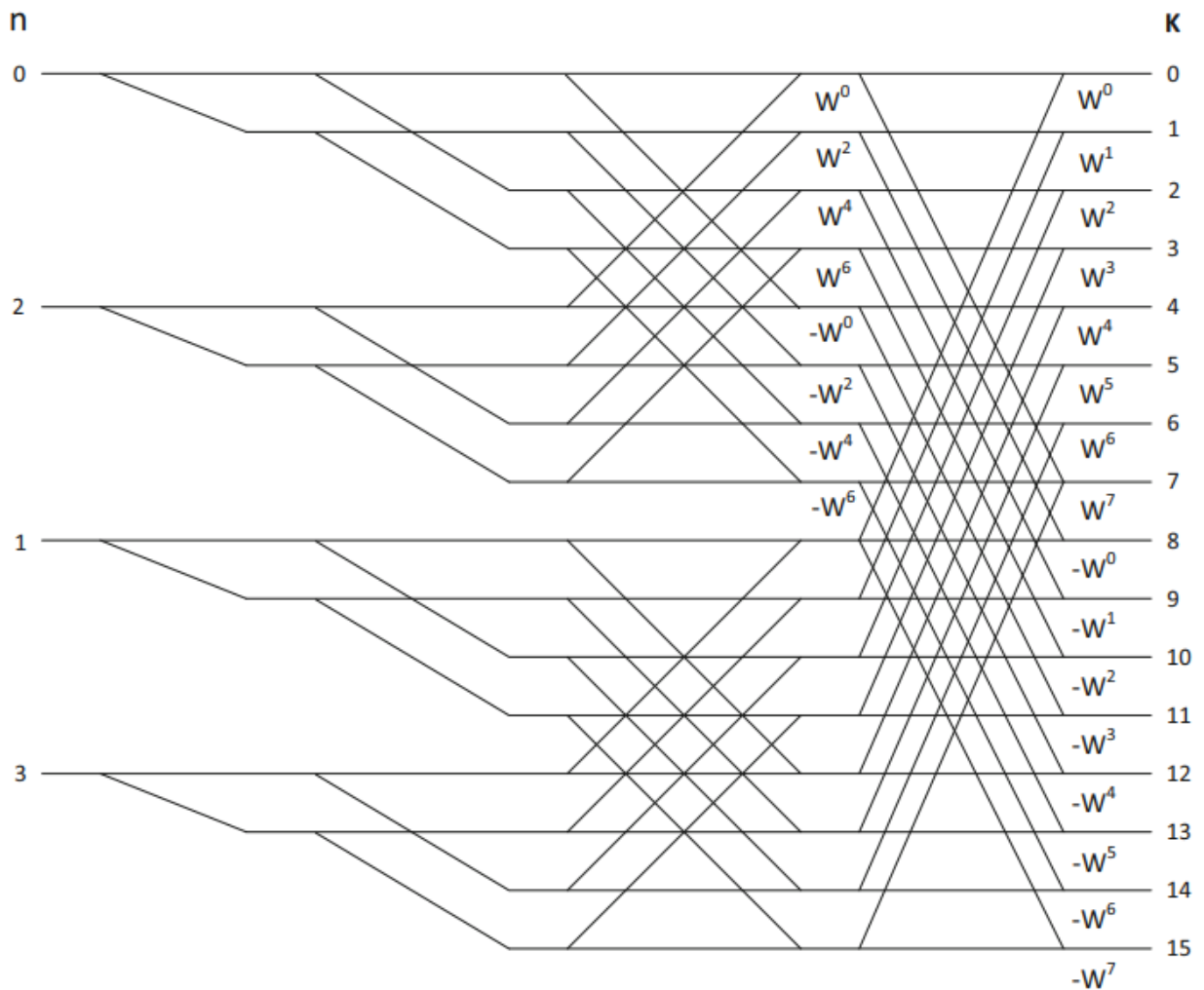


Рисунок 2.9 – Обрізання ШПФ

Обрізка ШПФ на $l = 2$, $m = 3$ показана на рисунку 2.9. Є чотири ненульові точки даних та три стадії. Обрізку застосовують на першій стадії, але другу та третю стадії не можна підрізати. Якщо застосовується обрізка, ми обчислюємо лише часткові метелики замість цілих метеликів. Загалом, якщо в наборі $2m$ точок даних є $2l$ ненульових точок даних, то кількість стадій, на яких можна застосувати обрізку $(m - l)$ стадій. Обрізка ШПФ застосовується, коли на вході відомо велика кількість нульових значень. Асимптотичний час запуску обрізаного ШПФ дорівнює $O(N \cdot \log z)$, де N - довжина ШПФ, z - кількість ненульових входів. Основний недолік обрізки полягає в тому, що послідовність даних повинна бути відома заздалегідь, щоб можна було знайти ненульові вхідні значення.

2.3.3.2 Алгоритм Герцеля

Алгоритм Герцеля може бути корисним для обчислення лише декількох компонентів, обраних частотою. Наприклад, для конкретних застосувань, таких як демодуляція клавіш зі зсувом частоти та розпізнавання двотональної багаточастотності, тонів вимагає лише декількох частот ДПФ. Алгоритм Герцеля може бути отриманий шляхом перетворення ДПФ рівняння. (1) в еквівалентну форму у вигляді згортки, яку можна ефективно реалізувати як цифровий фільтр, як показано на рисунку 2.10.

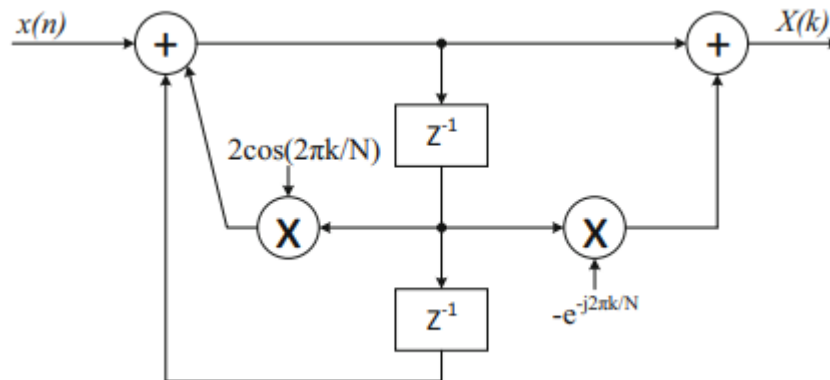


Рисунок 2.10 - Реалізація алгоритму Герцеля прямої форми

В алгоритмі Герцеля не потрібно оцінювати $X(k)$ при всіх N значеннях k , як у рівнянні. (1). Він може оцінювати $X(k)$ для будь-яких S -значень k , при цьому кожне значення ДПФ обчислюється рекурсивною системою форми рисунку 2.10 з відповідними коефіцієнтами. У цьому випадку загальне обчислення пропорційне $N S$. Цей алгоритм є більш кращим, коли S невеликий; однак, коли потрібні всі N значень $X(k)$, алгоритм ШПФ є більш ефективним, ніж алгоритм Герцеля.

2.3.3.3 Швидке перетворення Фур'є з рідким введенням

Час обчислення ДПФ, як правило, відповідає його розміру N . Однак у більшості застосувань, таких як спектральне зондування та обробка радіолокаційного сигналу, використовується лише декілька вибраних результатів ШПФ. Алгоритм обчислення цих коефіцієнтів перетворення Фур'є називається розрідженим ШПФ (РоШПФ), час виконання якого є сублінійним за розміром сигналу N .

Як правило, РоШПФ проходить у три етапи ідентифікації частотних розташувань головних елементів з великою величиною; оцінка коефіцієнтів цих елементів на першому кроці

(Оцінка частоти); видалення атрибуції результату Фур'є, обчисленого першими двома кроками, з вихідного сигналу (роздільна здатність зіткнення). Ці три етапи повторюються до тих пір, поки не знайдуться всі розріджені елементи. Складність додатково зменшується за рахунок віднімання реконструйованої часткової розрідженості від підвибірки сигналів, таким чином, що складність зворотного ШПФ також мінімізується. У [12] запропонована інша схема під назвою РоШПФ -ЗЧ (зниження частоти у часовій області), яка спочатку подає вибірку вихідного сигналу, а потім усі наступні операції проводяться на сигналах, що не мають вибірки.

2.3.3.4 Алгоритми безсилості двох довжин

Як правило, алгоритми потужності двох ШПФ використовуються традиційно для більшості програм обробки сигналів. Однак для деяких застосувань, таких як мультимедіа, вхідна довжина ДПФ відрізняється від потужності вдвічі. Такі програми вимагають швидкого алгоритму з меншою складністю обладнання. У цьому підрозділі описано популярний алгоритм ШПФ Winograd (АШПФВ) та алгоритм основного фактора (АОФ), які використовуються для інших випадків.

1) Алгоритм претворення Фур'є Виноград.

АШПФВ вимагає найменшої кількості множень серед практичних алгоритмів ШПФ для ДПФ середньої довжини. WFTA використовує картографування Good для претворення 1-D ДПФ довжини (N_1N_2) у довжину (N_1N_2) 2-D ДПФ. Структура алгоритмів малої довжини використовується для розміщення всіх множень у центрі загального алгоритму. На відміну від алгоритму Кулі – Тукі, АШПФВ спочатку обчислює попередні доповнення всіх суб- ДПФ, потім обчислює всі множення суб- ДПФ i , нарешті, обчислює після додавання всіх суб- ДПФ. Виноград показав, що ДПФ можна обчислити лише на множення $O(N)$. Однак цей алгоритм вимагає більшої кількості доповнень та великої пам'яті для більшої довжини. Таким чином, АШПФВ може бути ефективним лише у впровадженні ДПФ невеликих розмірів.

2) Алгоритм основного фактора.

Усунення множення коефіцієнту повертання для з'єднання двох розкладених ДПФ можна отримати, використовуючи алгоритм основного коефіцієнта для обчислення ШПФ. Це також називається алгоритмом Good Thomas ШПФ. Цей алгоритм використовує аналогічний підхід розділення та перемоги, де розміри розкладених ДПФ є відносно простими. Два числа є відносно простими або спільними простими, якщо їх спільний дільник один. У таких випадках формується спеціальне відображення індексів, засноване на китайській теоремі про залишки, щоб з'єднати ці розкладені ДПФ. Якщо довжину послідовності можна розподілити на два взаємно простих коефіцієнта $N = N_1N_2$, коефіцієнт ДПФ рівняння. (1) можна записати так:

$$X(k) = \underbrace{\sum_{n_1=0}^{N_1-1} \left(\underbrace{\sum_{n_2=0}^{N_2-1} x(N_2 n_1 + N_2 n_2) W_{N_2}^{n_2 k_2}}_{N_2 \text{ - точкове ДПФ}} \right) W_{N_1}^{n_1 k_1}}_{N_1 \text{ - точкове ДПФ}} \quad (2.42)$$

де не існує коефіцієнта повертання множників для з'єднання двох розкладених ДПФ.

2.3.3.5 Масштабовані ДПФ

У певних додатках, таких як ортогональне розподілення частоти, мультиплексування демодуляції та сучасна мікроскопія, вхідна довжина ДПФ дорівнює довжині- $q * 2^M$, де q - непарне число. Однак швидкі алгоритми для такої довжини послідовностей зазвичай вимагають складної обчислювальної структури і менш ефективні, ніж алгоритми ДПФ з двома довжинами. Для отримання ДПФ такої довжини послідовності часто застосовували техніку нульового набивання. Однак ця методика вимагає більше обчислень.

2.3.3.6 Багатовимірні ШПФ

Багатовимірні швидкі перетворення Фур'є (ШПФ в 2D або більше вимірах) використовуються в багатьох областях, таких як обробка зображень, прикладна фізика. Ці програми вимагають великої кількості обчислень.

Загальна форма багатовимірного ШПФ така:

$$X(u_1, u_2, \dots, u_m) = \sum_{v_1=0}^{N_1-1} \sum_{v_2=0}^{N_2-1} \dots \sum_{v_m=0}^{N_m-1} W_{N_1}^{u_1 v_1} W_{N_2}^{u_2 v_2} \dots W_{N_m}^{u_m v_m} x(v_1, v_2, \dots, v_m) \quad (2.43)$$

де $W_{N_k} = \exp\left(\frac{-2\pi j}{N_k}\right)$, $u_k=0, 1, \dots, u_k-1$; u_k - довжина k -го виміру $k=1, 2, \dots, m$ і $x(v_1, v_2, \dots, v_m)$ - це складні послідовності вхідних даних.

Рівняння (43) перетворюється в m одновимірних ШПФ наступним чином:

$$X(u_1, u_2, \dots, u_m) =$$

$$\sum_{v_1=0}^{N_1-1} W_{N_1}^{u_1 v_1} \sum_{v_2=0}^{N_2-1} W_{N_2}^{u_2 v_2} \dots \sum_{v_m=0}^{N_m-1} W_{N_m}^{u_m v_m} x(v_1, v_2, \dots, v_m) \quad (2.44)$$

Це забезпечує найпростіший алгоритм, коли кожен одновимірний ШПФ може бути обчислений ШПФ Кулі – Тьюкі, тому цей алгоритм відомий як алгоритм рядок-стовпчик.

Для багатовимірних ШПФ запропоновано декілька алгоритмів, такі як vector-radix алгоритми, алгоритми поліноміального перетворення та алгоритми розщеплення vector-radix. Ці алгоритми зменшують складність алгоритму рядок-стовпчик.

2.3.3.7 Квантове перетворення Фур'є

Закон Мура був послідовним протягом декількох десятиліть, але підтримувати темпи масштабування в останні роки стає все складніше. Для задоволення вимог щодо продуктивності та потужності в екскавальних системах, квантові комп'ютери можуть бути однією з альтернатив, яка, можливо, може запропонувати експоненціальне прискорення для певних типів обчислень.

У квантових комп'ютерах використовується квантове перетворення Фур'є (КПФ), яка подібна до ШПФ. Але КПФ працює на квантових бітах, а не на векторних елементах. Якщо елементи 2^p враховуються для обох перетворень, вони можуть приймати операції 2^{2^p} та $p(p+1)/2$ для обчислення ШПФ та КПФ відповідно. Порівнюючи кількість операцій, можна засвідчити, що КПФ є ефективнішим, ніж ШПФ. Нині значна увага приділяється дослідженням щодо впровадження алгоритмів КПФ.

Основні квантові комп'ютери розроблені в багатьох лабораторіях по всьому світу. Такі компанії, як Microsoft, IBM та Google, розробили власні прототипи. Однак ці прототипи дуже прості з лише невеликою кількістю кубітів. Квантова емуляція обладнання також має вирішальне значення при розробці практичних алгоритмів КПФ, перш ніж масштабний квантовий комп'ютер стане життєздатним.

2.4 Застосування ШПФ

Області застосування ШПФ дуже широкі та різноманітні. Більше того, з'являється багато нових додатків. Тут окреслено лише декілька важливих та популярних застосунків.

Додатки для обробки цифрового сигналу.

Спектральний аналіз сигналів - одне з найважливіших та основних програм ШПФ. На малюнку 11 показана блок-схема аналізатора спектру на основі ШПФ. Аналізатор ШПФ перетворює дані часової області в дані частотної області. Аналізуючи спектри вхідних сигналів

в частотній області, можна спостерігати невідомі параметри, такі як частота, амплітуда, спектри потужності та фазові параметри сигналу, які легко виявити у формі хвилі часової області.

Спектральний аналіз сигналів - одне з найважливіших та основних застосувань ШПФ. На рисунку 2.11 показана блок-схема аналізатора спектру на основі ШПФ. Аналізатор ШПФ перетворює дані часової області в дані частотної області. Аналізуючи спектри вхідних сигналів в частотній області, можна спостерігати невідомі параметри, такі як частота, амплітуда, спектри потужності та фазові параметри сигналу, які легко виявити у формі хвилі часової області.

Фізична інтерпретація спектру виходить з амплітудно-фазовими спектрами як

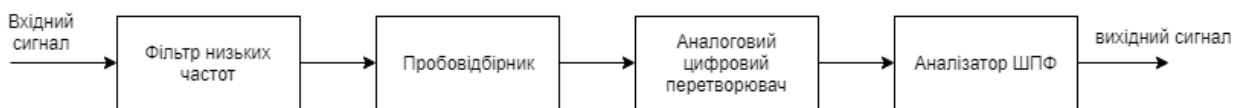


Рисунок 2.11 - Структурна схема аналізатора спектру ШПФ

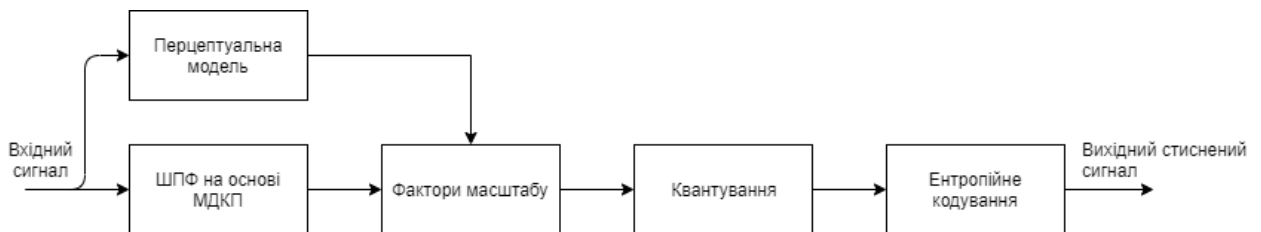


Рисунок 2.12 - Блок-схема аудіокодера

$$|X(k)| = \sqrt{\operatorname{Re}(X(k))^2 + \operatorname{Im}(X(k))^2}, \arg(x(k)) = \arctan\left(\frac{\operatorname{Im}(x(k))}{\operatorname{Re}(x(k))}\right) \quad (2.45)$$

Спектри амплітуди вказують на потужність сигналу в різній частоті. Спектр потужності випадкового сигналу можна обчислити шляхом підрахунку абсолютної величини складних компонентів ДПФ в сегменті сигналу як

$$P(k) = \frac{1}{N} |X(k)|^2 \quad (2.46)$$

Додатки для спілкування.

ШПФ є важливим функціональним блоком у сучасних системах зв'язку, зокрема для застосувань у системах мультимплексування з ортогональним частотним розділенням (МОЧР), таких як цифрове мовлення, всевітня сумісність для доступу до мікрохвильового доступу (WiMAX) [13], стандарти IEEE 802.11.

Потік послідовних вхідних даних $(d_0, d_1, \dots, d_{n-1})$ перетворюється в паралельний потік даних і відображається в символи з картографічного сузір'я. Нехай символи відображаються за допомогою бінарного фазового зсуву, а вихід представлений у вигляді X_0, X_1, \dots, X_{n-1} . Потім ці символи надаються ШПФ, який генерує цифровий символ МОЧП з N ортогональними піднесучими. Вихід ШПФ дорівнює x_0, x_1, \dots, x_{n-1} . Потім цей вихід серіалізується і перетворюється на аналоговий сигнал, використовуючи цифро-аналоговий перетворювач (ЦАП). Повний символ МОЧП $x(t)$ передається через канал. З боку приймача цей символ МОЧП перетворюється назад в паралельний потік і перетворюється на цифровий сигнал за допомогою аналогового цифрового перетворювача (АЦП). ШПФ використовується для декодування піднесучих OFDM, а потім відображається на зворотній перетворювач сузір'я. Нарешті, отриманий сигнал серіалізується для отримання вхідних даних $(\widehat{d}_0, \widehat{d}_1, \dots, \widehat{d}_{n-1})$. Основними перевагами МОЧП в системах бездротового зв'язку є висока ефективність пропускну здатності, стійкість до РЧ-перешкод і надійність багат шарового згасання.

Програми для обробки зображень.

ШПФ використовується в медичній візуалізації для фільтрації зображень, аналізу зображення та реконструкції зображення. У зображенні Фур'є зображення, що використовують ШПФ, спектральна величина і фаза, як правило, грають різні ролі. Оппенгейм та Лім продемонстрували важливість фазового спектру для спектру величин. Далі показано, що фазова інформація відіграє важливішу роль, ніж величина. Структурна схема для реалізації фазової кореляції (ФК) за допомогою 2D-ШПФ та 2D-ЗШПФ показана на рисунку 2.13. Кореляція між фазовими версіями двох зображень, які слід вирівняти, використовується для узгодження зображень. Деякі важливі програми, засновані на відповідності зображень на основі ШПФ, включають розпізнавання обличчя, розпізнавання райдужної оболонки, розпізнавання долонного друку, відповідність відбитків пальців та відповідність форми хвилі. В роботі [14] запропоновано новий алгоритм оцінки якості зображення, заснований на фазі та величині двовимірного ШПФ, який вказував, що для ідеальної реконструкції зображення необхідні як амплітуда, так і фазовий спектр.

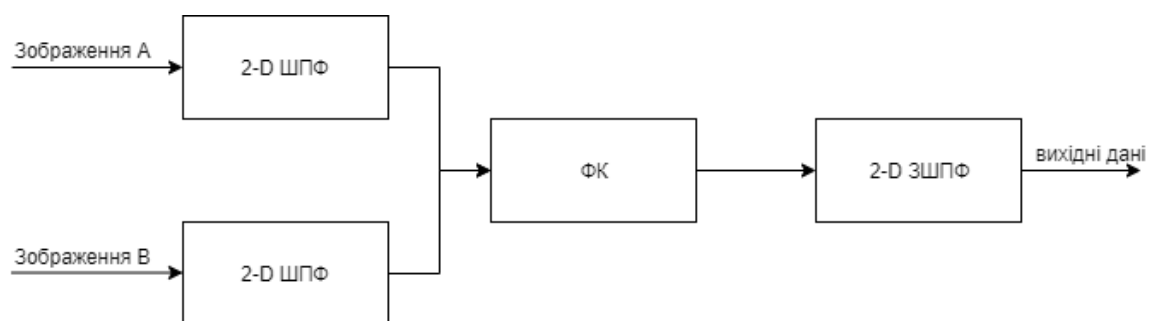


Рисунок 2.13 - Структурна схема фазового співвідношення

Фріго та ін. [8] розробили найшвидше перетворення Фур'є на Заході (ШПФЗ), яке містить бібліотеку ШПФ (C код). Ця бібліотека використовується для обчислення ДПФ в 1-D або більше вимірах різного вхідного розміру як для вхідних даних, так і для реального і комплексного значення. Програми для обробки зображень використовують ШПФЗ. Однак сьогоднішньому додатку потрібно обробити величезну кількість наборів даних, що вимагають швидкого обчислення ШПФ. Алгоритм Sparse ШПФ [15] вирішує цю проблему, надаючи сублінійну складність, яка використовує розрідженість в області Фур'є (враховуючи лише частоти з інформацією). Швидке обчислення робить рідкий ШПФ багатообіцяючим інструментом для багатьох застосувань, що вимагають великого об'єму даних, таких як 4D світлові поля [16] та 2D спектральна магнітно-резонансна спектроскопія (MPC) [17].

Інше застосування.

Множення довгих цілих чисел широко використовується в алгоритмах криптографії з відкритим ключем, таких як RSA (Rivest, Shamir та Adelman) [18], які використовують арифметику з цілими числами, що мають сотні цифр. Множення довгих цілих чисел використовується в деяких інших програмах, де потрібні високоточні обчислення і переповнення небажано. Метод множення в шкільній книзі передбачає складність у часі на $O(n^2)$ для множення n -бітного цілого числа "a" з m -бітовим цілим числом "b". У [19] Шонгаге та Страссен використовували алгоритм швидкого множення поліномів на основі ШПФ для вчасного обчислення цілих добутків $O(n \log_2 n \log(\log n))$. Потік даних алгоритму Шонгаге та Страссена показаний на рисунку 2.14.

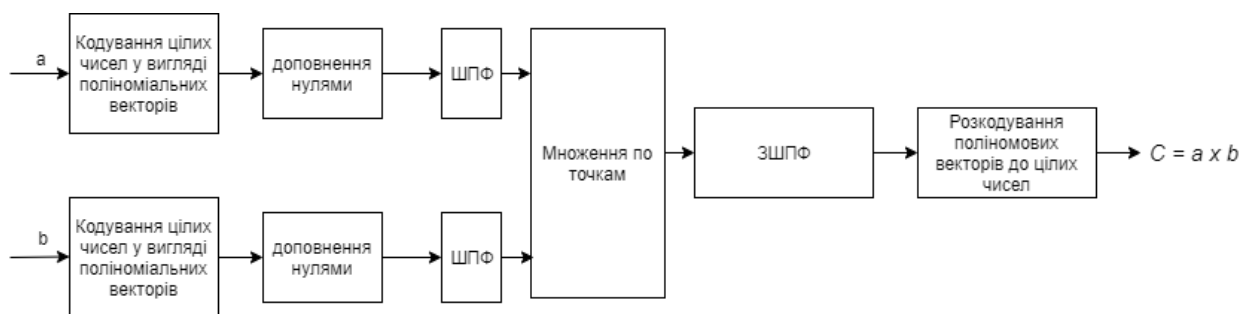


Рисунок 2.14 - Потік даних алгоритму Шонгаге та Страссена за допомогою ШПФ та ЗШПФ

Пару цілих чисел a і b у поліноміальне представленні можна записати як

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1} \quad (2.47)$$

$$q(x) = b_0 + b_1x + b_2x^2 + \dots + b_{m-1}x^{m-1} \quad (2.48)$$

де x - базове ціле число.

Зазначені вище рівняння можна представити у векторному форматі як

$$p = [a_0, a_1, a_2, \dots, a_{n-1}] \quad (2.49)$$

$$q = [b_0, b_1, b_2, \dots, b_{m-1}] \quad (2.50)$$

Коли вектори різної довжини, коротший доповнюється нулями, поки довжина не стане такою ж, як і інші. Для використання ШПФ radix-2 довжина повинна бути потужністю два. Тому, коли довжина векторів не є потужністю двох чисел, нулі забиваються ними, поки довжина не стане наступним цілим числом двох потужностей. Нехай p' і q' , відповідно, - вектори, одержані нульовим накладенням векторів p і q .

2.5 Висновки до розділу 2

Краса ШПФ - це ефективне обчислення ДПФ. Пряме обчислення ДПФ вимагає обчислення порядку $O(N^2)$, тоді як ШПФ включає лише операції $O(N \log_2 N)$. Протягом останніх 50 років було розроблено кілька алгоритмів ШПФ, які використовуються в різних програмах у широких сферах обробки сигналів, зв'язку та обробки зображень. У цьому розділі представлено короткий огляд важливих досягнень в алгоритмах та застосуванні ШПФ.

РОЗДІЛ 3

ПРОЕКТУВАННЯ ПІДСИСТЕМИ РОЗПІЗНАВАННЯ МУЗИКИ

3.1 Структура підсистеми

Перш, ніж приступити до створення системи необхідно сформулювати її основне завдання і вирішити, яку цінність вона представлятиме для користувачів. Іншими словами, яку функцію дана система буде виконувати.

В контексті даної магістерської роботи було виділено наступні цілі:

1. Дослідження схожих підсистем.
2. Ознайомлення з інструментами та ресурсами веб розробника.
3. Розробка системи для розкладення аудіофайлів на хеші.
4. Розробка системи для автоматичного розпізнання аудіо, яке вже є у БД.

Підсистема повинна функціонувати на операційній системі Windows (або Linux) і надавати можливість користувачу розпізнавати невідомі для нього пісні. Загальна структура роботи системи зображена на рисунку 3.1.

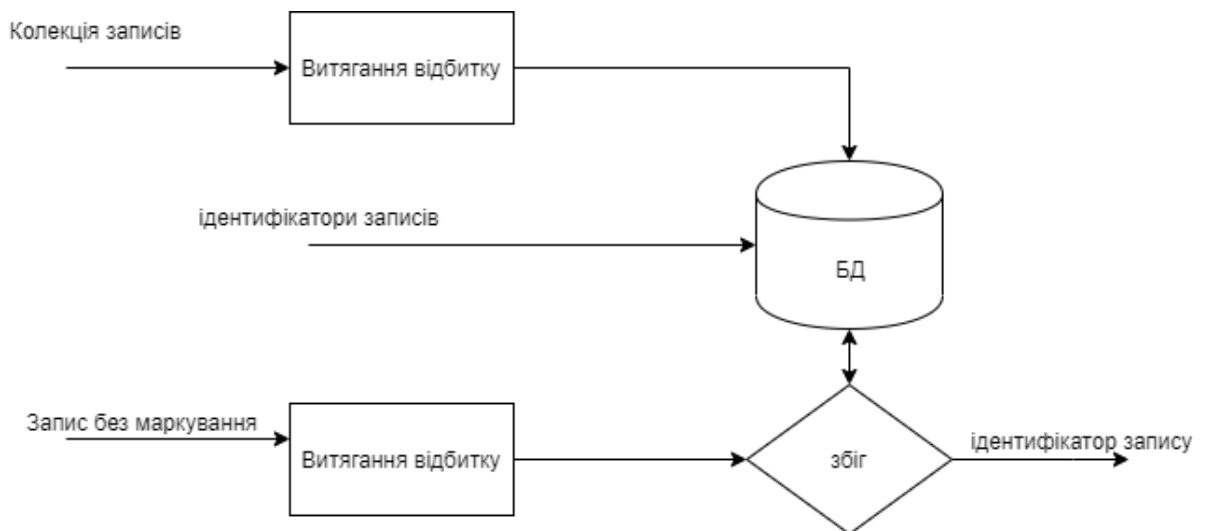


Рисунок 3.1 - Загальна структура роботи системи

3.2 Музика як сигнал

Музика, кодується у цифровий вид як просто довгий список чисел. У нестисненому .wav файлі дуже багато цих чисел - 44100 в секунду на канал. Це означає, що 2-хвилинна пісня має більше десяти з половиною мільйонів вибірок.

$$2 \text{ хвилини} * 60 \text{ секунд} * 44100 \text{ вибірок у секунду} * 2 \text{ канали} = 10584000$$

Канал - це окрема послідовність зразків, яку може грати динамік. Два навушника - це "стерео" або два канали. Один канал називається "моно". Сьогодні сучасні системи об'ємного звучання можуть підтримувати ще більше каналів. Але якщо звук не записаний або змішаний з однаковою кількістю каналів, додаткові динаміки зайві, і деякі динаміки будуть просто відтворювати той самий потік вибірок, що й інші динаміки.

3.3 Вибірки (Sampling)

Швидкість вибірки або частота дискретизації визначає кількість вибірок в секунду (або на іншу одиницю), взяті з безперервного сигналу для отримання дискретного або цифрового сигналу. Для сигналів часової області, таких як форми хвиль для звуку (та інших типів аудіовізуального контенту), частоти вимірюються в герцах (Гц). Теорема дискретизації Найкіста – Шеннона (принцип Найквіста) стверджує, щоб відновити сигнал за його відліками без втрат, необхідно, щоб частота дискретизації була хоча б у два рази більша за максимальну частоту первинного неперервного сигналу. Наприклад, якщо звуковий сигнал має верхню межу в 20000 Гц (приблизна верхня межа слуху людини), частота дискретизації, що перевищує 40 000 Гц (40 кГц), уникне згладжування і дозволить теоретично досконало відновити.

Багато органів влади щодо збереження звукових записів, як-от Міжнародна асоціація звукових та аудіовізуальних архівів (IASA), рекомендують частоти дискретизації, що кодують аудіо поза діапазону слуху людини, тобто частоти, що перевищують 40 кГц. У документі IASA TC-04 зазначається, що більш високі показники вибірки "можуть бути вигідними для багатьох типів вмісту. Чистий ефект від вищої швидкості дискретизації та технології перетворення покращує якість звуку в ідеальному діапазоні слуху для людини. Невідповідні та небажані артефакти притаманні в записі також є частиною звукового документа, незалежно від того, чи вони були притаманні виробництву запису, або згодом були додані до оригінального сигналу шляхом неправильного поводження або поганого зберігання. Обидва повинні зберігатися з максимальною точністю. Для деяких типів шуму, Частота дискретизації, що перевищує 48 кГц, може бути вигідною. Для будь-яких більш високих частот дискретизації IASA рекомендує 96 кГц.

Що стосується запису аудіо, прийнятим правилом є те, що частотах понад 22000 Гц упускаються, оскільки люди не можуть навіть чути частоти понад 20 000 Гц. Таким чином, за Найквістом, необхідно зробити вибірку двічі:

$$\begin{aligned} \text{Необхідні зразки в секунду} &= \text{Найвища частота} * 2 = \\ &= 22050 * 2 = 44100 \end{aligned} \quad (3.1)$$

3.4 Спектрограма

Спектрограма (рис. 3.2) - зображення, що показує залежність спектральної щільності потужності сигналу від часу. Спектрограми застосовуються для ідентифікації мови, аналізу звуків тварин, в різних областях музики, обробці мови, сейсмології і в інших областях.

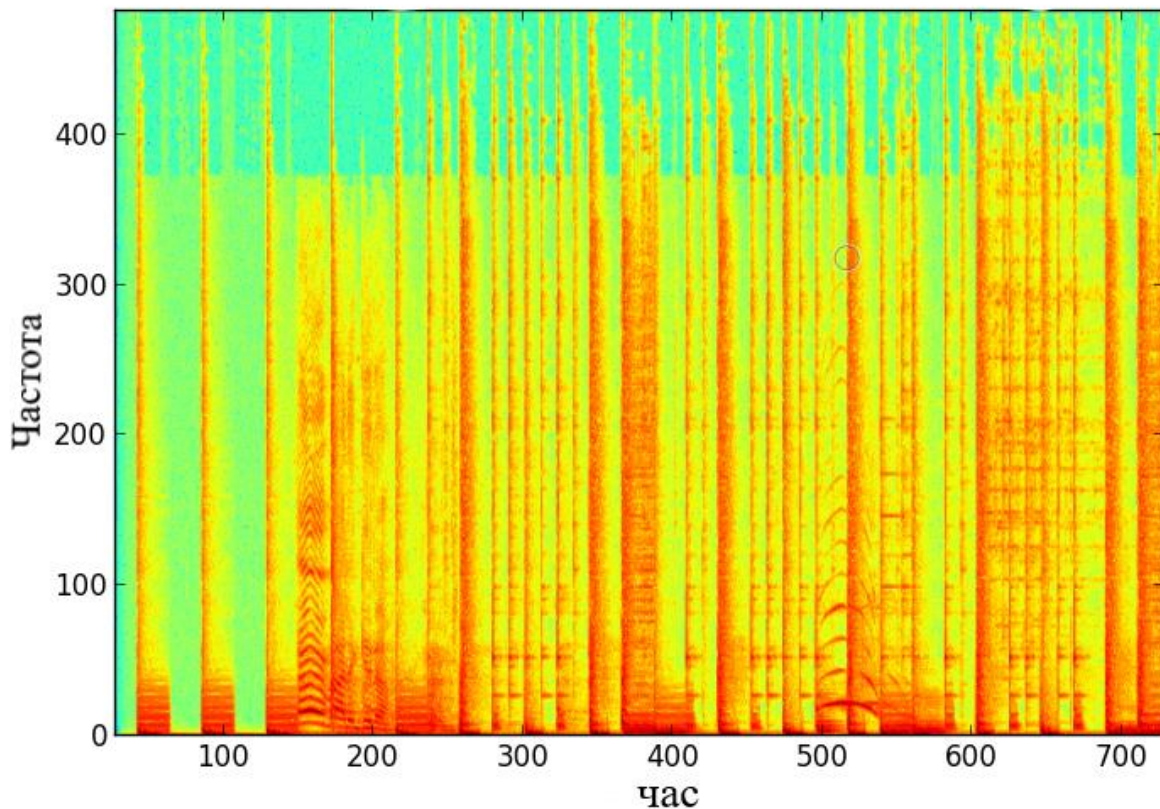


Рисунок 3.2 – Спектрограма аудіо

Спектрограма це просто двовимірний масив з амплітудою як функції часу і частоти. ШПФ показує нам силу (амплітуду) сигналу саме на цій частоті, даючи нам колонку. Якщо зробити це досить разів за допомогою розсувного вікна ШПФ, можна отримати 2D спектрограму масиву.

Важливо зауважити, що значення частоти та часу дискретні, кожен з яких представляє собою "відро", а амплітуди - реально оцінені. Колір показує реальне значення (червоне -> вище, зелене -> нижче) амплітуди за дискретизованою (часом, частотою) координатою.

То як це допомагає розпізнавати звук? Необхідно використовувати цю спектрограму для унікальної ідентифікації цієї пісні. Проблема полягає в тому, що при використанні, наприклад, телефону для розпізнавання музики по радіо (чи у людному місці), з'явиться шум. Необхідно знайти надійну спосіб зафіксувати унікальні «відбитки» від звукового сигналу.

Перш, ніж знаходити піки, необхідно відфільтрувати спектрограми для видалення шуму. Для фільтрації було використано функцію бібліотеки `scipy` - `maximum_filter`. Вона розраховує багатовимірний максимальний фільтр.

```
from scipy import ndimage, misc
import matplotlib.pyplot as plt
fig = plt.figure()
plt.gray() # показати відфільтрований результат у відтінках сірого
ax1 = fig.add_subplot(121) # ліва сторона
ax2 = fig.add_subplot(122) # права сторона
ascent = misc.ascent()
result = ndimage.maximum_filter(ascent, size=20)
ax1.imshow(ascent)
ax2.imshow(result)
plt.show()
```

Рисунок 3.3 – Приклад роботи функції `maximum_filter`

Приклад роботи фільтру:

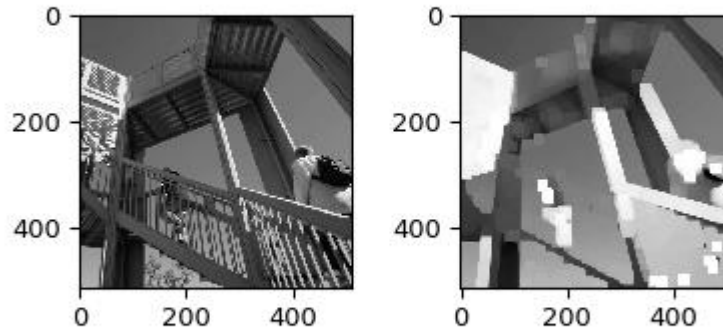


Рисунок 3.4 – Приклад роботи фільтру

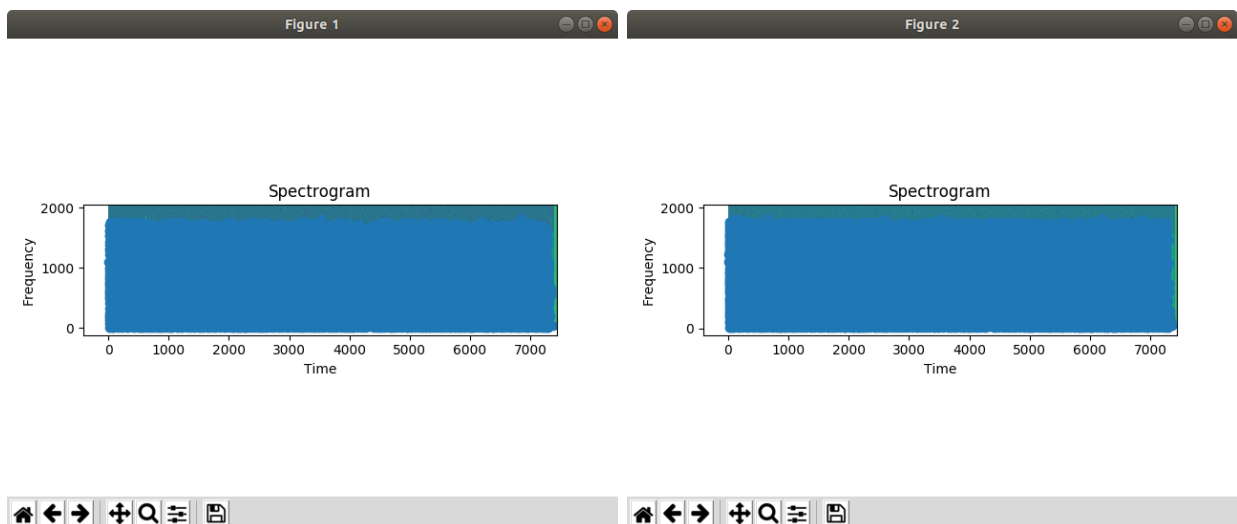


Рисунок 3.5 – Відфільтровані спектрограми першого і другого каналів аудіозапису

Після фільтрації спектрограми можна приступати до знаходження піків.

3.5 Знаходження піків

Тепер, після отримання спектрограми аудіосигналу, є можливість почати зі знаходження «піків» за амплітудою. Пік визначається як пара (час, частота), відповідаючи значенню амплітуди, що є найвищою в локальній «околиці» навколо нього. Інші пари (час, частота) навколо мають меншу амплітуду і, слідуючи, менше схильні до впливу шуму.

Пошук піків - це сама проблема. У підсумку було оброблено спектрограму як зображення і використовував інструментарій обробки зображень і методи scіру щоб знайти піки. Комбінація фільтра верхніх частот (підкреслюючи високі амплітуди) і структури scіру локальних максимумів зробили свою справу.

Після того, як було витягнуто ці стійкі до шуму піки, знайдено точки «інтересу» в пісні, які ідентифікують її. Як тільки піки знайдено, спектрограму необхідно ефективно «здавити». Амплітуди досягли своєї мети і більше не потрібні. Піки зображено на рисунку 3.6.

3.6 Хешування відбитків

Наступним кроком буде хешування сусідніх найвищих піків. Створювати хеш лише за одним значенням піку немає сенсу, бо, є вірогідність, що в інших аудіозаписах можуть бути схожі піки. Щоб такого не сталося, необхідно об'єднаємо піки в відбитки! У даній дипломній роботі робиться це за допомогою хеш-функції.

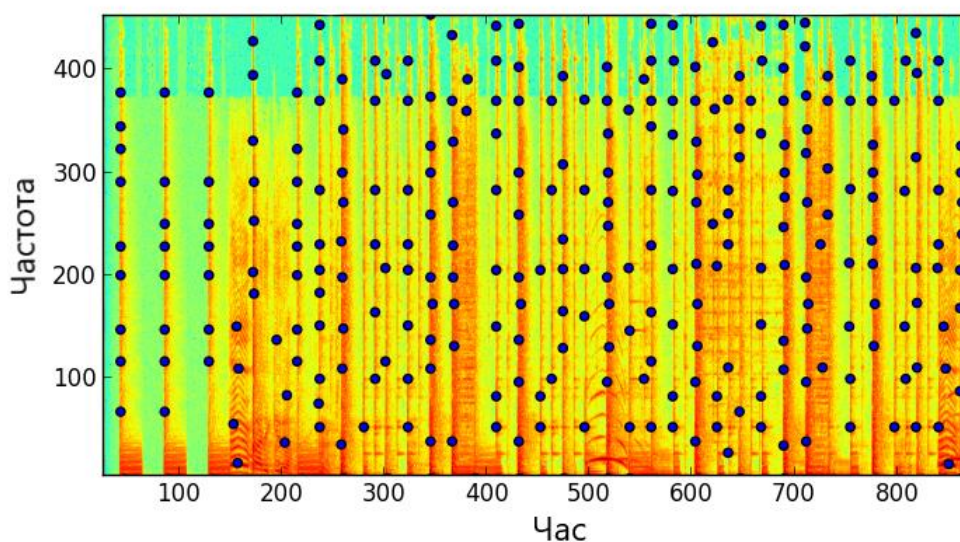


Рисунок 3.6 – Спектрограма аудіо з поміченими піками

Хеш - функція приймає ціле введення і повертає інше ціле число в якості висновку. Привабливість у тому, що хороша хеш-функція буде повертати не тільки одне і те ж ціле число кожного разу, коли вхідні дані однакові, але також і те, що далеко не всі різні вхідні дані будуть мати однаковий вихід.

Дивлячись на піки спектрограми і комбінуючи пікові частоти і різницю в часі між ними (рис 3.7), можна створити хеш, який представляє унікальний відбиток для цієї пісні. Загальний зміст хешу представлено нижче.

хеш (частота першого піку | частота другого піку | різниця у часі між піками) = хеш-значення відбитків

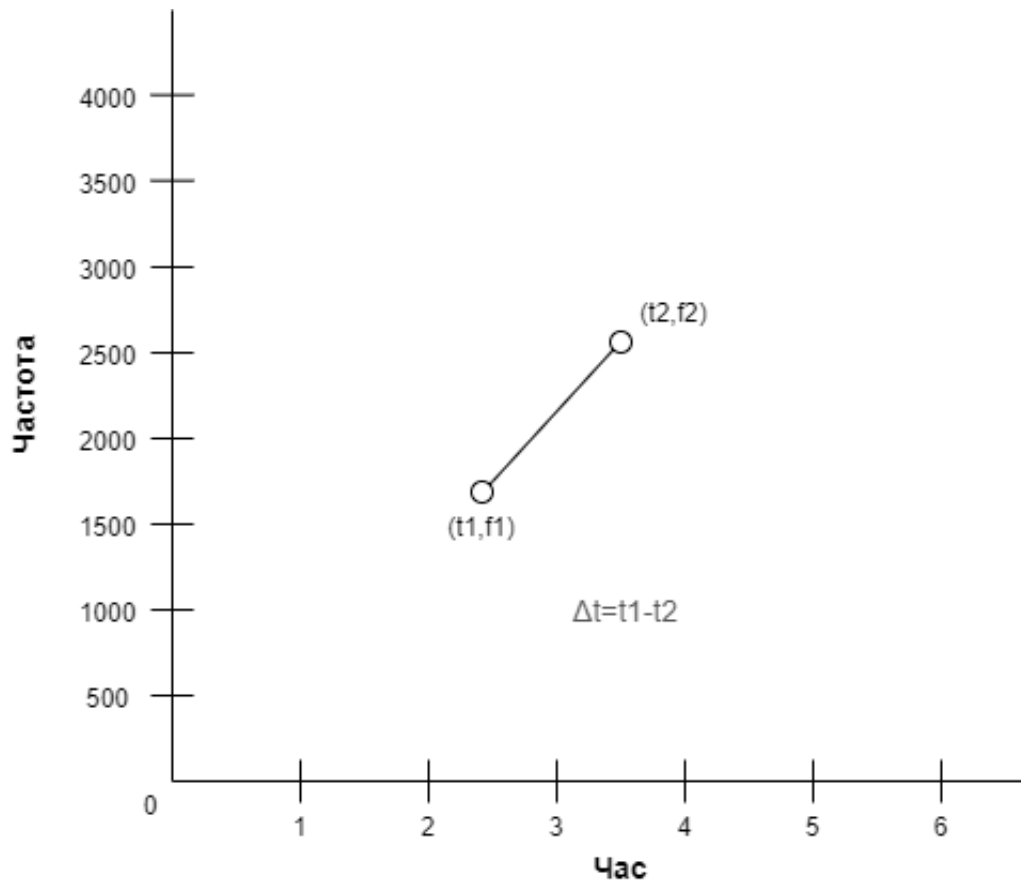


Рисунок 3.7 – Приклад вибірки двох сусідніх піків

Завдяки такій кількості відбитків, потрібно скоротити дисковий простір для зберігання значення. Для хешу відбитків було використано хеш SHA-1, а потім його необхідно скоротити до половини його розміру (лише перші 20 символів). Це скорочує використання байтів наполовину на один хеш: $\text{char}(40) \Rightarrow \text{char}(20)$ переходить від 40 байт до 20 байт.

Приймаючи до уваги більш одного значення піку, створюються відбитки, які мають меншу ентропію і, отже, містять більше інформації. Таким чином, вони є більш потужними ідентифікаторами пісень, оскільки вони будуть менше стикатися.

Візуалізація того, що відбувається за допомогою збільшеної анотованої спектрограми, показаної нижче:

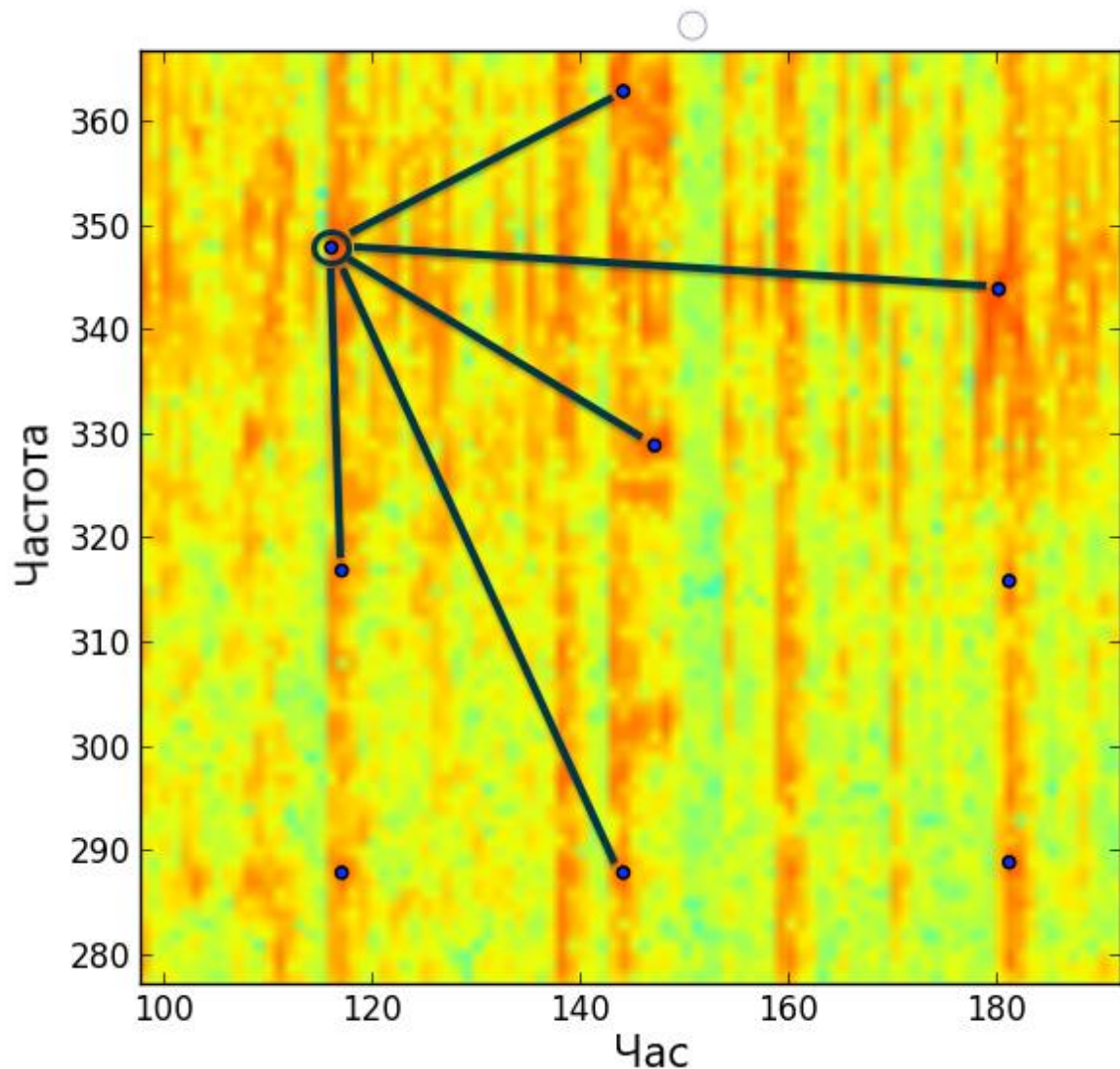


Рисунок 3.8 – Спектрограмма аудіо з поміченим «сузір'ям» піків

Система порівнює ці групи піків як свого роду «сузір'я» піків, які використовуються для ідентифікації пісні. Насправді вона використовує пари піків поряд з дельтою часу між ними. Можна уявити безліч різних способів угруповання точок і відбитків. З одного боку, більше піків в відбитку означає більш рідкісний відбиток, який більш сильно ідентифікував би пісню. Але більша кількість піків також означає меншу стійкість до шуму.

Тепер є можливість розпочати роботу з такою системою. Аудіо-система відбитків має два завдання:

- Розібрати нову пісню на хеші.
- Розпізнати невідомі пісні, шукаючи їх у базі пісень які були внесені до БД раніше.

3.7 Послідовність роботи програми

Для функціональності БД було використано систему керування реляційними базами даних MySQL. База даних буде містити дві таблиці:

- відбитки (fingerprints)
- пісні (songs)

У таблиці «відбитки» є не тільки хеш і ідентифікатор пісні, а й зміщення. Це відповідає часовому вікну спектрограми, звідки походить хеш. Це знадобиться пізніше, коли потрібно буде фільтрувати відповідні хеші. Тільки хеші, які "вирівнюють", будуть з істинного сигналу, який треба ідентифікувати (докладніше про це в підрозділі "Вирівнювання відбитків" нижче).

```
CREATE TABLE `fingerprints` (  
  `id` INTEGER PRIMARY KEY AUTOINCREMENT,  
  `song_fk` INTEGER,  
  `hash` TEXT,  
  `offset` INTEGER  
);
```

У таблиці «пісні» буде зберігатися інформація про аудіозапис.

```
CREATE TABLE songs (  
  id INTEGER PRIMARY KEY AUTOINCREMENT,  
  name TEXT,  
  filehash TEXT  
);
```

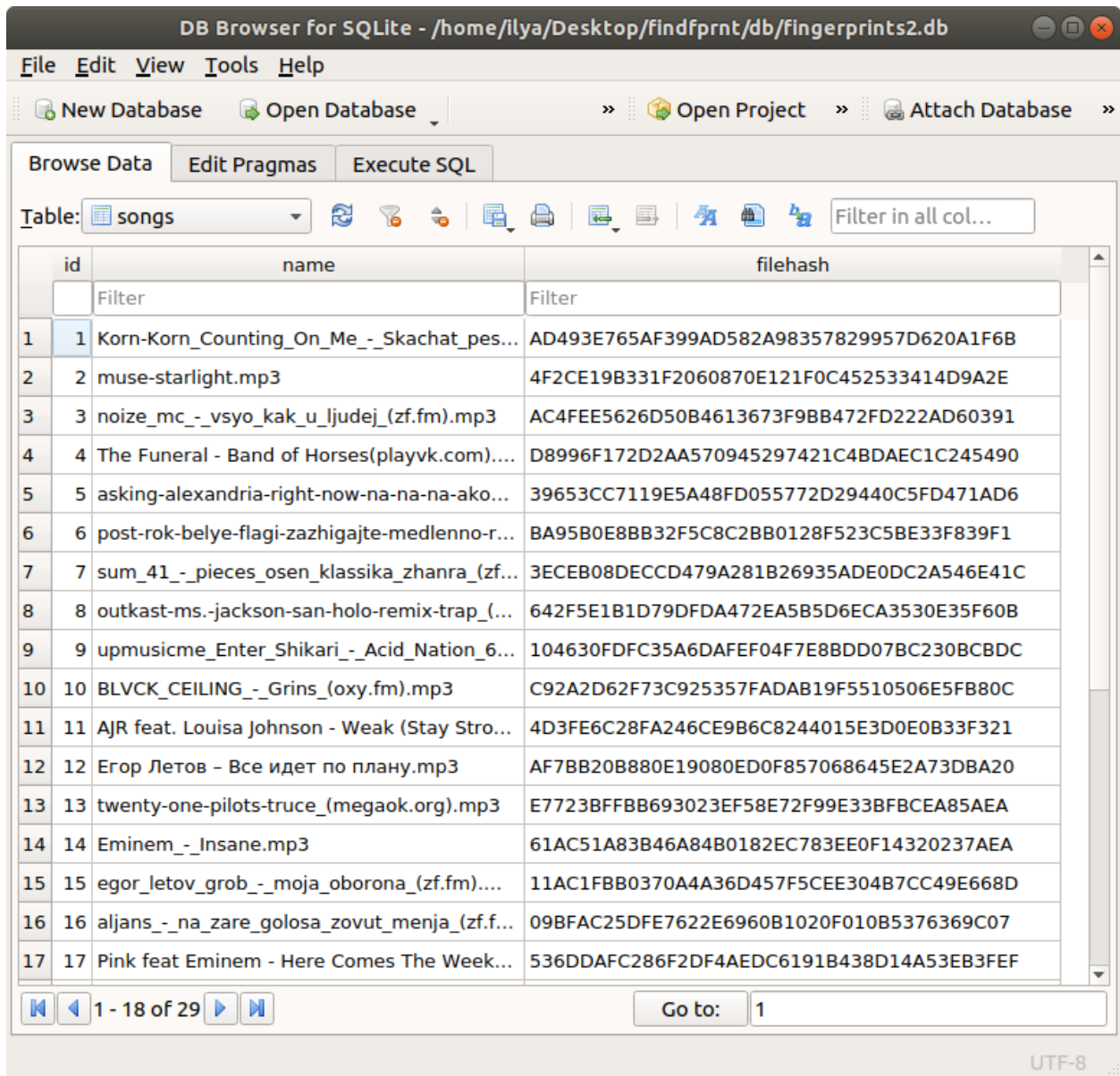


Рисунок 3.9 – Вікно програми DB Browser, у якому виведено таблицю «songs»

Для спрощення роботи звичайному користувачу для створення таблиць у БД необхідно лише запустити скрипт «reset-database.py». Після запуску скрипту, таблиці «відбитки» і «пісні» будуть видалені (якщо вони існували) і знову створені (рис. 3.12).

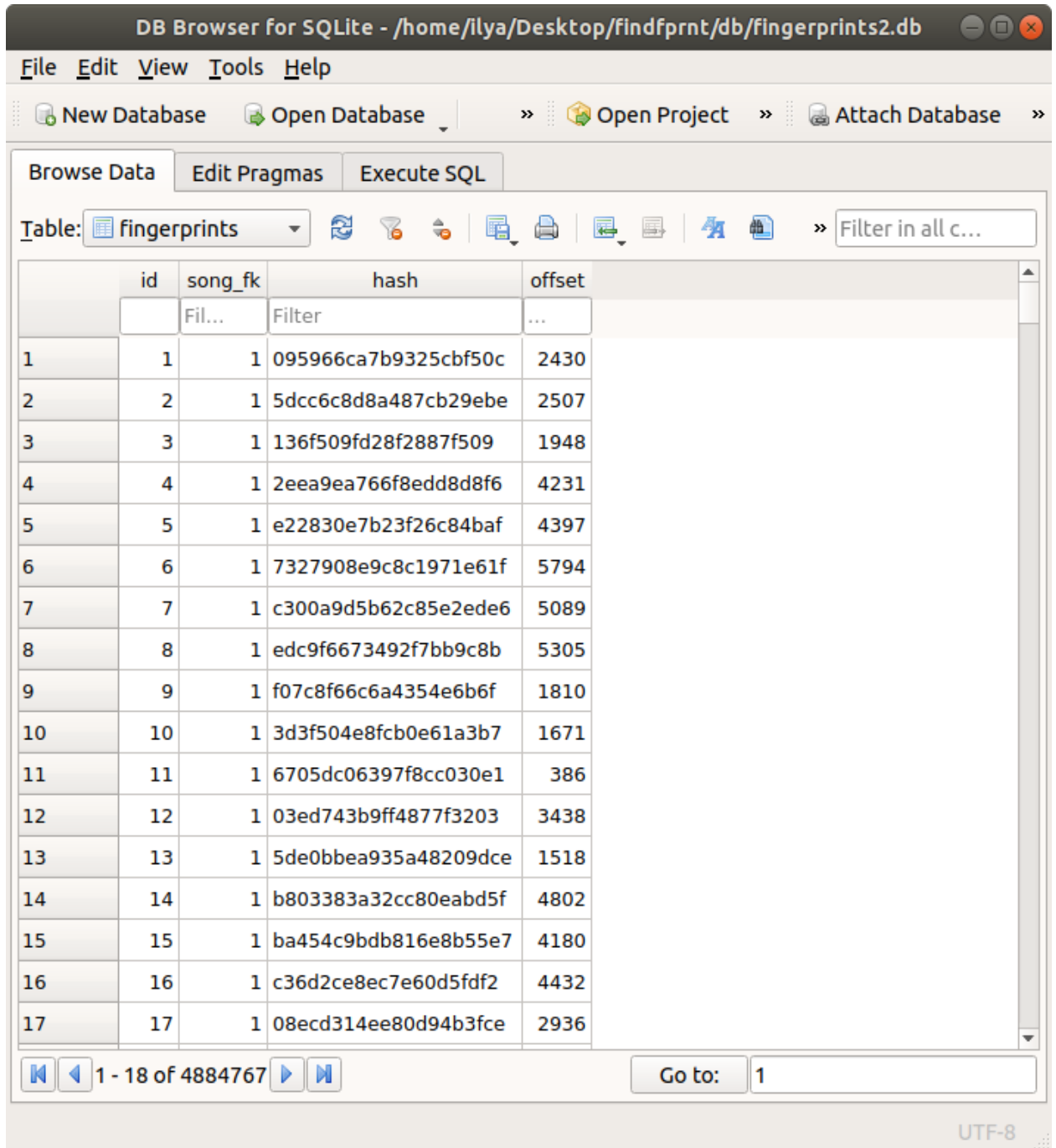


Рисунок 3.10 – Вікно програми DB Browser, у якому виведено таблицю «fingerprints»

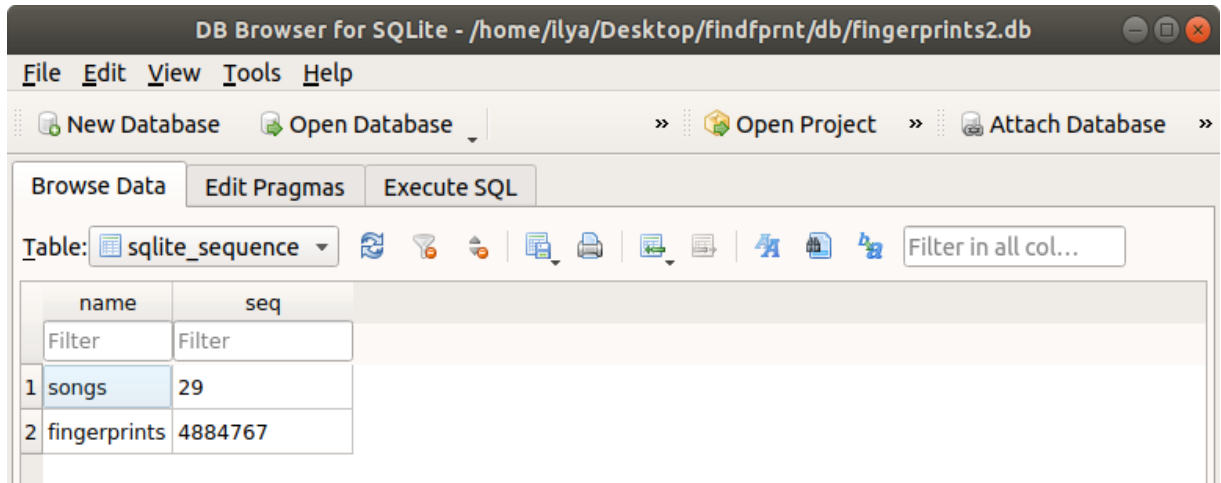


Рисунок 3.11 – Вікно програми DB Browser, у якому виведено таблицю «sqlite_sequence»

The screenshot shows a terminal window with the prompt 'ilya@ilya-System: ~/Desktop/findfprnt'. The terminal output shows the execution of 'python reset-database.py'. The script successfully connects to the SQLite database, removes and recreates the 'songs' and 'fingerprints' tables, and then closes the connection.

```

ilya@ilya-System: ~/Desktop/findfprnt
File Edit View Search Terminal Help
ilya@ilya-System:~/Desktop/findfprnt$ python reset-database.py
sqlite - connection opened
removed db.songs
created db.songs
removed db.fingerprints
created db.fingerprints
done
sqlite - connection has been closed
ilya@ilya-System:~/Desktop/findfprnt$

```

Рисунок 3.12 – Результат роботи скрипту

Перш ніж запускати розпізнавання аудіо, необхідно заповнити створені таблиці БД даними. Для цього необхідно у папку mp3 перенести и mp3 файли (саме ці пісні і будуть розкладені на хеші, за якими система і буде встановлювати яка саме пісня програвється).

```

ilya@ilya-System: ~/Desktop/findfprnt
File Edit View Search Terminal Help
канал 1/2, получено 103901 хешей
поиск отпечатков для 2/2-го канала
канал 2/2, получено 105777 хешей
В БД сохранено 203146 хешей
* id=27 каналов=2: justice_vs_simian_-_we_are_your_friends_ost_128_udarov_serdc
a_v_minutu_(zv.fm).mp3
анализ новой песни..
поиск отпечатков для 1/2-го канала
канал 1/2, получено 128079 хешей
поиск отпечатков для 2/2-го канала
канал 2/2, получено 127295 хешей
В БД сохранено 192956 хешей
* id=28 каналов=2: my-chemical-romance-helena(mp3-top.info).mp3
анализ новой песни..
поиск отпечатков для 1/2-го канала
канал 1/2, получено 91273 хешей
поиск отпечатков для 2/2-го канала
канал 2/2, получено 91735 хешей
В БД сохранено 168685 хешей
* id=29 каналов=2: Data feat. Benny Sings - Don't Sing [midi-mp3.org].mp3
анализ новой песни..
поиск отпечатков для 1/2-го канала
канал 1/2, получено 110271 хешей
поиск отпечатков для 2/2-го канала
канал 2/2, получено 108129 хешей
В БД сохранено 210812 хешей
Завершено!
sqlite - connection has been closed
ilya@ilya-System:~/Desktop/findfprnt$ ^C
ilya@ilya-System:~/Desktop/findfprnt$

```

Рисунок 3.13 – Результат успішного внесення до БД 22 аудіо

Після перенесення файлів необхідно у консолі перейти до загальної папки проекту і запустити скрипт «collect-fingerprint-of-songs». Після чого у консолі буде виводитися інформація про кількість створених хешей для кожного каналу і їх загальну кількість для кожної пісні (рис. 3.14). Для перегляду загальної кількості всіх хешей і аудіофалів, що містяться у БД, необхідно запустити скрипт «get-database-stat.py».


```

ilya@ilya-System: ~/Desktop/findfprnt
File Edit View Search Terminal Help
ilya@ilya-System:~/Desktop/findfprnt$ python get-database-stat.py
sqlite - connection opened

* Всего: 28 песен (4691027 отпечатков)
** id=30 aljans_-_na_zare_golosa_zovut_menja_(zf.fm).mp3: 274643 хешей
** id=26 rammstein-unheilig-freiheit_(mp3CC.biz).mp3: 260893 хешей
** id=27 post-rok-belye-flagi-zazhigajte-medlenno-raketnymi-dvigatelyami-v-te
lezhhkah-iz-supermarketa_(Cool.DJ).mp3: 253403 хешей
** id=19 sum-41-with-me(zvuqa.net).mp3: 243345 хешей
** id=15 Bi-2-feat-0xxxymiron_-_Pora-Vozvraschatsya-Domoy.mp3: 242217 хешей
** id=25 rammstein_-_donaukinder_deti_dunaja_(zf.fm).mp3: 224327 хешей
** id=2 M83 - Gone.mp3: 224073 хешей
** id=16 Korn-Korn_Counting_On_Me_-_Skachat_pesnyu_ili_slushat_onlayn_korn_ko
unting_on_me_skachat_mp3.mp3: 219180 хешей
** id=29 i_see_stars_murder_mitten.mp3: 203146 хешей
** id=24 Pink feat Eminem - Here Comes The Weekend (promusic.me).mp3: 202741
хешей
** id=21 Sia_-_Unstoppable_(mp3-muzon.com).mp3: 189007 хешей
** id=5 30 Seconds To Mars - From Yesterday.mp3: 186136 хешей
** id=20 salem_-_frost.mp3: 176562 хешей
** id=1 Monika-Linkyte_-_Po-Dangum.mp3: 173244 хешей
** id=18 splin_-_ves_etot_bred_(zvukoff.ru).mp3: 164051 хешей
** id=10 noize_mc_-_vsyo_kak_u_ljudej_(zf.fm).mp3: 159411 хешей
** id=11 egor_letov_grazhdanskaja_oborona_-_vse_idet_po_planu_1988_(zf.fm).mp
3: 144091 хешей
** id=28 quest_pistols_-_belaja_strekoza_ljubvi_(zvukoff.ru).mp3: 137882 хеше
й
** id=12 grazhdanskaja_oborona_grob_-_moja_oborona_(zf.fm).mp3: 120919 хешей
** id=23 salem-sick_(mp3CC.biz).mp3: 116444 хешей
** id=17 splin_-_plastmassovaja_zhizn_(zvukoff.ru).mp3: 115472 хешей
** id=4 m-not-sorry.mp3: 111944 хешей
** id=8 cold_hart_h_lil_peep_-_dying_(zf.fm).mp3: 108534 хешей
** id=9 scarlxrd_-_sx_sad_(zf.fm).mp3: 84803 хешей
** id=3 milky-chance-running_(mp3CC.biz).mp3: 80183 хешей
** id=13 sharlot_-_mechta_affinazh_cover_(zf.fm).mp3: 66707 хешей
** id=14 1540125020_1540124410_billie-eilish-when-the-party-s-over.mp3: 56316
хешей
** id=22 salem-king-night_(mp3CC.biz).mp3: 0 хешей

* Повторов: 0 песен

Готово
sqlite - connection has been closed
ilya@ilya-System:~/Desktop/findfprnt$

```

Рисунок 3.14 – Вікно консолі після введення команди get-database-stat.py

Також можна включити відображення спектрограми. Вона буде з'являтися після аналізу кожного каналу аудіо.

3.8 Вирівнювання відбитків

Тепер, коли аудіозапис був прослуханий, виконано ШПФ у вікнах, що перекриваються, по всій довжині пісні, видобули піки та створили відбитки.

Якщо припустити, що вже виконано це зняття відбитків на відомих доріжках, тобто відбитки додані до бази даних, позначені ідентифікаторами пісні, можна просто зіставити.

Код виглядає так:

```
channels = capture_audio()
fingerprints_matching = [ ]
for channel_samples in channels
    hashes = process_audio(channel_samples)
    fingerprints_matching += find_database_matches(hashes)

predicted_song = align_matches(fingerprints_matching)
```

Рисунок 3.15 – код вирівнювання відбитків

Якщо представити зразок, який прослуховується, як підрозділ оригінальної аудіодоріжки, то хеші, які витягуються з вибірки, матимуть зміщення, яке буде відносно початку вибірки.

Проблема, полягає в тому, що коли спочатку бралися відбитки, фіксувалось абсолютне зміщення хешу. Відносні хеші зразка та абсолютні хеші з бази даних ніколи не збігаються, якщо не почати записувати зразок саме з початку пісні, але це досить малоімовірно.

Хоча вони можуть бути не однаковими, дещо відомо про відповідність із реального сигналу, що стоїть за шумом. Відомо, що всі відносні зрушення будуть на однаковій відстані один від одного. Це вимагає припущення, що трек відтворюється та знімається з такою ж швидкістю, з якою він записувався та випускався у студії. Насправді, у випадку, якщо швидкість відтворення була іншою, не вдалося б розпізнати аудіо, яке вже міститься у БД, оскільки це вплине на частоту відтворення і, отже, на піки спектрограми. У будь-якому випадку припущення про швидкість відтворення є важливим.

Згідно з цим припущенням, для кожної відповідності обчислюємо різницю між відхиленнями:

$$\begin{aligned} \text{різниця} = & \text{зміщення БД від вихідної доріжки} - \\ & - \text{зсув вибірки від запису} \end{aligned} \quad (3.2)$$

Різниця завжди буде давати позитивний ціле число, тому що доріжка бази даних завжди буде мати довжину принаймні зразка. Всі справжні збіги мають таку саму різницю. Таким чином, збіги з бази даних змінені так: (Song_id, difference)

Тепер необхідно просто переглянути всі співпадіння та спрогнозувати ідентифікатор пісні, для якої виходить найбільша кількість різниць.

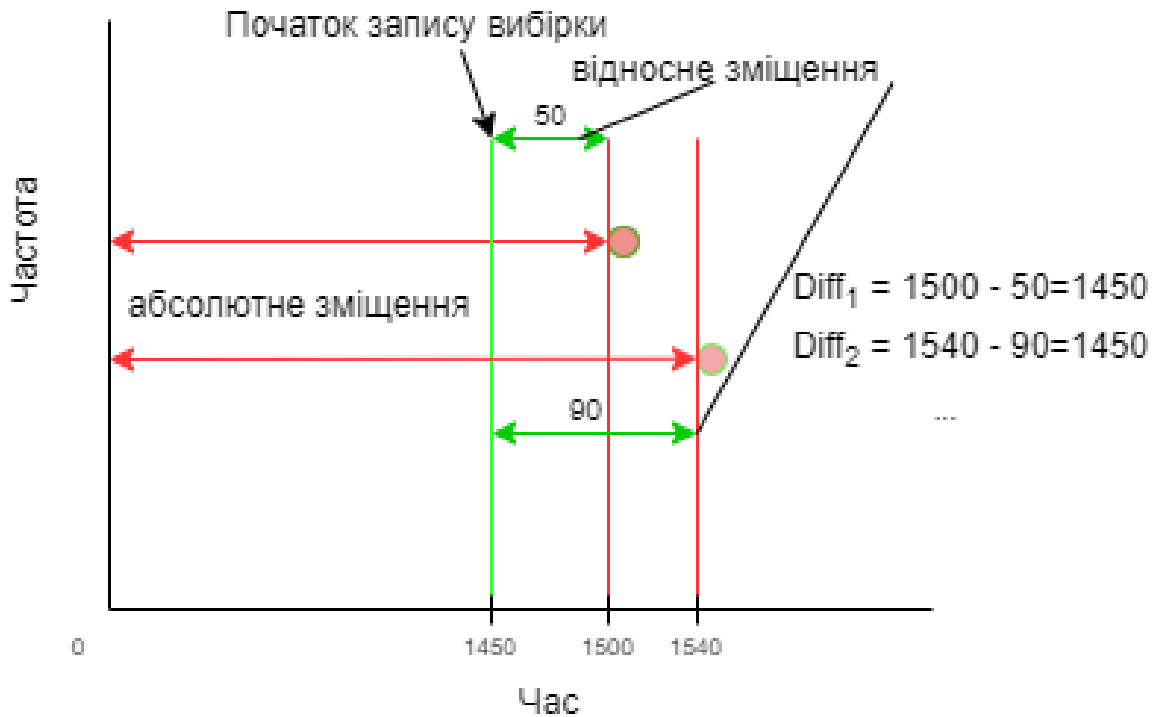


Рисунок 3.16 – Обчислення різниці між відхиленнями

3.9 Використання системи за призначенням

Після вищеописаних процедур можна тестувати систему. У БД було додано двадцять дев'ять пісень. Це зайняло приблизно сім хвилин. Після заповнення БД настав час тестувань. Аналіз фрагменту пісні запускається командою (рис. 3.17): «python recognize-from-microphone.py -s 5» (аналіз аудіо триватиме 5 секунд). Під час тестування одні й ті ж пісні розпізнавалися протягом 1,2,3,4 і 5 секунд. В результаті було обрано оптимальний час для розпізнавання аудіо. Ідеальним вибором виявилися записи, які аналізувалися протягом п'яти секунд.

```

ilya@ilya-System: ~/Desktop/findfpnt
File Edit View Search Terminal Help
00486 #
00452 #
00552 #
00358 #
00295
00793 ##
00728 ##
00558 #
00636 #
00942 ##
* запись остановлена
* записано 245760 виборок
канал 1/2
** найдено 26 хеш совпадений (шаг 426/426)
обработка канала завершена 1/2, получено 26 хешей
канал 2/2
** найдено 20 хеш совпадений (шаг 370/370)
обработка канала завершена 2/2, получено 46 хешей

** полностью найдено 46 хеш-совпадений
=> песня: noize_mc_-_vsyo_kak_u_ljudej_(zf.fm).mp3 (id=3)
смещение: 353 (16 сек)
confidence: 2
sqlite - connection has been closed

```

Рисунок. 3.17 - Результати роботи програми

Таблиця 3.1 - Результати роботи програми

Кількість секунд	Правильні визначення	Точність у відсотках
1	13/29	45%
2	18/29	62%
3	23/29	79%
4	26/29	89%
5	29/29	100%
6	29/29	100%

Навіть з однією секундою, обраною випадковим чином в будь-якому місці пісні, система розпізнає 45%. Від однієї додаткової секунди до 2 секунд розпізнається близько 62% пісень, в той час як на досконале розпізнавання необхідно мінімум 5 секунд.

3.10 Продуктивність зберігання даних

Для 29 пісень, для яких було я відібрав, база даних використовує 202 МБ місця на жорсткому диску для 4,8 мільйонів відбитків. Для порівняння, використання диска наведено нижче (таб. 3.2):

Таблиця 3.2 – Використання пам'яті на жорсткому диску.

Тип аудіоінформації	Розмір у МБ
MP3	322
Відбитки	202

Існує досить прямий компроміс між необхідним часом запису і обсягом необхідного сховища. Регулювання порога амплітуди для піків і значення вентилятора для зняття відбитків додасть більше відбитків і підвищить точність за рахунок більшого простору.

Відбитки займають надзвичайну кількість місця (трохи менше, ніж сирі файли MP3). Це викликає певну тривогу, поки не стає зрозуміло, що в кожній пісні десятки, а іноді і сотні тисяч хешів. Також включення зіставлення пісень впродовж п'яти секунд є дуже надійним, тому компроміс між простором і швидкістю, схоже, окупився.

3.11 Висновки до розділу 4

У даному розділі описано реалізацію підсистеми. Детально прокоментовано кожен пункт створення підсистеми. На наборі даних, який було обрано для тестування, підсистема продемонструвала стовідсоткове розпізнання невідомого файлу, при прослуховуванні запису хоча б 5 секунд.

Прослухавши аудіо один раз і занісши відбитки до БД підсистема запам'ятовує аудіо. При відтворюванні пісні і запису її через мікрофон, підсистема намагається порівняти звук з відбитками, які зберігаються у БД, повертаючи відтворювану пісню (звісно якщо така є у БД).

РОЗДІЛ 4

ОХОРОНА ПРАЦІ ТА БЕЗПЕКА В НАДЗВИЧАЙНИХ СИТУАЦІЯХ

В даному розділі проведено аналіз потенційних небезпечних та шкідливих виробничих факторів, причин пожеж. Розглянуті заходи, які дозволяють забезпечити гігієну праці і виробничу санітарію. На підставі аналізу розроблені заходи з техніки безпеки та рекомендації з пожежної профілактики.

Завданням даної роботи бакалавра було спроектувати систему управління продажами комп'ютерних компонент в мережі Інтернет, і як результат було створено повноцінний проект, котрий можна використовувати у розробці такої системи. За цим проектом в подальшому розроблятиметься реальна система, яка значно полегшить процес продажу комп'ютерних компонент у мережі Інтернет. Так як в процесі проектування використовувалося спеціалізоване приміщення, то аналіз потенційно небезпечних і шкідливих виробничих чинників виконується для персонального комп'ютера на якому буде розроблятися розроблена система.

4.1 Аналіз стану умов праці

Робочі місця офісних працівників, обладнані персональними комп'ютерами (далі – робочі місця), повинні відповідати вимогам «Правил охорони праці під час експлуатації електронно-обчислювальних машин», затверджених Наказом Державного комітету України з промислової безпеки, охорони праці та гірничого нагляду від 26.03.2010 року № 65 (Правила), та «Державних санітарних правил і норм роботи з візуальними дисплейними терміналами електронно-обчислювальних машин», затверджених постановою Головного державного санітарного лікаря України від 10.12.98 N 7 (ДСанПіН 3.3.2-007-98 [20]). Правила поширюються на всіх суб'єктів господарювання незалежно від форм власності, які у своїй діяльності здійснюють роботу, пов'язану з персональними комп'ютерами, у тому числі на тих, які мають робочі місця, обладнані персональними комп'ютерами і периферійними пристроями. Зазначені нормативно-правові акти встановлюють санітарно-гігієнічні вимоги до приміщення, в якому розташоване робоче місце, власне до робочого місця, освітлення, рівнів вібрації і шуму, мікроклімату в приміщенні тощо.

Приміщення

Будівлі та приміщення, де розміщені робочі місця, повинні відповідати вимогам нормативно-технічної та експлуатаційної документації виробника персональних комп'ютерів ДСанПіН 3.3.2-007-98 [20] та Правил. Будівлі та приміщення, де розміщені робочі місця

операторів, мають бути не нижче другого ступеня вогнестійкості. Для всіх будівель і приміщень, де знаходяться робочі місця, повинно бути визначено клас зони згідно з НПАОП 40.1-1.01-97 [21]. Відповідне позначення повинно бути нанесено на вхідних дверях кожного приміщення. Не дозволяється розташування приміщень з робочими місцями у підвалах і цокольних поверхах. Неприпустимим є розташування приміщень категорій А і Б, а також виробництв з мокрими технологічними процесами поряд з приміщеннями, де розташовуються робочі місця, а також над ними чи під ними. При цьому площа приміщення має бути не менше 6,0 кв. м. із розрахунку на одне робоче місце, а об'єм – не менше 20,0 куб. м.

Віконні прорізи приміщень для роботи з персональними комп'ютерами мають бути обладнані регульованими пристроями (жалюзі, завіски, зовнішні козирки. Для внутрішнього оздоблення приміщень з персональними комп'ютерами слід використовувати дифузно-відбивні матеріали з коефіцієнтами відбиття для стелі 0,7-0,8, для стін 0,5-0,6. Покриття підлоги повинне бути матовим з коефіцієнтом відбиття 0,3-0,5. Поверхня підлоги має бути рівною, неслизькою, з антистатичними властивостями. Забороняється для оздоблення інтер'єру приміщень з персональними комп'ютерами застосовувати полімерні матеріали (деревинно-стружкові плити, шпалери, що миються, рулонні синтетичні матеріали, шаруватий паперовий пластик тощо), що виділяють у повітря шкідливі хімічні речовини. Полімерні матеріали для внутрішнього оздоблення приміщень з персональними комп'ютерами можуть бути використані при наявності дозволу органів та установ державної санітарно-епідеміологічної служби. Приміщення можуть обладнуватись шафами для зберігання документів, магнітних дисків, полицями, стелажми, тумбами тощо з урахуванням вимог до площі приміщень.

У приміщеннях з джерелами шкідливих виробничих факторів робочі місця операторів мають розміщуватися в ізольованих кабінах, які обладнані повітрообміном. Заземлені конструкції, що знаходяться в приміщеннях, де розміщені робочі місця (батареї опалення, водопровідні труби, кабелі із заземленим відкритим екраном), мають бути надійно захищені діелектричними щитками або сітками з метою недопущення потрапляння працівника під напругу. Приміщення, де розміщені робочі місця, мають бути оснащені системою автоматичної пожежної сигналізації і вогнегасниками відповідно до вимог чинного законодавства України. Проходи до засобів пожежогасіння мають бути вільними. У приміщеннях, в яких розташовані робочі місця, слід щоденно робити вологе прибирання. Крім того, ці приміщення мають бути оснащені аптечками першої медичної допомоги, а при них мають бути обладнані побутові приміщення для відпочинку під час роботи, кімната психологічного розвантаження.

Організація та обладнання робочого місця

При розміщенні робочих столів з персональними комп'ютерами слід дотримувати:

- відстань між бічними поверхнями персональних комп'ютерів 1,2 м.;
- відстань від тильної поверхні одного персонального комп'ютера до екрана іншого – 2,5 м.

За потреби особливої концентрації уваги під час виконання робіт суміжні робочі місця операторів необхідно відділяти одне від одного перегородками висотою 1,5 – 2м.

Конструкція робочого місця користувача персонального комп'ютера має забезпечити підтримання оптимальної робочої пози офісного працівника. Конструкція робочого столу має відповідати сучасним вимогам ергономіки і забезпечувати оптимальне розміщення на робочій поверхні використовуваного обладнання (дисплея, клавіатури, принтера) і документів. Висота робочої поверхні робочого столу має регулюватися в межах 680-800 мм, а ширина і глибина – забезпечувати можливість виконання операцій у зоні досяжності моторного поля (рекомендовані розміри: 600-1400мм, глибина – 800-1000мм).

Робочий стіл повинен мати простір для ніг заввишки не менше ніж 600мм, завширшки не менше ніж 500мм, завглибшки (на рівні колін) не менше ніж 450мм, на рівні простягнутої ноги не менше ніж 650мм. Робочий стілець має бути підйомно-поворотним, регульованим за висотою, з кутом і нахилу сидіння та спинки і за відстанню від спинки до переднього краю сидіння поверхня сидіння має бути плоскою, передній край – заокругленим. Регулювання за кожним із параметрів має здійснюватися незалежно, легко і надійно фіксуватися. Шаг регулювання елементів стільця має становити: для лінійних розмірів – 15-20мм, для кутових – 2-5 градусів. Зусилля регулювання має не перевищувати 20Н. Висота поверхні сидіння має регулюватися в межах 400-500мм, а ширина і глибина становити не менше ніж 400мм. Кут нахилу сидіння – до 15 градусів вперед і до 5 градусів назад. Висота спинки стільця має становити (300+/-20) мм, ширина – не менше ніж 380 мм, радіус кривизни горизонтальної площини – 400мм. Кут нахилу спинки має регулюватися в межах 1-30 градусів від вертикального положення. Відстань від спинки до переднього краю сидіння має регулюватися в межах 260-400мм. Для зниження статичного напруження м'язів верхніх кінцівок слід використовувати стаціонарні або змінні підлокітники завдовжки не менше ніж 250мм, завширшки 50-70мм, що регулюються за висотою над сидінням у межах 230-260мм і відстанню між підлокітниками в межах 350-500мм. Поверхня сидіння і спинки стільця має бути напівм'якою з нековзним, повітронепроникним покриттям, що легко чиститься і не електризується. Робоче місце має бути обладнане підставкою для ніг завширшки не менше ніж 300мм, завглибшки не менше ніж 400мм, що регулюється за висотою в межах до 150мм і за кутом нахилу опорної поверхні підставки до 20 градусів. Підставка повинна мати рифлену поверхню і бортик по передньому краю заввишки 10мм.

Робочі місця слід розташовувати відносно світових прорізів так, щоб природне світло падало переважно з лівого боку. Монітор має розташовуватися на оптимальній відстані від очей користувача, що становить 600-700мм, але не ближче ніж за 600мм з урахуванням розміру літерно-цифрових знаків і символів. Розташування екрана монітору має забезпечувати зручність зорового спостереження у вертикальній площині під кутом +30 градусів до нормальної лінії погляду працівника. Клавіатуру слід розташовувати на поверхні столу на відстані 100-300 мм від краю, звернутого до працюючого. У конструкції клавіатури має передбачатися опорний пристрій (виготовлений із матеріалу з високим коефіцієнтом тертя, що перешкоджає мимовільному її зсуву), який дає змогу змінювати кут нахилу поверхні клавіатури у межах 5-15 градусів. Висота середнього рядка клавіш має не перевищувати 30мм. Поверхня клавіатури має бути матовою з коефіцієнтом відбиття 0,4. Розташування пристрою введення – виведення інформації має забезпечувати добру видимість монітору, зручність ручного керування в зоні досяжності моторного поля і за висотою – 900-1300мм, за шириною 400-500мм. Під матричні принтери потрібно підкладати вібраційні килимки для гасіння вібрації та шуму.

Робоче місце з персональним комп'ютером слід обладнати пюпітром для документів, що легко переміщуються.

Для забезпечення захисту і досягнення нормованих рівнів комп'ютерних випромінювань необхідно застосування приєкраних фільтрів, локальних світлофільтрів (засобів індивідуального захисту очей) та інших засобів захисту, що пройшли випробування в акредитованих лабораторіях і мають щорічний гігієнічний сертифікат.

Безпека під час роботи з персональним комп'ютером

Щодня перед початком роботи необхідно очищати монітор від пилу та інших забруднень. Після закінчення роботи персональний комп'ютер і периферійні пристрої повинні бути відключені від електричної мережі. У разі виникнення аварійної ситуації необхідно негайно відключити персональний комп'ютер і периферійні пристрої від електричної мережі.

Не допускається:

- виконувати обслуговування, ремонт та налагодження персонального комп'ютеру та периферійних пристроїв безпосередньо на робочому місці оператора;
- зберігати біля персонального комп'ютеру та периферійних пристроїв папір, будь-які носії інформації (диски, флешки тощо), запасні блоки, деталі тощо, якщо вони не використовуються для поточної роботи;
- відключати захисні пристрої, самочинно проводити зміни у конструкції та складі персонального комп'ютеру та периферійних пристроїв або їх технічне налагодження;

- працювати з персональним комп'ютером, у яких під час роботи з'являються нехарактерні сигнали, нестабільне зображення на моніторі тощо;
- працювати з матричним принтером за відсутності вібраційного килимка та зі знятою (піднятою) верхньою кришкою.

4.2 Виробнича санітарія

На підставі аналізу небезпечних та шкідливих факторів при виробництві (експлуатації), пожежної безпеки можуть бути надалі вирішені питання необхідності забезпечення працюючих достатньою кількістю освітлення, вентиляції повітря, організації заземлення, тощо.

4.2.1 Аналіз небезпечних та шкідливих факторів при виробництві (експлуатації) виробу

Роботу, пов'язану з ЕОП з ВДТ, у тому числі на тих, які мають робочі місця, обладнані ЕОМ з ВДТ і ПП, виконують із забезпеченням виконання НПАОП 0.00-7.15-18 [22] «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями», які встановлюють вимоги безпеки до обладнання робочих місць, до роботи із застосуванням ЕОМ з ВДТ і ПП. Переважно роботи за проектами виконують у кабінетах чи інших приміщеннях, де використовують різноманітне електрообладнання, зокрема персональні комп'ютери (ПК) та периферійні пристрої.

- робоча напруга $U=+220\text{В} \pm 5\%$;
- робочий струм $I=2\text{А}$;
- споживана потужність $P=350\text{ Вт}$.

Робочі місця мають відповідати вимогам Державних санітарних правил і норм роботи з візуальними дисплейними терміналами електронно-обчислювальних машин, затверджених постановою Головного державного санітарного лікаря України від 10.12.98 N 7 [20]. За умов роботи з ПК виникають наступні небезпечні та шкідливі чинники: несприятливі мікрокліматичні умови, освітлення, електромагнітні випромінювання, забруднення повітря шкідливими речовинами (джерелом, яких можуть бути: принтер, сканер та інші джерела виділення багатьох хімічних речовин - напр., озону, оксидів азоту та аерозолів високодисперсних частинок тонера), шум, вібрація, електричний струм, електростатичне поле, напруженість трудового процесу та інше.

4.2.2 Пожежна безпека

Небезпека розвитку пожежі на обчислювальному центрі обумовлюється застосуванням розгалужених систем електроживлення ЕОМ, вентиляції і кондиціонування. Небезпека загоряння пов'язана з особливістю комп'ютерів - із значною кількістю щільно розташованих на монтажній платі і блоках електронних вузлів і схем, електричних і комутаційних кабелів, резисторів, конденсаторів, напівпровідникових діодів і транзисторів. Надійна робота окремих елементів і мікросхем в цілому забезпечується тільки в певних інтервалах температури, вологості і при заданих електричних параметрах. При відхиленні реальних умов експлуатації від розрахункових можуть виникнути пожежонебезпечні ситуації.

Запобігти утворенню горючого середовища (замінити горючі речовини і матеріали на негорючі і важко горючі) не надається технічно можливим. Тому проектом передбачаються способи і засоби запобігання утворення (або внесення) в горюче середовище джерел запалювання, таких як: застосування електроустаткування, відповідної пожежонебезпечної і вибухонебезпечної зонам відповідно до ПУЕ;

Згідно ГОСТ 12.1.004-91 ССБТ [23] таке приміщення, площею 25 м², відноситься до категорії "В" (пожежонебезпечної) та для протипожежного захисту в ньому проектом передбачено устаткування автоматичною пожежною сигналізацією із застосуванням датчиків-сповіщувачів РІД-1 (сповіщувач димовий ізоляційний) в кількості 1 шт., і застосуванням первинних засобів пожежогасіння.

4.2.3 Електробезпека

На робочому місці виконуються наступні вимоги електробезпеки: ПК, периферійні пристрої та устаткування для обслуговування, електропроводи і кабелі за виконанням та ступенем захисту відповідають класу зони за ПУЕ (правила улаштування електроустановок), мають апаратуру захисту від струму короткого замикання та інших аварійних режимів. Лінія електромережі для живлення ПК, периферійних пристроїв і устаткування для обслуговування, виконана як окрема групова три-провідна мережа, шляхом прокладання фазового, нульового робочого та нульового захисного провідників. Нульовий захисний провідник використовується для заземлення (занулення) електроприймачів. Штепсельні з'єднання та електророзетки крім контактів фазового та нульового робочого провідників мають спеціальні контакти для підключення нульового захисного провідника. Електромережа штепсельних розеток для живлення персональних ПК, укладено по підлозі поруч зі стінами відповідно до затвердженого плану розміщення обладнання та технічних

характеристик обладнання. Металеві труби та гнучкі металеві рукави заземлені. Захисне заземлення включає в себе заземлюючих пристроїв і провідник, який з'єднує заземлюючий пристрій з обладнанням, яке заземлюється - заземлюючий провідник.

4.3 Гігієнічні вимоги до параметрів виробничого середовища

– площа приміщення повинна бути не менше 6,0 м² на 1 робоче місце; робочі місця повинні бути розташовані на відстані не менше ніж 1 м від стіни з вікном, і 1,4 м від звичайної стіни; відстань між бічними поверхнями комп'ютерів має бути не меншою за 1,2 м; відстань між тильною поверхнею одного комп'ютера та екраном іншого не повинна бути меншою 2,5м.

– відповідні робочі місця заборонено облаштовувати у підвальних або цокольних приміщеннях будинків. В обладнанні приміщень забороняється використання полімерних матеріалів (деревинно-стружкові плити, шпалери, що миються, рулонні синтетичні матеріали, шаруватий паперовий пластик тощо), що виділяють у повітря шкідливі хімічні речовини. Покриття підлоги повинно бути матовим, а поверхня – рівною, неслизькою, з антистатичними властивостями.

– особливу увагу необхідно приділити колірній гармонії офісних приміщень. Колір є засобом створення психологічного комфорту та підвищення продуктивності праці. Найбільш сприятливі для нервової системи світлі, пастельні тони – зеленувато-блакитний, ясно-сірий, золотавий. Яскраві, контрастні поєднання (синій і жовтогарячий, червоний і фіолетовий) викликають втому, роздратування.

-у приміщеннях, де здійснюється робота з комп'ютерами, щодня має проводитися вологе прибирання з метою недопущення запиленості підлоги та меблів. Крім того, має бути обладнана кімната психологічного розвантаження.

-конструкція робочого столу та крісла користувача персонального комп'ютера має забезпечити підтримання оптимальної робочої пози та забезпечувати оптимальне розміщення на робочій поверхні використовуваного обладнання (дисплея, клавіатури, принтера) і документів.

-приміщення для роботи з персональними комп'ютерами мають бути обладнані системами опалення, кондиціонування повітря, або припливно-витяжною вентиляцією. У приміщеннях на робочих місцях мають забезпечуватись оптимальні значення параметрів мікроклімату: температура повітря повинна становити 22–25°C, відносна вологість повітря — 40–60%, швидкість руху повітря — не більше 0,1 м/с. При недотриманні вказаних показників мікроклімату в офісних приміщеннях робочий день для робітників повинен бути скорочений мінімум на 10%.

-досить важливим є вимоги до освітлення приміщень, оскільки відомо, що тривала робота за комп'ютером та з документами при недостатньому рівні освітленості може призвести до значного перенапруження зору. Природне освітлення має забезпечувати коефіцієнт природної освітленості (КПО) не нижче ніж 1,5%. Для регулювання рівня освітлення природним світлом бажано застосовувати жалюзі. Робоче місце, обладнане ПК повинно бути розташоване так, щоб уникнути попадання в очі прямого сонячного світла. Штучне освітлення приміщення має бути обладнане системою загального рівномірного освітлення. Застосування світильників без розсіювачів та екрануючих сіток забороняється. Рівень освітленості на робочому столі в зоні розташування документів має бути в межах 300–500 лк.

-в офісних приміщеннях нормуються також еквівалентні рівні звуку (для програмістів – 50 дБА, а для операторів в залах обробки інформації на ПК та операторів комп'ютерного набору – 65 дБА).

-вимоги щодо рівня неіонізуючих електромагнітних випромінювань, електростатичних і магнітних полів, а також інтенсивність потоків інфрачервоного та ультрафіолетового випромінювань встановлюються відповідно до ДСанПіН 3.3.2-007-98 [20] і ДСанПіН 3.3.6.096-2002 [24].

4.4 Рекомендації із пожежної профілактики

У виробничому приміщенні на організм людини та його працездатність впливають мікрокліматичні фактори. Мікроклімат виробничих приміщень визначається поєднанням температури, вологості й швидкості руху повітря, а також температури навколишніх поверхонь.

Для робіт категорії 1а для робочої зони виробничих приміщень забезпечуються наступні метеорологічні умови:

- в холодний і перехідний період року температура повітря $22 \div 24^{\circ}\text{C}$;
- відносна вологість повітря $40 \div 60\%$, швидкість руху повітря не більше 0.1 м/с;
- у теплий період року температура повітря $23 \div 25^{\circ}\text{C}$, відносна вологість повітря 40-60%, швидкість руху повітря не більше 0.1 м/с.

Рівень шуму не перевищує санітарних норм. Тому застосування захисту від шуму в роботі не передбачається.

Ще одна проблема полягає в тому, що спектр випромінювання комп'ютерного монітора включає в себе рентгенівську, ультрафіолетову та інфрачервону області, а також широкий діапазон хвиль інших частот. Небезпека рентгенівських променів мізерно мала, оскільки цей вид випромінювання поглинається речовиною екрану. Однак велику увагу слід приділяти біологічним ефектам низькочастотних електромагнітних полів.

Для зниження стомлюваності персоналу в приміщеннях, де розташовані обчислювальні засоби, передбачається використовувати кольорові поєднання й покриття, що не дають відблисків.

У проекті, що розробляється передбачається використовувати сполучне освітлення. У світлий час доби приміщення буде висвітлюватися через віконні прорізи, в решту часу буде використовуватися штучне освітлення. Штучне освітлення створюється лампами розжарювання або газорозрядними лампами. Штучне освітлення в робочому приміщенні передбачається здійснювати з використанням люмінесцентних джерел світла в світильниках загального освітлення, оскільки люмінесцентні лампи мають високу світловою віддачу (до 75 лм/Вт і більше), тривалим терміном служби (до 10000 годин), спектральним складом випромінюваного світла, близьким до сонячного. При експлуатації комп'ютерів виробляється зорова робота IV в розряді точності (середня точність). При цьому нормована освітленість на робочому місці дорівнює 200 лк. Джерелом природного світла (освітлення) є сонячне світло. У приміщенні, де розташовані комп'ютери, передбачається природне бічне освітлення, рівень якого відповідає ДБН В.2.5-28-2015 [25] "Будівельні норми і правила".

Регулярно проводиться контроль освітленості, який підтверджує, що рівень освітленості задовольняє ДБН і для данного приміщення в світлий час доби достатньо природного освітлення. Для організації освітлення в темний час доби передбачається обладнати приміщення, довжина якого становить 5 м, ширина 5 м, висота 4 м, світильниками ЛПО2П, облаштованими лампами типу ЛБ (дві по 80 Вт) зі світловим потоком 5400 лм кожна.

Розрахунок штучного освітлення проводиться за коефіцієнтами використання світлового потоку, яким визначається потік, необхідний для створення заданої освітленості при загальному рівномірному освітленні.

Розрахунок кількості світильників здійснюється за формулою:

$$N = (E \times S \times Z \times K) / (F \times U \times M) \quad (4.1)$$

де N - число світильників;

E - нормоване освітлення;

S - площа підлоги, м²;

Z - поправний коефіцієнт світильника (для стандартних світильників Z=1.1÷1.3);

K - коефіцієнт запасу, що враховує зниження освітленості в процесі експлуатації;

U - коефіцієнт використання, що залежить від типу світильника, показника індексу приміщення і т. п. (U= 0.55 ÷ 0.6);

M - число люмінесцентних ламп у світильнику;

F - світловий потік.

Згідно вимог ДБН В.2.5-28-2015 [25], освітлення робочого місця оператора обчислювальної техніки повинно бути не менше 200 лк.

$$N = (200 \times 25 \times 1,1 \times 1,5) / (5500 \times 0,6 \times 2) = 1,27 \approx 1$$

Обираємо кількість світильників, що дорівнює 1.

4.5 Охорона навколишнього природного середовища

Діяльність за темою магістерської роботи, а саме робота за комп'ютером в процесі її виконання впливає на навколишнє природне середовище і регламентується нормами діючого законодавства.

Основним екологічним аспектом в процесі діяльності за даними спеціальностями є процеси впливу на атмосферне повітря та процеси поводження з відходами, які утворюються, збираються, розміщуються, передаються на видалення (знешкодження), утилізацію, тощо в ІТ галузі.

Вплив на атмосферне повітря при нормальних умовах праці не оказує, бо не має в приміщенні сканерів, принтерів та інших джерел викиду забруднюючих речовин в повітря робочої зони.

В процесі діяльності комп'ютера виникають процеси поводження з відходами ІТ галузі. Нижче надано перелік відходів, що утворюються в процесі роботи:

Відпрацьовані люмінесцентні лампи - I клас небезпеки

Батарейки та акумулятори (малі) -III клас небезпеки

Змінні носії інформації - IV клас небезпеки

Відходи друкуючих пристроїв - IV клас небезпеки

Відпрацьований ізолюючий матеріал, дроти та кабелі - IV клас небезпеки

Макулатура - IV клас небезпеки

Побутові відходи - IV клас небезпеки

Відходи в міру їх накопичення збирають у тару, відповідну класу небезпеки, з дотриманням правил безпеки, після чого доставляють до місця тимчасового зберігання відходів відповідно до затвердженої схеми їх розміщення. Зазначені для зберігання відходів місця чи об'єкти повинні використовуватися лише для заявлених відходів.

Не допускається зберігання відходів у невстановлених схемою місцях, а також перевищення норм тимчасового зберігання відходів.

Способи тимчасового зберігання відходів визначаються видом, агрегатним станом і класом небезпеки відходів:

- Відходи I класу небезпеки зберігаються в герметичній тарі (сталеві бочки, контейнери). У міру наповнення тару з відходами закривають герметично сталевий кришкою;

- Відходи III класу небезпеки зберігаються в тарі, яка забезпечує локалізацію зберігання, дозволяє виконувати вантажно-розвантажувальні і транспортні роботи і виключає поширення в ОС шкідливих речовин;

- Відходи IV класу небезпеки можуть зберігатися відкрито на промисловому майданчику у вигляді конусоподібної купи, звідки їх автонавантажувачем перевантажують у самоскид і доставляють на місце утилізації або захоронення.

З метою визначення та прогнозування впливу відходів на навколишнє середовище, своєчасного виявлення негативних наслідків, їх запобігання відповідно до Закону України «Про відходи» [26] повинен здійснюватися моніторинг місць утворення, зберігання, і видалення відходів.

4.6 Мікроклімат

Суттєвий вплив на стан організму працівника, його працездатність здійснює мікроклімат (метеорологічні умови) у виробничих приміщеннях, під яким розуміють умови внутрішнього середовища цих приміщень, що впливають на тепловий обмін працюючих з оточенням. Мікроклімат визначається сукупністю фізичних параметрів повітряного середовища, таких як температура, швидкість руху, вологість і барометричний тиск повітря, температура поверхонь, що оточують людину, та інтенсивність інфрачервоного випромінювання.

Властивість організму людини підтримувати тепловий баланс із навколишнім середовищем називаються терморегуляцією.

Нормальне протікання фізіологічних процесів, а отже, і хороше самопочуття можливе лише тоді, коли тепло, що виділяється організмом людини, постійно відводиться в навколишнє середовище. Кількість тепла, що утворюється в організмі людини, залежить від фізичних навантажень, а рівень тепловіддачі — від мікрокліматичних умов, переважно температури повітря.

Теплообмін організму людини з навколишнім середовищем здійснюється такими способами (шляхами): конвекція, кондукція (теплопровідність), випромінювання та випаровування вологи з поверхні шкіри.

Основним нормативним документом, що регламентує параметри мікроклімату виробничих приміщень, є ДСН 3.3.6.042-99 [27]. Цей документ встановлює оптимальні і допустимі значення температури, відносної вологості та швидкості руху повітря, допустиму температуру внутрішніх поверхонь приміщення (стіни, стеля, підлога) і зовнішніх поверхонь технологічного обладнання, а також допустиму інтенсивність теплового випромінювання нагрітих поверхонь у приміщенні та відкритих джерел тепла (нагрітий метал, скло, відкритий вогонь тощо) для робочої зони — визначеного простору, в якому знаходяться робочі місця постійного або непостійного (тимчасового) перебування працівників.

Табл. 4.1 Оптимальні величини температури, відносної вологості та швидкості руху повітря в робочій зоні.

Період року	Категорія робіт	Температура повітря, град. С	Відносна вологість повітря, %	Швидкість руху повітря, м/с
		оптимальна	оптимальна	оптимальна
Холодний період року	легка 1 а	22 – 24	40 – 60	0.1
	легка-1 б	21 – 23	40 – 60	0.1
Теплий період року	легка-1 а	23 - 25	40 – 60	0.1
	легка-1 б	22 – 24	40 – 60	0.2

4.6 Висновки до розділу 4

У розділі «Охорона праці» проаналізовано потенційні небезпеки при роботі з засобами обчислювальної техніки, на підставі якого розроблено заходи з техніки безпеки, заходи, що забезпечують виробничу санітарію та гігієну праці, рекомендації з пожежної профілактики, які підтверджені відповідними розрахунками.

Була наведена схема, розміри приміщення та наведено значення температури, вологості й рухливості повітря, необхідна кількість і потужність ламп та інші параметри, значення яких впливає на умови праці робітника, а також – наведені інструкції з охорони праці, техніки безпеки при роботі на комп'ютері.

А також визначені основні екологічні аспекти впливу на навколишнє природне середовище та зазначені заходи щодо поводження з ними.

ВИСНОВКИ

У магістерській роботі було розроблено підсистему розпізнавання аудіозаписів. Створена підсистема задовольняє всім вимогам поставленого завдання.

Детально прокоментовано кожен пункт створення підсистеми. На наборі даних, який було обрано для тестування, підсистема продемонструвала стовідсоткове розпізнання невідомого файлу, при прослуховуванні запису хоча б 5 секунд.

Прослухавши аудіо один раз і занісши відбитки до БД підсистема запам'ятовує аудіо. При відтворюванні пісні і запису її через мікрофон, підсистема намагається порівняти звук з відбитками, які зберігаються у БД, повертаючи відтворювану пісню (звісно, якщо така є у БД).

Щодо перспектив розвитку, у майбутньому планується зробити «оболонку» для підсистеми – сайт чи прикладну програму для смартфонів, щоб у користувачів була можливість використовувати підсистему де б їм було зручно. Заплановано значно розширити базу даних відбитків, проаналізувавши велику кількість аудіовідбитків. Надати перевіреним користувачам можливість поповнювати БД новими творами.

ПЕРЕЛІК ДЖЕРЕЛ І ПОСИЛАНЬ

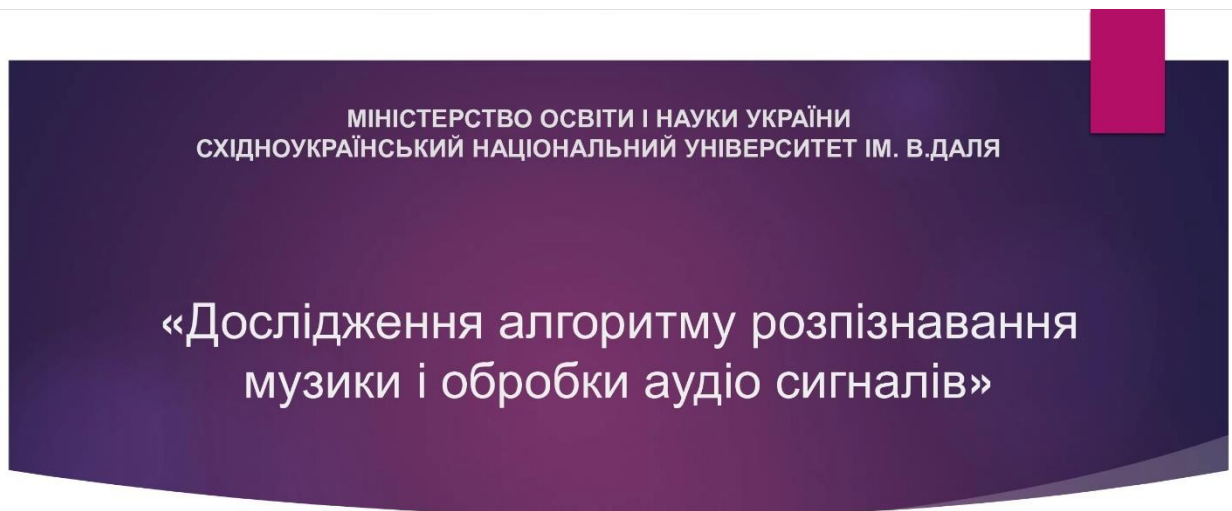
1. Панфілов И.П., Дирда В.Е. Теорія електричного зв'язку [Текст]. — М.: Радіо і зв'язок, 1991. — 344 с.
2. Обчислювальна техніка. Термінологія: Довідковий посібник. Випуск 1 [Текст] / Рецензент канд. техн. наук Ю. П. Селіванов. — М.: Видавництво стандартів, 1989. — 168 с.
3. Хоровиц П., Хілл У. Мистецтво схемотехніки: В 3-х томах: Т. 2. Пер. с англ.— 4-е изд., переп. и доп.— М.: Мир, 1993.—371 с
4. Л. Рабінер, Б. Гоулд Теория и применение цифровой обработки сигналов — М.: Мир, 1978, с. 327.
5. Жан М. Рабаї, Ананта Чандракасан, Боривож Ніколич. Цифрові інтегральні схеми. Методологія проектування = Digital Integrated Circuits. — 2-е изд. — М.: Уільямс, 2007. — 912 с.
6. Смірнов С. В. Засоби і системи технічного забезпечення обробки, зберігання та передачі інформації. — МГІУ, 2011. — С. 260
7. Генадієва Е. Г., Дождіков В. Г., Кульба А. В. и др. Короткий енциклопедичний словник з радіоелектроніки і радіопромисловості / Под ред. В. Н. Сабліна. М.: Диво, 2006. С. 276.
8. M. Frigo, S.G. Johnson, FFTW: an adaptive software architecture for the FFT, in Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, 1998, vol. 3 (IEEE, 1998), pp. 1381–1384
9. C. Runge, Zeit. f. Math. u. Phys 48, 443–456 (1903)
10. J. Cooley, J. Tukey, An algorithm for the machine calculation of complex Fourier series. Math. Comput. 19(90), 297–301 (1965)
11. O. Jung-Yeol, L. Myoung-Seob, New radix-2 to the 4th power pipeline FFT processor. IEICE Trans. Electron. 88(8), 1740–1746 (2005)
12. S.-H. Hsieh, C.-S. Lu, S.-C. Pei, Sparse fast Fourier transform by downsampling, in 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (IEEE, 2013), pp. 5637–5641
13. C.-P. Fan, M.-S. Lee, G.-A. Su, A low multiplier and multiplication costs 256-point FFT implementation with simplified radix-24 SDF architecture, in IEEE Asia Pacific Conference on Circuits and Systems, 2006. APCCAS 2006 (2006), pp. 1935–1938. — Режим доступу: <https://doi.org/10.1109/APCCAS.2006.342239>
14. M. Narwaria, W. Lin, I.V. McLoughlin, S. Emmanuel, L.-T. Chia, Fourier transform-based scalable image quality measure. IEEE Trans. Image Process. 21(8), 3364–3377 (2012)

15. B. Ghazi, H. Hassanieh, P. Indyk, D. Katabi, E. Price, L. Shi, Sample-optimal average-case sparse Fourier transform in two dimensions. arXiv preprint arXiv:1303.1209 (2013)
16. L. Shi et al., Imaging Applications of the Sparse FFT. Ph.D. thesis, Massachusetts Institute of Technology (2013)
17. J.-R. Choi, S.-B. Park, D.-S. Han, S.-H. Park, A 2048 complex point FFT architecture for digital audio broadcasting system, in The 2000 IEEE International Symposium on Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva, vol. 5 (IEEE, 2000), pp. 693–696
18. R.L. Rivest, A. Shamir, L. Adleman, A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM 21(2), 120–126 (1978)
19. A. Schönhage, V. Strassen, Schnelle multiplikation grosser zahlen. Computing 7(3–4), 281–292 (1971)
20. ДСанПіН 3.3.2-007-98 «Державні санітарні правила і норми. Гігієнічні вимоги до організації роботи з візуальними дисплейними терміналами електронно-обчислювальних машин». Затверджено постановою Головного державного санітарного лікаря України 10 грудня 1998 р. № 7. Режим доступу: [www. URL: https://zakon.rada.gov.ua/rada/show/v0007282-98](http://www.zakon.rada.gov.ua/rada/show/v0007282-98)
21. НПАОП 40.1-1.01-97 «Правила безпечної експлуатації електроустановок». Наказ Державного комітету України по нагляду за охороною праці від 6 жовтня 1997 р. № 257 Зареєстровано в Міністерстві юстиції України 13 січня 1998 р. за № 11/2451. Режим доступу: [www. URL: https://zakon.rada.gov.ua/laws/show/z0011-98](http://www.zakon.rada.gov.ua/laws/show/z0011-98)
22. НПАОП 0.00-7.15-18 «Вимоги щодо безпеки та захисту здоров'я працівників під час роботи з екранними пристроями». Зареєстровано в Міністерстві юстиції України 25 квітня 2018 р. За № 508/31960. Режим доступу: [www. URL: https://zakon.rada.gov.ua/laws/show/z0508-18](http://www.zakon.rada.gov.ua/laws/show/z0508-18)
- 23.ГОСТ 12.1.004-91 ССБТ «Пожежна безпека. Загальні вимоги». Затверджено і введено в дію Ухвалою Державного комітету СРСР з управління якістю продукції і та стандартів від 14.06.91 №875. Режим доступу: [www.URL: http://docs.cntd.ru/document/9051953](http://docs.cntd.ru/document/9051953)
24. ДСанПіН 3.3.6.096-2002 «Державні санітарні норми і правила при роботі з джерелами електромагнітних полів». Зареєстровано в Міністерстві юстиції України 13 березня 2003 р. за № 203/7524. Режим доступу: [www. URL: https://zakon.rada.gov.ua/laws/show/z0203-03](http://www.zakon.rada.gov.ua/laws/show/z0203-03)
25. ДБН В.2.5-28-2015 «Природне і штучне освітлення». Режим доступу: [www. URL: http://www.minregion.gov.ua/wp-content/uploads/2018/12/V2528-1.pdf](http://www.minregion.gov.ua/wp-content/uploads/2018/12/V2528-1.pdf)
- 26.Закон України «Про відходи». Відомості Верховної Ради України (ВВР), 1998, № 36-37, ст.242. Режим доступу: [www. URL: https://zakon.rada.gov.ua/laws/show/187/98-вр](http://www.zakon.rada.gov.ua/laws/show/187/98-вр)

27. ДСН 3.3.6.042-99 «Державні санітарні норми мікроклімату виробничих приміщень». Вводиться в дію Постановою Головного Державного санітарного лікаря України від 01.12.1999 № 42. Режим доступу: [www. URL:](http://www.url)
<https://zakon.rada.gov.ua/rada/show/va042282-99>

Додаток А

Електронні плакати



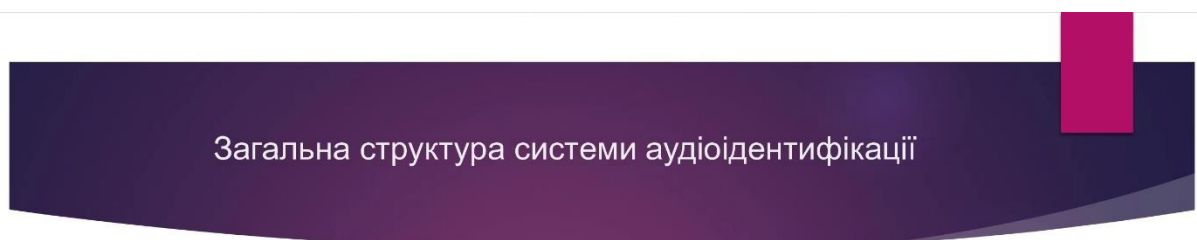
Студент гр. КН-18дМ

Хоткін І.О.

Науковий керівник

доц. Щербакова М.Є

Рисунок А.1 - Титульний лист



В контексті даної магістерської роботи було виділено наступні цілі:

1. Ознайомлення з інструментами та ресурсами веб розробника.
2. Розробка системи для розкладення аудіофайлів на хеші.
3. Розробка системи для автоматичного розпізнавання аудіо, яке вже є у БД.

Підсистема повинна функціонувати на операційній системі Windows (або Linux) і надавати можливість користувачу розпізнавати невідомі для нього пісні. Загальна структура роботи системи зображена на рисунку.

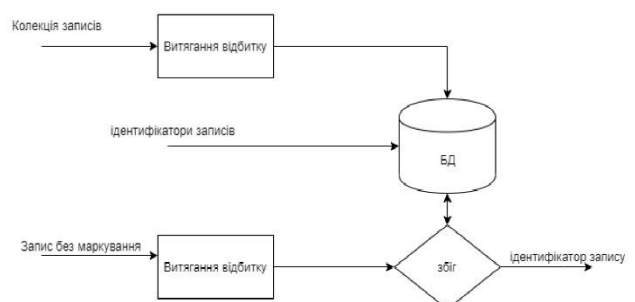


Рисунок А.2 – Загальна структура системи аудіоідентифікації

Швидке перетворення Фур'є

Швидке перетворення Фур'є (ШПФ) - важливий метод вимірювання в науці про вимірювання аудіо та акустики. Воно перетворює сигнал в окремі спектральні компоненти і тим самим забезпечує інформацію про частоту про сигнал. ШПФ - це оптимізований алгоритм реалізації дискретного перетворення Фур'є (ДПФ). Сигнал відбирається за певний проміжок часу і ділиться на його частотні компоненти. Ці компоненти являють собою поодинокі синусоїдальні коливання з різними частотами, кожна зі своєю амплітудою та фазою. Ця трансформація проілюстрована на малюнку праворуч. За вимірюваний період часу сигнал містить 3 чітко домінуючих частоти.

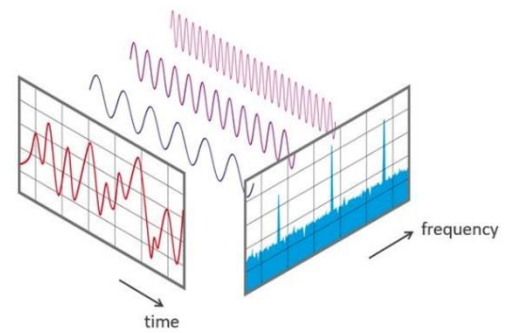


Рисунок А.3 – Швидке перетворення Фур'є

Спектрограма аудіо

Спектрограма це просто двовимірний масив з амплітудою як функції часу і частоти. ШПФ показує нам амплітуду сигналу саме на цій частоті, даючи нам колонку. Якщо зробити це досить разів за допомогою розсунутого вікна ШПФ, можна отримати 2D спектрограму масиву.

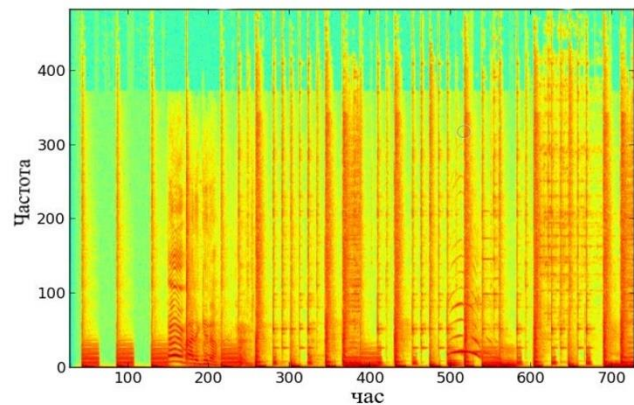


Рисунок А.4 – Спектрограма аудіо

Видалення шуму

- Фільтр замінює кожне значення пікселя цифрового зображення на максимальне значення (тобто значення найяскравішого пікселя) вікна сусіднього пікселя.
- Застосування фільтра видаляє негативний зовнішній шум, присутній у цифровому зображенні.
- Якщо застосовано максимальний фільтр, темніші предмети, присутні на зображенні, стираються.

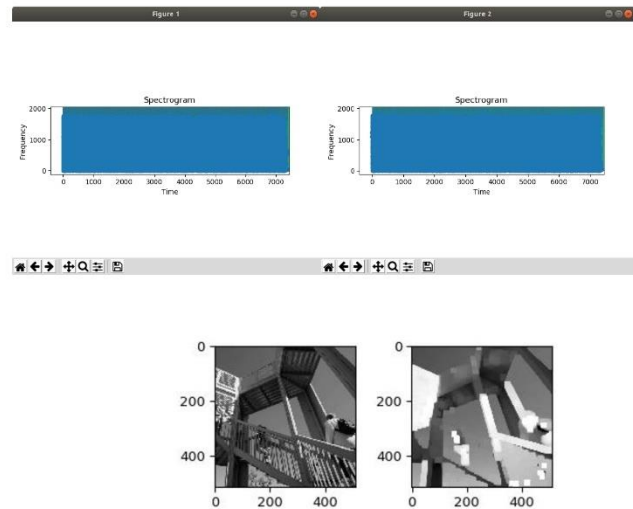


Рисунок А.5 – Видалення шуму

Пошук піків спектрограми

Пік визначається як пара (час, частота), відповідаючи значенню амплітуди, що є найвищою в локальній «околиці» навколо нього. Інші пари (час, частота) навколо мають меншу амплітуду і, слідуєчи, менше схильні до впливу шуму.

Пошук піків - це головна проблема. У підсумку було оброблено спектрограму як зображення і використовував інструментарій обробки зображень і методи ссіру щоб знайти піки. Комбінація фільтра верхніх частот і структури ссіру локальних максимумів зробили свою справу.

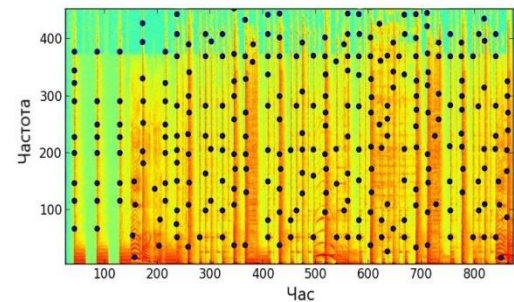


Рисунок А.6 – Пошук піків спектрограми

Аудіо відбиток

Аудіо-відбиток - відомий завдяки його здатності пов'язувати неназначений звук до відповідних метаданих (наприклад, ім'я виконавця та назва пісні), незалежно від формату звуку. Аудіо-відбитки витягають перцептивний довідник аудіовмісту, тобто відбитка і зберігають його в базі даних. Коли він подається з маркованим аудіо, його відбиток обчислюється і порівнюється з тими, що зберігаються в базі даних. Використовуючи відбитки та алгоритми відповідності, спотворені версії запису можуть бути ідентифіковані як однаковий звуковий вміст.

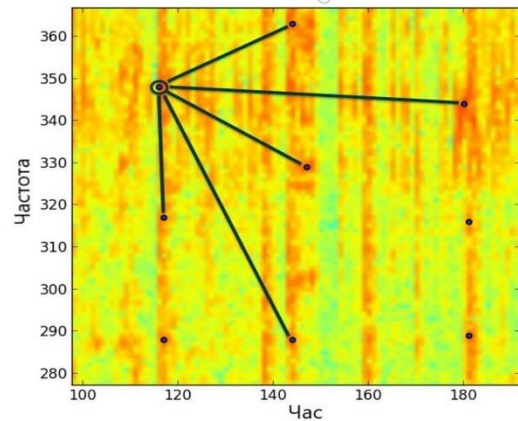
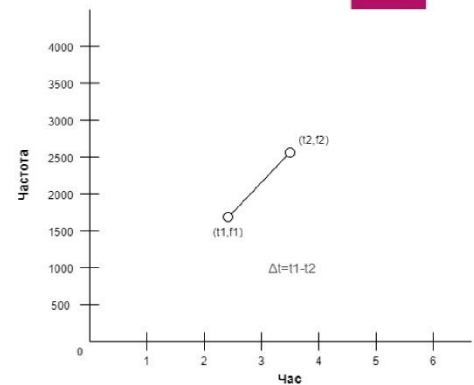


Рисунок А.7 – Аудіо відбиток

Хешування відбитків

Дивлячись на піки спектрограми і комбінуючи пікові частоти і різницю в часі між ними, можна створити хеш, який представляє унікальний відбиток для цієї пісні. Загальний зміст хешу представлено нижче.

Завдяки такій кількості відбитків, потрібно скоротити дисковий простір для зберігання значення. Для хешу відбитків було використано хеш SHA-1, а потім його необхідно скоротити до половини його розміру (лише перші 20 символів).



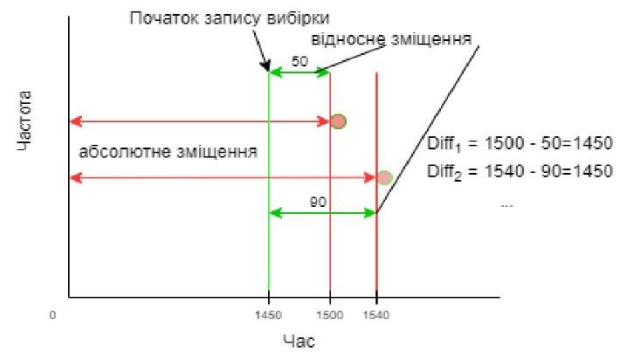
$\text{sha1}(\text{частота першого піку} \mid \text{частота другого піку} \mid \text{різниця у часі між піками}) = \text{хеш-значення відбитку}$

Рисунок А.8 – Хешування відбитків

Вирівнювання відбитків

Проблема, полягає в тому, що коли спочатку бралися відбитки, фіксувалось абсолютне зміщення хешу. Відносні хеші зразка та абсолютні хеші з бази даних ніколи не збігаються, якщо не почати записувати зразок саме з початку пісні, але це досить малоймовірно.

Всі відносні зміщення будуть на однаковій відстані один від одного.



різниця = зміщення БД від вихідної доріжки –
– зсув вибірки від запису

Рисунок А.9– Вирівнювання відбитків

Результати роботи програми

Кількість секунд	Правильні визначення	Точність у відсотках
1	13/29	45%
2	18/29	62%
3	23/29	79%
4	26/29	89%
5	29/29	100%
6	29/29	100%

Рисунок А.10 – Результати роботи програми

Продуктивність зберігання даних

Існує досить прямий компроміс між необхідним часом запису і обсягом необхідного сховища. Регулювання порога амплітуди для піків і значення певної змінної (яка відповідає за максимально можливий кут вибору сусіднього піку для зняття відбитків) додасть більше відбитків і підвищить точність за рахунок більшого простору.

Тип аудіоінформації	Розмір у МБ
MP3	322
Відбитки	202

Рисунок А.11 – Продуктивність зберігання даних

Висновки

У магістерській роботі було розроблено підсистему розпізнавання аудіозаписів. Створена підсистема задовольняє всім вимогам поставленого завдання.

Детально прокоментовано кожен пункт створення підсистеми. На наборі даних, який було обрано для тестування, підсистема продемонструвала стовідсоткове розпізнавання невідомого файлу, при прослуховуванні запису хоча б 5 секунд.

Прослухавши аудіо один раз і занісши відбитки до БД підсистема запам'ятовує аудіо. При відтворенні пісні і запису її через мікрофон, підсистема намагається порівняти звук з відбитками, які зберігаються у БД, повертаючи відтворену пісню (звісно, якщо така є у БД).

Щодо перспектив розвитку, у майбутньому планується зробити «оболонку» для підсистеми – сайт чи прикладну програму для смартфонів, щоб у користувачів була можливість використовувати підсистему де б їм було зручно. Заплановано значно розширити базу даних відбитків, проаналізувавши велику кількість аудіовідбитків. Надати перевіреним користувачам можливість поповнювати БД новими творами.

Рисунок А.12 – Висновки