

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання
(повна назва)

Кафедра інформаційно-мережної інженерії
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти другий (магістерський)
(рівень вищої освіти)

Оптимізація параметрів кодування відеоданих при передачі по каналу
зв'язку з обмеженою пропускнуою здатністю
(тема)

Виконав:

студент 2 курсу, групи ІМІзм-19-2
Мороз С.О.
(прізвище, ініціали)

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

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

Освітня програма Інформаційно-
мережна інженерія
(повна назва освітньої програми)

Керівник доц. Ширяєв А.В.
(прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____ Безрук В.М.
(підпис) (прізвище, ініціали)

2021 p.

Не містить відомостей, заборонених до відкритого публікування

Студент _____ */Мороз С.О./*
Керівник _____ */Ширяєв А.В./*

Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання
Кафедра інформаційно-мережної інженерії
Рівень вищої освіти другий (магістерський)
Спеціальність 172 Телекомунікації та радіотехніка
(код і повна назва)
Тип програми Освітньо-наукова
(освітньо-професійна або освітньо-наукова)
Освітня-наукова програма Інформаційні мережі зв'язку
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«_____» _____ 2021 р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Морозу Сергію Олексійовичу
(прізвище, ім'я, по батькові)

1. Тема роботи Оптимізація параметрів кодування відеоданих при передачі по каналу зв'язку з обмеженою пропускну здатністю

затверджена наказом по університету від «25» березня 2021 р. № 33 Стз

2. Термін подання студентом роботи до екзаменаційної комісії «25» травня 2021 р.

3. Вихідні дані до роботи Дослідити проблематику розвитку інфокомунікацій з позицій можливості забезпечення ефективної передачі відеоданих з максимальною якістю.

Визначити умови ефективної передачі відеоданих. На основі аналізу існуючих технологій запропонувати концепцію підходу до управління бітовою швидкістю відеопотоку.

Розглянути можливість реалізації запропонованої концепції у існуючих обчислювальних системах. Розробити механізм контролю швидкість передачі у каналу зв'язку

4. Перелік питань, що потрібно опрацювати в роботі _____

1. Особливості відеотрафіку як типу даних

2. Загальні підходи до обробки відеоданих в умовах необхідності забезпечення якості відеосервісу

3. Концептуальна модель методу контролю об'єму мультимедійних даних на базі тривимірної моделі трансформованого фрагменту кадру

4. Організація службових даних для методу управління бітовою швидкістю

5. Механізм контролю швидкість передачі

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (слайдів) Слайди у форматі Power Point у кількості 18 слайдів

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів роботи	Терміни виконання етапів роботи	Примітка
1	<i>Ознайомлення із завданням. Уточнення ТЗ.</i>	25.03.2021	
2	<i>Підбір літератури за темою роботи.</i>	26.03 – 02.04.21	
3	<i>Виконання розділу 1</i>	03.04 – 15.04.21	
4	<i>Виконання розділу 2</i>	16.04 – 22.04.21	
5	<i>Виконання розділу 3</i>	23.04 – 28.04.21	
6	<i>Виконання розділу 4</i>	29.04 – 02.05.21	
7	<i>Виконання розділу 5</i>	03.05 – 10.05.21	
8	<i>Оформлення пояснювальної записки</i>	11.05 – 19.05.21	
9	<i>Оформлення презентаційного матеріалу, підготовки до захисту у ЕК</i>	20.05 – 25.05.21	

Дата видачі завдання _____ 25.03.2021

Студент _____

(підпис)

Керівник роботи _____

(підпис)

доц. Ширяєв А.В.

(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка: с., рис., табл., джерел, додаток

ПРИНЦИП ОПТИМАЛЬНОСТІ БЕЛМАНА, КОЛІРНА СУБДИСКРЕТИЗАЦІЯ, JPEG, ТРАНСФОРМАНТА ДКП, СЕМАНТИЧНІ ОСОБЛИВОСТІ ЗОБРАЖЕННЯ

Об'єкт дослідження – методи управління бітовою швидкістю відеопотоку на рівні внутрішньокадрової обробки структурних одиниць відеоданих.

Мета роботи – оптимізація вибору параметрів кодування відео потоку у процесі передавання у мережу.

Проведено аналіз існуючих підходів до адаптації бітовою швидкості відеоданих, що знаходяться у мережу, до її пропускну здатності. Досліджено їх особливості та з урахуванням цього розроблено загальну концепцію побудови методу управління інтенсивністю відеоданих, орієнтованого на забезпечення максимальної якості відеоданих.

THE ABSTRACT

Explanatory note: p., fig., tabl., sources, app.

PRINCIPLE OF BELMAN'S OPTIMALITY, COLORED
SUBDISCRATING, JPEG, TRANSFORMATION DKP, SEMANTIC
FEATURES OF IMAGES

Object of research - methods of control the bit rate of video stream at the level of intra-frame processing of video units.

The purpose of the work is to optimize the choice of video stream encoding parameters in the process of transmission to the network. The analysis of existing approaches to the adaptation of the bit rate of video data found in the network to its bandwidth is carried out. Their features are investigated and in this context the general concept of building a method for managing the intensity of video data, focused on ensuring the maximum quality of video data.

ЗМІСТ

С.

ПЕРЕЛІК СКОРОЧЕНЬ.....			
ВСТУП.....			
1 ОСОБЛИВОСТІ ВІДЕОТРАФІКУ ЯК ТИПУ ДАНИХ.....			
1.1	Загальні	характеристики	відеопотоку
.....			
1.1.1		Формат	пікселя
.....			
1.1.2		Розмір	кадру
.....			
1.1.3		Формат	кадру
.....			
1.1.4		Частота	кадрів
.....			
1.1.5		Бітова	швидкість
.....			
1.2	Метрики	оцінювання	якості відео потоку
.....			
1.2.1	Показники	оцінки процесу	передавання відеотрафіку
.....			
1.2.2	Показники	оцінки прийнятих	відеоданих
.....			
2 ЗАГАЛЬНІ ПІДХОДИ ДО ОБРОБКИ ВІДЕОДАНИХ В УМОВАХ НЕОБХІДНОСТІ ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ВІДЕОСЕРВІСУ.....			
2.1	Базові	напрямки	обробки відеоданих
.....			
2.2	Вимоги до технологій кодування відеоданих на рівні окремих кадрів відеопотоку.....		
2.3	Підходи, спрямовані на забезпечення зменшення просторової надмірності відеоданих.....		
2.4	Режими	роботи	JPEG
.....			

2.5	Методи адаптації інтенсивності відео потоку до пропускну	здатності мережі.....
3	КОНЦЕПТУАЛЬНА МОДЕЛЬ МЕТОДУ КОНТРОЛЮ ОБ'ЄМУ МУЛЬТИМЕДІЙНИХ ДАНИХ НА БАЗІ ТРИВИМІРНОЇ МОДЕЛІ ТРАНСФОРМОВАНОГО ФРАГМЕНТУ КАДРУ.....	
3.1	Вимоги до методу контролю об'єму мультимедійних даних	
3.2	Представлення фрагментів кадру у вигляді тривимірних об'єктів	
3.3	Реалізація алгоритму на базі метода динамічного програмування	
3.4	Оцінка складності алгоритму на базі динамічного програмування	
4	ОРГАНІЗАЦІЯ СЛУЖБОВИХ ДАНИХ ДЛЯ МЕТОДУ УПРАВЛІННЯ БІТОВОЮ ШВИДКІСТЮ.....	
4.1	Вимоги до необхідної службової інформації	
4.2	Формування службових даних	
4.3	Оцінка об'єму службових даних	
5	МЕХАНІЗМ КОНТРОЛЮ ШВИДКІСТЬ ПЕРЕДАЧІ.....	
	ВИСНОВКИ.....	
	ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.....	
	ДОДАТОК А СЛАЙДИ ПРЕЗЕНТАЦІЇ.....	

ПЕРЕЛІК СКОРОЧЕНЬ

FPS – (frames per second) – частота слідування кадрів;

PSNR – (Pulse Signal to Noise Ratio) пікове відношення сигнал/шум;

ДКП – дискретно-косинусне перетворення;

MSE – (mean square error) – середньоквадратичне відхилення;

HD – (high definition) – відео високої роздільної здатності;

UHD - (high definition) – відео ультра-високої роздільної здатності;

SD - (small definition) – відео низької роздільної здатності;

DVD – (digital versatile disc) – формат відеокодування;

RGB – (red, green, blue) – трьохколірна модель;

YCrCb – (y, chromatic red, chromatic blue) – яскравісно-хроматична колірна модель;

EBCOT – (embedded block coding with optimal truncation) – режим внутрішньо кадрового кодування з оптимальним відсіченням структурних одиниць у схемі Jpeg2000;

SBR – (scalable bit rate) – масштабована бітова швидкість – рекомендація щодо контролю бітової швидкості у стандарті MPEG4;

MPEG4 – стандарт відео кодування.

ВСТУП

Розвиток сучасних інфокомунікацій відзначається постійною динамікою та здійснюється під впливами багатьох чинників, серед яких головними є наступні:

- потреби суспільства у тих чи інших сервісах;
- розвиток та удосконалення технологій;
- зростання вимог користувачів відносно можливостей та якісних характеристик наявних сервісів.

На сьогодні тенденція розвитку глобальної інформаційно-комунікаційної мережі відбувається у таких напрямках:

- збільшення кількості клієнтських мобільних та стаціонарних пристроїв, що мають можливість підключення до Всесвітньої мережі [1];
- поява нових мережевих сервісів;
- зростання обсягів інформації, що передається мережею.

За прогнозами Cisco, на 2022 рік кількість користувачів Всесвітньої мережі буде складати порядку 4,8 млрд. у розпорядженні яких буде близько 28,5 млрд. пристроїв для доступу.

Окрім цього, очікується подальше збільшення відсоткової долі мультимедійного трафіку, зокрема – відеоданих, які на 2022 рік складатимуть більш ніж 82%, усього об'єму даних, а саме – 325 ексабайт на місяць проти 93,5 ексабайт на місяць 2017 року.

У той же самий час, на фоні збільшення обсягів відеотрафіку, спостерігається стійка тенденція до зростання клієнтських вимог відносно якості відеоданих.

Як наслідок, зокрема, прогнозується поступове зменшення об'єму відео формату SD з 50,3% на 2017 рік до 20,9%, при цьому прогнозується також незначне зростання обсягів відео HD формату з 46,4% до 56,8 %.

У свою чергу, відсоток відео формату UHD може збільшитися у 6 разів з 3,3% до 22,3%.

Стосовно окремих груп сервісів, зокрема, “відео на запит”, прогноз є більш радикальним, а саме – зростання обсягів UHD відео майже у 15 разів (з 2,1% до 35,1%) на фоні зменшення відео HD формату (з 75% до 60,6%) та майже 5 кратного зменшення об’єму SD відео (з 23% до 4,3%).

Окрім тенденції зміни форматів відео, що передається, у бік його зростання, очікується також ріст вимог до інших параметрів відеопотоку, зокрема таких, як частота кадрів за секунду та показник якості за однією з метрик оцінювання.

Разом з тим, прогнозується збільшення середньої пропускної здатності для фіксованих мереж з 39 Мбіт/с до 75,4 Мбіт/с.

За результатом аналіз прогнозу тенденції розвитку глобальної інфокомунікаційної мережі можна зробити такі висновки:

- прогнозується ріст більш ніж втричі відсотку передаваних у мережі відеоданих;
- отримують подальший розвиток більшість технологій та сервісів надання мультимедійних послуг;
- спостерігається стійка тенденція зростання вимог до якості прийнятих відеоданих;
- тенденція росту обсягу даних, що передаються у мережі, суттєво випереджає зростання пропускної здатності мережі.

Тобто, у таких умовах необхідно:

- створити умови для забезпечення гарантованої доставки відеотрафіку кінцевим користувачам;
- передбачити ефективні механізми запобігання перевантажень мережі;
- забезпечити можливість підтримки якості відеоданих на максимально можливому рівні.

Таким чином, розробка, вдосконалення та застосування методів та технологій, спрямованих на оптимізацію процесу обробки відеоданих є актуальним науково-прикладним завданням.

Сьогодні для раціонального використання мережевих каналів застосовуються 2 основних підходи:

- управління трафіком на рівні мережевих вузлів (встановлення пріоритетів трафіку, резервування каналів, маршрутизація та інше) [2];
- зменшення об'ємів даних, що поступають у мережу на рівні кінцевих вузлів [3].

Перший підхід не може у достатній мірі забезпечити раціонального використання мережі, оскільки сприяє лише перерозподілу загального навантаження на мережу, яке у сумі залишається сталим.

У цьому випадку більш ефективним є другий підхід, який базується на використанні технологій кодування відео, що здатні зменшити об'єм відеотрафіку, який, у свою чергу, є однією з причин перевантажень.

Скачки бітової швидкості можуть призвести до великих проблем для багатьох протоколів транспортування і зберігання даних. Так мережі на основі комутації пакетів можуть підтримувати змінну швидкість передачі, але середня пропускна здатність в будь-який момент часу обмежена певними факторами, залежними від швидкості передачі і перевантаженості каналу.

Тому актуальною науковою задачею є розвиток підходів з контролю бітової швидкості кодера для її відповідності швидкостям транспортування в мережі передачі даних.

В даний час, для вирішення завдань адаптації бітової швидкості при передачі відеоданих по каналах з постійною швидкістю використовуються буфери в кодері і декодері. При цьому, чим вище варіація бітової швидкості, тим більшим буде необхідний обсяг буфера і тим довше затримка декодування. Крім того, таким методом можна впоратися з довільними стрибками бітової швидкості, так як обсяги буферів і затримки при передачі і декодування мають обмеження. Тому необхідно реалізувати таку технологію

контролю бітової швидкості з використанням зворотного зв'язку, яка при зниженні бітової швидкості також враховує якість реконструкції і час обробки кадру відеопослідовності.

Для отримання на виході кодера потоку, який відповідає пропускній здатності каналу передачі даних, застосовують алгоритми управління бітової швидкістю. Необхідна вихідна швидкість досягається шляхом регулювання параметра квантування. Залежно від області застосування може використовуватися стиснення як з постійною вихідною швидкістю (CBR - Constant Bit Rate), так і з пере-змінної (VBR - Variable Bit Rate) [18]. Для стиснення з постійною швидкістю в залежності від призначення алгоритму може варіюватися розмір буфера для кодера і декодера. Якщо необхідний більш точний алгоритм контролю вихідний швидкості кодера використовується многопроходное кодування (multi-pass encoding). Великого поширення набула двопрхідні схема кодування: перший прохід - етап попереднього аналізу і встановлення необхідних параметрів кодування; другий прохід - кодування з раніше встановленими параметрами. При цьому вибрані параметри можуть застосовуватися як до окремого блоку в кадрі, так і до всього кадру.

1 ОСОБЛИВОСТІ ВІДЕОТРАФІКУ ЯК ТИПУ ДАНИХ

1.1 Загальні характеристики відеопотоку

У спрощеному розгляді відео - це сукупність статичних зображень (кадрів), які змінюють один одного у часі. Ця впорядкована множина кадрів у прив'язці до часу називається відеопотоком. У загальному вигляді для опису характеристик відеопотоку використовуються ряд най розповсюджених показників, до яких відносяться:

- формат пікселя;
- розмір кадру;
- формат кадру;
- частота слідування кадрів;
- бітова швидкість;
- одна або кілька метрик якості прийнятих відеоданих.

1.1.1 Формат пікселя

Піксель не несе у собі ніякої інформації, крім його кольору. Однак, сприйняття кольору людським зором є суб'єктивним, тому було створено системи опису та передачі кольору, які були б прийнятні для більшості людей.

Зокрема колір, видимий нами, є досить складним за спектром частот світла, тому передати його у цифровому вигляді вкрай складно, а, у свою чергу, відобразити ще складніше. Тому зараз наближено з достовірністю майже 99,9% для опису майже будь-якого кольору видимого діапазону використовується поєднання трьох колірних каналів - червоного, зеленого та синього. Такий спосіб представлення зветься форматом RGB [4].

Частіше за все, коли йде мова про формат RGB, мова йде про формат RGB24, у якому для опису кожного з колірних каналів відводиться рівно по 8 біт інформації.

Таким чином ми можемо передати 256 градацій кожного кольору й усього 16,777,216 усіляких відтінків. Але на практиці, при зберіганні такий формат представлення кольору практично не використовується, унаслідок великих об'ємів графічних файлів.

Для компактного зберігання або у випадку відео - кадр, початково представлений палітрою RGB, переводиться у простір опису YCrCb, де Y - ярісна компонента, Cr і Cb, відповідно, хроматична синя й хроматична червона компоненти.

Таке перетворення виконується з огляду на те, що людське око більш сприйнятливим до яскравісного сигналу, ніж до колірному. Тому це перетворення дозволить досягти певного зменшення обсягу даних при мінімальних візуальних втратах.

Процес перетворення колірної моделі кадру представлено наступною системою співвідношень:

$$\begin{cases} Y = 0,299 \cdot R + 0,587 \cdot G + 0,114 \cdot B, \\ Cb = (B - Y) \cdot 0,5643 + 128, \\ Cr = (R - Y) \cdot 0,7132 + 128. \end{cases} \quad (1.1)$$

У свою чергу, зворотнє колірне перетворення має наступний вигляд:

$$\begin{cases} G = Y - 0,714 \cdot (Cr - 128) - 0,334 \cdot (Cb - 128), \\ R = Y + 1,402 \cdot (Cr - 128), \\ B = Y + 1,772 \cdot (Cb - 128). \end{cases} \quad (1.2)$$

Такий спосіб опису окремого пікселя (і усього зображення) вносить незворотні зміни у початкове зображення та веде до втрати до 75%

інформації про розподіл колірності у зображенні. Таким чином, забезпечується ефективне зниження об'єму даних [3-5].

1.1.2 Розмір кадру

Розмір, або роздільна здатність кадру характеризує ширину та висоту його видимої частини.

У загальному випадку, ці розміри можуть бути довільними [6]. Проте їх конкретні значення регламентуються базовими стандартами відео. Зокрема, зараз широко розповсюджені розміри 640x480, 720x480, 720x576, 1280x720 та 1920x1080 пікселів.

Наприклад, роздільну здатність 720x576 має більша частина європейських DVD-стандартів. У свою чергу, 1280x720 та 1920x1080 – роздільні здатності відео форматів HD-Ready та FullHD

1.1.3 Формат кадру

Навіть знаючи розміри кадру, ми не можемо представити масив пікселів у більш зручній формі, не маючи інформації про спосіб "упакування" кадру.

У найпростішому випадку, у сутності, для кожного рядку пікселів робиться опис підряд бітів кожного закодованого пікселя. Тобто здійснюється опис стількох рядків, скільки у наявності за висотою кадру і при цьому кожен піксель описується окремо. Процес порядкового опису кадру має назву розгортання [6].

Таке розгортання, що передбачає окремий опис пікселя у кожному рядку називається прогресивним (Progressive).

У випадку, коли виконується черезрядкове розгортання (Interlaced), досягається зменшення об'єму даних для опису кадру, однак при цьому може виникати "ефект гребінки", коли той самий об'єкт перебуває в різних положеннях щодо парних і непарних рядків.

Такий ефект може бути у певній мірі компенсовано методами деінтерлісинга.

З огляду на це, представлення формату кадру та типу розгорнення фіксують відповідними позначеннями. Наприклад: 576i, 720p, 1080i, 1080p, де зазначена кількість рядків (висота кадру) і тип розгорнення.

1.1.4 Частота кадрів

Як було зазначено, під відеопотоком мається на увазі множина окремих кадрів, що змінюються у часі. Тобто, частота слідування кадрів є однією з ключових характеристик відео [3,4].

Говорячи про певні величини показника частоти слідування кадрів, або fps (frames per second), слід зазначити, що сьогоднішні значення на рівні 25 та 30 кадрів, властиві більшості технологіям кодування відео, отримані експериментальним шляхом та являють собою “золоту середину”.

Це пояснюється тим, що з одного боку, зменшення показника fps веде до зниження візуальної якості відео, яке найбільш проявлюється у сценах з високою динамікою.

З іншого боку, ріст fps веде до збільшення обчислювального навантаження та зростання об'єму файлу.

Що стосується телебачення, то європейський показник fps=25 к/с. У той же час, аналогічний американський показник для багатьох стандартів має значення 23.976, 24 та 29.97 кадрів у секунду [6].

Такі значення fps є наслідком того, що, зокрема, у стандарт американського телемовлення NTSC закладений дільник 1.001. Тобто: $23.976 = 24/1.001$, а $29.97 = 30/1.001$.

Відповідно, при показі відео відбудеться зовсім невелике вповільнення, яке не буде помітно глядачеві, але якщо це музичний концерт, то швидкість показу настільки критична, що краще зрідка пропускати кадри й глядач нічого не помітить.

Хоча, частіше за все американське телебачення трансліює 30 черезрядкових кадрів (або 59.94 напівкадра у секунду, що відповідає частоті американської електромережі). При цьому кадри виходять "методом спуску" (3:2 pulldown).

Сутність методу полягає у тому, що існує 2 повних кадра й 5 напівкадрів. Отже, даними з першого кадру заповнюються перші 3 напівкадра, а із другого, що залишилися 2. Тобто послідовність напівкадрів така: [1 top, 1 bottom], [1 top, 2 bottom], [2 top, 3 bottom], [3 top, 3 bottom], [4 top, 4 bottom] і т.д. Де top - верхні рядки (поля, fields), а bottom нижні, тобто, непарні й парні починаючи зверху відповідно.

Частота кадрів може бути й змінною, проте це потребує додаткових обчислювальних навантажень на реалізуючи систему.

1.1.5 Бітова швидкість

У загальному випадку, стосовно відеоданих - під бітовою швидкістю розуміється об'єм компактно представлених даних, що передаються у одиницю часу.

Розглянемо докладніше, з яких складників формується даний показник.

Як було сказано вище, відеопоток складається з певної кількості кадрів визначеного розміру та формату, що трансліюється у одиницю часу t [4,5].

Оцінка бітової швидкості відео потоку, відносно якого не було задіяно технології кодування здійснюється відповідно наступному виразу:

$$V(t) = \frac{h \times w \times b \times fps}{t} \quad (1.3)$$

де h та w - відповідно ширина та висота одного кадру у пік селях;

b - кількість біт для опису одного пікселя кадру;

fps - частота слідування кадрів;

Як можемо бачити з виразу (1.3), бітова швидкість S відеопотоку роздільної здатності 1280×720 при $\text{fps} = 30$ становитиме $663,55$ Мбіт/с, що взагалі виключає можливість передавання даного відео мережею.

Зрозуміло, що за таких умов необхідно задіяти методи кодування відеоданих.

Тоді загальний вираз для оцінки бітової швидкості може бути записано наступним чином:

$$V(t) = \frac{h \times w \times b \times \text{fps} \times f(\text{enc}) \times f(r)}{t} + V(c), \quad (1.4)$$

$f(r)$ - функціонал, що визначає спосіб розгорнення, у режимі *progressive* дорівнює 1 а у *interlaced*, відповідно, 0,5;

$f(\text{enc})$ - функціонал, що визначається технологією кодування та наближено може бути прирівняним до коефіцієнта стиснення, зазвичай $f(\text{enc}) < 1$;

$V(c)$ - об'єм службових даних для декодування відео.

1.2 Метрики оцінювання якості відео потоку

Для оцінки ефективності передавання відео потоку сьогодні існує ряд показників, що стосуються як самого процесу трансляції, так і якісних показників відеоданих на прийомній стороні з позицій цілісності прийнятих даних відносно вихідних.

1.2.1 Показники оцінки процесу передавання відеотрафіку

Вимоги до якості сервісів, що надаються в інфокомунікаційних мережах, регламентуються системою вимог QoS (Quality of Service).

Стосовно відеопотоку, найбільш суттєвими показниками QoS, які характеризують трансляцію відеоданих у мережі, є наступні [7]:

- затримка Δt передачі пакету;
- джиттер Δ_t , або різниця між максимальним $\Delta_{t(\max)}$ та мінімальним $\Delta_{t(\min)}$ значеннями затримки передачі пакетів;
- відсоток втрачених пакетів.

Затримка Δt передачі пакету визначається різницею часу відправки та прийому, тобто:

$$\Delta t = T_{\text{res}} - T_{\text{trns}}, \quad (1.5)$$

де T_{res} та T_{trns} – відповідно абсолютні часові відмітки прийому та початку передачі пакету.

У свою чергу, величина джиттера J_t у момент часу t визначається за значеннями величин $RTT_{t,\min}$ та $RTT_{t,\max}$, як максимальне відхилення односторонніх затримок від середнього значення, що показується наступним виразом:

$$J_t = \frac{1}{2} \max \left[RTT_{t,\max} - RTT_{t,\text{cp}}, \quad RTT_{t,\text{cp}} - RTT_{t,\min} \right], \quad (1.6)$$

де $RTT_{t,\min}$, $RTT_{t,\max}$ та $RTT_{t,\text{cp}}$ - максимальний, мінімальний, та, відповідно, середній час двохсторонньої затримки (round trip time).

Відсоток втрачених у ході передавання відеопотоку пакетів, у момент вимірювання t визначається, відношенням середньої кількості втрачених

пакетів $\theta_{втр,t}$ до середньої кількості прийнятих $\theta_{пр,t}$ пакетів на прийомній стороні, а саме:

$$\delta_t = \frac{\theta_{втр,t}}{\theta_{пр,t}} \times 100\% \quad (1.7)$$

1.2.2 Показники оцінки прийнятих відеоданих

Метрики QoS, що характеризують якість надання відеосервісу, розглядають відеопотік як специфічний тип трафіку.

У той же час, для оцінювання якості прийнятого окремого кадру, або фрагменту відео потоку, утвореного сукупністю кадрів, використовується ряд інших метрик.

При цьому, у залежності від певних умов передавання, обробки та застосування відеоданих можуть бути обрані ті чи інші метрики оцінки оцінювання.

Основна розбіжність у виборі викликана тим, що оцінювання за одними й тими показниками якості зазвичай може не корелюватися з візуальною оцінкою якості [3, 4, 8, 9].

Отже, можуть виникати умови, коли метрика оцінювання свідчить про те, що прийняті данні мають високу якість, у той же час, відеопотік є візуально спотвореним. Те ж саме стосується випадку, коли попри низьку оцінку якості за однією з метрик відеопотік відзначається візуально прийнятною якістю.

Нехай є два зображення: оригінал та відновлене зображення розміром $H \times W$, тоді одним із простих критеріїв оцінки втрати якості є середньоквадратичне відхилення значень пікселів кодованого зображення від оригіналу (СКВ або mean square error -MSE), а саме:

$$M = \frac{1}{HW} \sum_{x=1, y=1}^{H, W} (I(x, y) \times I'(x, y))^2, \quad (1.8)$$

де $I(x, y)$ та $I'(x, y)$ - відповідно, початковий та відновлений пікселі зображення.

Згідно цього критерію, зображення буде значно спотворене при зміні яскравості всього на 5% (око цього не помітить - у різних моніторів налаштування яскравості варіюється набагато сильніше).

У той же час, зображення з "снігом" - різкою зміною кольору окремих пікселів, смугами або "муаром" будуть визнані "такими, що майже не змінилися".

Наступною метрикою оцінювання якості є максимальне відхилення від оригіналу:

$$\Delta D = |I(x, y) - I'(x, y)| \quad (1.9)$$

Дана метрика у край чутлива до спотворення окремих пікселів, тобто, зображення може змінити тільки один піксель, проте, за цією метрикою зображення буде визнано значно зіпсованим.

На практиці частіше за все використовується показник якості зображення такий, як пікове відношення сигнал/шум (PSNR - peak-to-peak signal-to-noise ratio).

У даному випадку даний показник може бути описано наступним виразом:

$$\text{PSNR} = 10 \log_{10} \frac{255^2 HW}{\sum_{x,y} (I(x, y) - I'(x, y))^2} \quad (1.10)$$

Цей показник, у сутності, аналогічний середньоквадратичному відхиленню, однак користуватися ним зручніше за рахунок логарифмічного

масштабу шкали. Йому властиві ті ж недоліки, що й середньоквадратичному відхиленню.

Найкраще втрати якості зображень оцінюються зоровим апаратом людини.

У цьому випадку відмінним вважається кодування, при якому неможливо візуально розрізнити вихідне й відновлене зображення. Прийнятним - коли сказати, яке із зображень зазнало кодування, можна тільки порівнюючи два зображення, що перебувають поруч.

При подальшому збільшенні ступеня стиснення, як правило, стають помітні побічні ефекти, характерні для даного алгоритму.

На практиці, навіть при відмінному збереженні якості, у зображення можуть бути внесені регулярні специфічні зміни.

Тому алгоритми кодування з втратами не рекомендується використовувати при стисненні зображень, які надалі будуть друкуватися з високою якістю, або оброблятися програмами розпізнавання образів. Неприємні ефекти з такими зображеннями можуть виникнути навіть при простому масштабуванні зображення.

До інших найбільш уживаним критеріїв оцінки якості зображення, які найчастіше збігаються з візуальними оцінками, відносяться [4, 9]:

Коефіцієнт крос-кореляції, а саме:

$$K_{cc} = \frac{\sum_{x=1, y=1}^{H, W} (I(x, y) \times I'(x, y))}{\sum_{x=1, y=1}^{H, W} (I(x, y))^2}, \quad (1.11)$$

та вірність зображення (image fidelity):

$$K_{if} = \frac{\sum_{x=1, y=1}^{H, W} (I(x, y) - I'(x, y))}{\sum_{x=1, y=1}^{H, W} (I(x, y))^2}. \quad (1.12)$$

На сьогодні найбільш часто оцінка якості прийнятих відео або графічних даних виконується за показником PSNR або MSE. Це стосується майже усіх сервісів та додатків, що пов'язані з передаванням та обробкою відео та графічних даних. У той же час, деякі спеціальні системи, такі, як:

- системи дистанційного зондування;
- біометричні системи високої чіткості;
- т.з. “дистанційна” або “відеомедицина”;

використовують для оцінки якості показники вірності зображення, крос-кореляції або SSIM (структурна подібність).

Така закономірність пояснюється рядом причин, серед яких наступні:

- спеціальні системи розроблювалися набагато пізніше вже традиційних сервісів, що за інерцією використовують показники PSNR та MSE;
- ряд спецсистем працюють локально, тому їх обчислювальні можливості мають більший ресурс, що може бути зосереджено на обробці даних за більш складними показниками.

2 ЗАГАЛЬНІ ПІДХОДИ ДО ОБРОБКИ ВІДЕОДАНИХ В УМОВАХ НЕОБХІДНОСТІ ЗАБЕЗПЕЧЕННЯ ЯКОСТІ ВІДЕОСЕРВІСУ

2.1 Базові напрямки обробки відеоданих

Як було зазначено раніше, доцільним є застосування методів, орієнтованих на обробку відеоданих на рівні кінцевих вузлів.

У загальному випадку такі методи умовно можна поділити на дві основні групи, а саме:

- технології кодування [4];
- технології, спрямовані на узгодження параметрів відеопотоку та пропускної здатності мережі [9, 10].

Особливістю першої групи є те, що методи, які відносяться до неї, виконують завдання зменшення об'єму відеоданих, що надходять до мережі, орієнтуючись на встановлені показники якості відео на прийомній стороні або на величину заданого коефіцієнта стиснення.

При цьому дані параметри мають обернено пропорційний характер відносно один одного, тобто, ріст коефіцієнта стиснення приводить у загальному випадку до зменшення якості як відео потоку взагалі, так і окремих кадрів, як його структурних одиниць.

Тут під коефіцієнтом стиснення будемо розуміти величину C , що характеризує відношення початкового V_d об'єму даних до об'єму V_c кодованих даних, тобто [4]:

$$C = \frac{V_d}{V_c}. \quad (2.1)$$

Сутність технологій кодування полягає у тому, що під час обробки зменшується надмірність представлення відео потоку як у просторі, так і у

часі. Тобто, зменшення надмірності відбувається як на рівні відеопотоку вцілому, так і на рівні окремих кадрів.

2.2 Вимоги до технологій кодування відеоданих на рівні окремих кадрів відеопотоку

До усіх без винятку існуючих технологій кодування окремих кадрів, спрямованих на усунення просторової надмірності зображень, можуть бути

висунені такі вимоги [8]:

- забезпечення достатнього рівня зменшення початкового об'єму зображення;
- внесення мінімальних спотворень початкових зображень під час обробки;
- відносно невисока складність алгоритму і, відповідно, вимогливість до обчислювальної здатності реалізуючих пристроїв.

З огляду на сферу застосування, перелічені вимоги можуть доповнюватися.

Зокрема, технології компресії потоку відеокадрів, які застосовуються на базі відеоінформаційних ресурсів, що передбачають безпосередню трансляцію даних у мережу або на локальній пристрій відображення, повинні забезпечувати обробку даних у реальному масштабі часу з мінімально можливими часовими затримками.

У той же самий час, дана вимога не є ключовою для технологій, які мають на меті кодування зображень для подальшого зберігання.

Для даного випадку можуть бути задіяні технології та методи, що забезпечують у першу чергу необхідне зменшення даних у процесі кодування та заданий рівень якості.

2.3 Підходи, спрямовані на забезпечення зменшення просторової надмірності відеоданих

На сьогодні переважна більшість існуючих технологій кодування для зменшення просторової надмірності використовує кодування окремих кадрів на базі платформи JPEG [5].

Загальна схема обробки за алгоритмом JPEG приведена на рис. 2.1

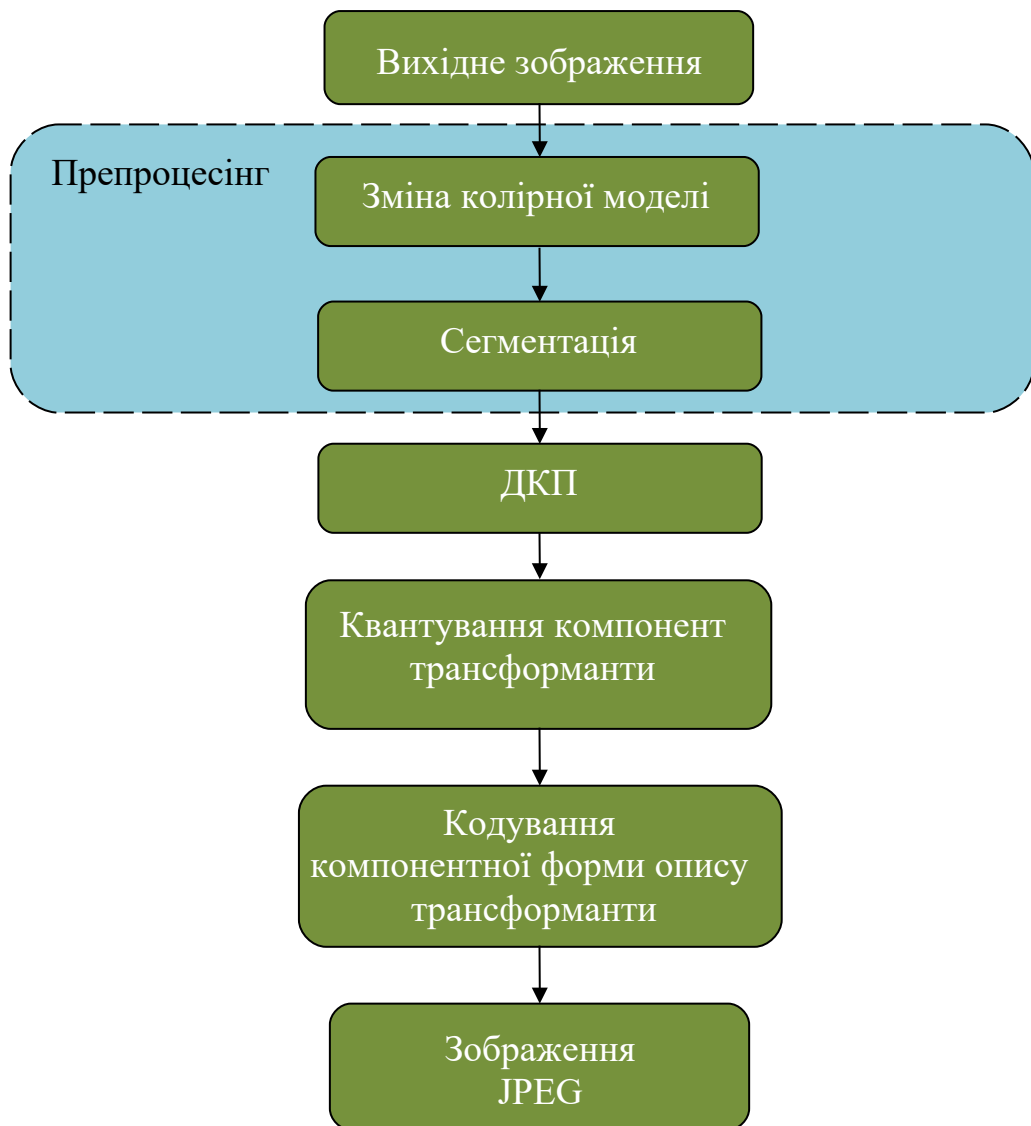


Рисунок 2.1 - Загальна схема обробки за алгоритмом JPEG

Обробка у схемі JPEG починається з етапу препроцесингу, на якому відбувається зміна колірної моделі та сегментація кадру.

На даному етапі обробки виконується переведення зображення з його компонентного представлення у хроматично-яскравісне представлення (ICT - Irreversible Color Transform) [4, 5], як показано виразом (1.1).

У цьому випадку, початковий кадр, представлений палітрою RGB, переводиться у простір опису YCrCb, де Y - яскісна компонента, Cr і Cb, відповідно, хроматична синя й хроматична червона компоненти.

На рис.2.2 а) наведено приклад такого перетворення для групи з 4 пікселів.

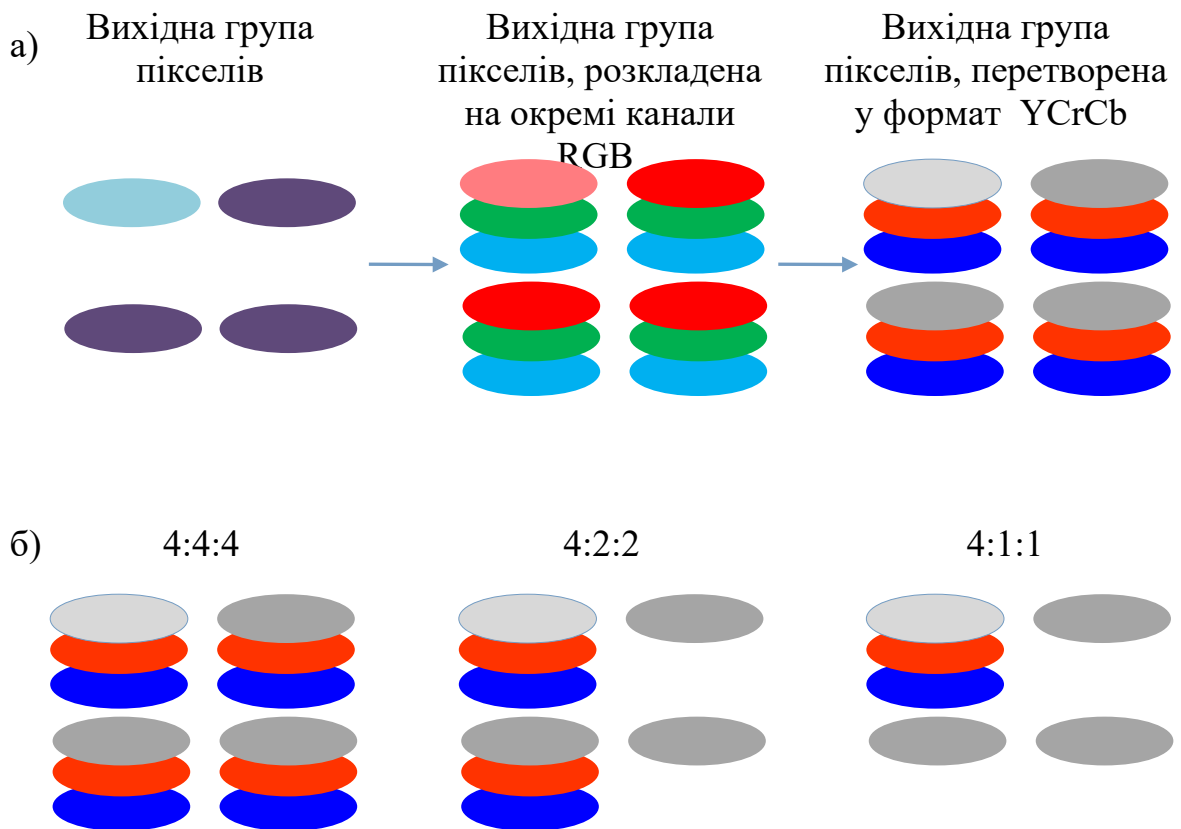


Рисунок 2.2 – Приклад зміни колірної моделі для групи пікселів (а) та приклади різних форматів колірної субдискретизації

Даний етап є підготовчим для подальшої зміни моделі колірної субдискретизації, під час якої одній компоненті яскравості Y ставиться у

відповідність декілька колірних компонент при тому, що у початковій схемі (4:4:4) одній компоненті Y відповідали по одній компоненті Cr і Cb.

На цьому етапі вносяться незворотні зміни у колірний опис пікселя, проте, це дає змогу знизити кількість даних щодо колірного опису зображення до 75% (у моделі 4:1:1).

Таким чином, даний етап обробки дозволяє ефективно знизити об'єм даних. Цей крок не робиться для півтонових зображень а також не вконується для компоненти яскравості.

Далі вихідне зображення розбивається на блоки 8x8, з кожного з яких формується три робочі матриці ДКП - по 8 біт окремо для кожної компоненти.

Зображення ділиться по компоненті Y - як і в першому випадку, а для компонент Cr і Cb виходить тільки одна робоча матриця ДКП, тому що матриці набираються через рядок і через стовпець.

Виконання такої операції можливе завдяки роботі в просторі YCrCb. На результуючому RGB зображенні, як показала практика, це позначається незначно.

Після колірних перетворень у схемі опису кадру виконується сегментація зображення.

Дана операція виконується з метою розподілу зображення на два й більш сегментів (підзображення). Це полегшує буферизацію даних зображення в пам'яті ЕОМ, прискорює їхню довільну вибірку з диска, та дозволяє зберігати зображення розміром понад 64x64 Кб.

Ключовим у сімействі технології JPEG є етап трансформації блоків 8x8 кадру шляхом дискретно-косинусного перетворення (ДКП).

Дане перетворення є різновидом перетворення Фур'є, що дає змогу із просторового представлення блоку перейти до спектрального [3].

ДКП перетворить матрицю пікселів розміру N x N у матрицю частотних коефіцієнтів (трансформанту) відповідного розміру.

Стандарт JPEG застосовує ДКП до одиниць даних (блоків) розміром 8x8 пікселів, при цьому використовується головна особливість ДКП-трансформації, яка полягає у тому, що основна енергія частотних складових концентрується в невеликій області низьких частот.

При цьому амплітуда високочастотних складових або відносно невелика у порівняння з НЧ областю, або просто дорівнює нулю.

Передачі підлягають тільки ті частотні коефіцієнти матриці ДКП, величини яких перевищують прийняті граничні значення. Коефіцієнти нижче граничного значення вважаються нульовими.

Трансформація ДКП є підготовчою операцією до етапу квантування.

Етап квантування, у свою чергу, сприяє найсуттєвішому зменшенню об'єму зображення.

При цьому виконується поелементне ділення компонент трансформант на відповідні коефіцієнти матриці квантування (МК).

Коефіцієнти квантування підібрані таким чином, що їх величина збільшується у напрямку від коефіцієнтів, що відповідають НЧ-компонентам трансформанти, до тих, що відповідають ВЧ-компонентам (рис. 2.3).

Таким чином, ВЧ-компоненти квантуються найбільше.

3	5	7	9	11	13	15	17
5	7	9	11	13	15	17	19
7	9	11	13	15	17	19	21
9	11	13	15	17	19	21	23
11	13	15	17	19	21	23	25
13	15	17	19	21	23	25	27
15	17	19	21	23	25	27	29
17	19	21	23	25	27	29	31

Рисунок 2.3 - Матриця квантування з фактором якості 2

Технологія JPEG використовує два підходи побудови МК:

- матриця коефіцієнтів, прийнята за замовчуванням. Дві такі таблиці, одна для компонента яскравості (і для градації сірого кольору), а інша - для хроматичних компонентів, є результатом досліджень і отримані дослідним шляхом. Вони є частиною стандарту JPEG (рис 1.8).

- обчислюється таблиця коефіцієнтів квантування, що залежить від параметра Q , який задається користувачем. Формула (5) забезпечує зменшення коефіцієнтів у напрямку з лівого верхнього кута у правий нижній.

Яскравість

Колірність

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Рисунок 2.4 - Рекомендовані таблиці квантування

Після округлення квантованих компонент (при цьому значення, менші певної величини, прирівнюються до 0) у ВЧ та СЧ областях трансформанти утворюються ланцюжки з великої кількості нулів, що на етапі ймовірностатистичного кодування (арифметичного або Хафмана) дає змогу ефективно знизити загальний об'єм даних.

2.4 Режими роботи JPEG

Технологія JPEG передбачає кілька режимів обробки зображень, а саме [5]:

- baseline (базовий, за замовчуванням);
- progressive (прогресивний);
- hierarchial (ієрархічний режим).

Під час обробки у базовому режимі, зображення кодується повністю, за один технологічний цикл.

На відміну від базового, прогресивний режим містить додаткові операції обробки зображення.

У даному режимі зображення, оброблене за алгоритмом JPEG, постає у вигляді ряду прошарків (сканів).

При цьому, виконується послідовне завантаження кожного скану, що у результаті формує цілісне зображення. Перші скани, завантажені на дисплей, ілюструють огрублене та позбавлене багатьох деталей зображення.

Кожен наступний скан при цьому послідовно покращує якість зображення.

Існують такі реалізації даного режиму:

- метод спектральної вибірки (spectral selection);
- метод послідовного наближення (successive approximation).

У випадку методу спектральної вибірки, увесь масив компонент трансформанти розбивається на L частотних підгруп (рис.2.5 а), кожна з яких займає певний діапазон.

Метод послідовного наближення [5] передбачає розбиття компонент трансформанти на двійкові розряди, при цьому один або Q двійкових розрядів будуть формувати скан (рис.2.5 б).

- вплив поточного рівня завантаження мережі, що визначається загальною кількістю одночасних користувачів та обсягами трафіку, що приходить на кожного з них;

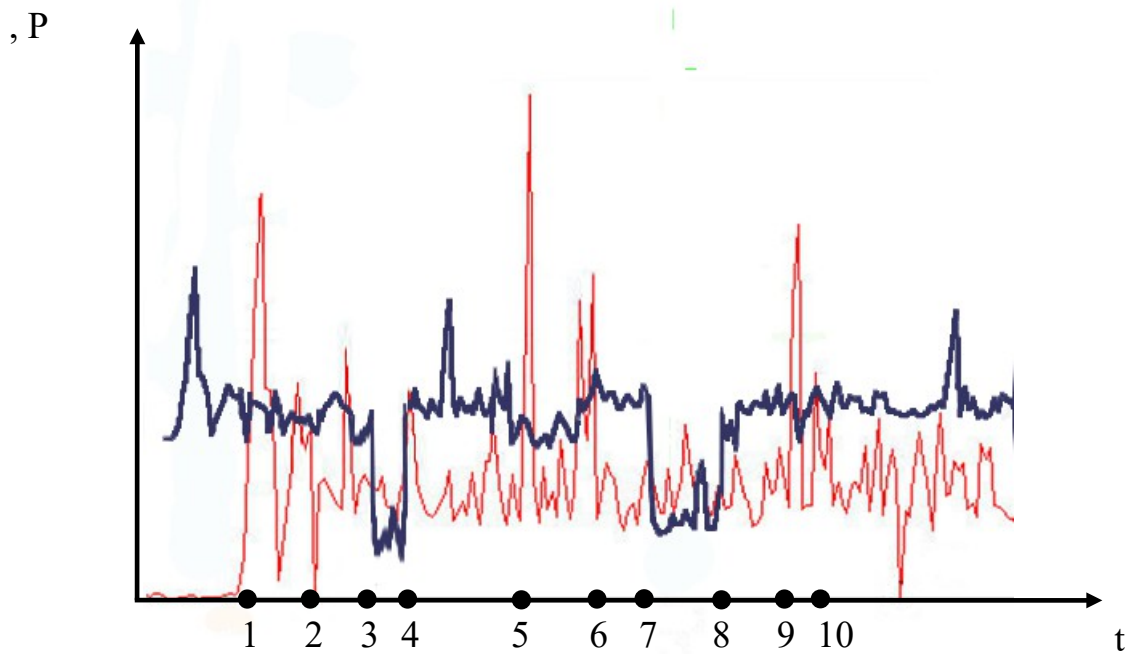


Рисунок 2.6 – Умови неузгодженості бітової швидкості відео потоку та пропускної здатності мережі

- характер трафіку, що необхідно опрацювати на кожному з мережевих вузлів. Зокрема, продуктивність роботи мережевих пристроїв залежить від розмірів окремих пакетів, які необхідно опрацювати.

Тут спостерігається чітка залежність, згідно якій зменшення розмірів пакетів веде до збільшення часу обробки масивів пакетів, та, як наслідок, росту ймовірності перевантаження окремих вузлів мережі.

Як показано рис. 2.3, на відрізках часу 3-4 та 7-8 спостерігається зниження пропускної здатності відносно інтенсивності трансльованого відеопотоку.

1) Коефіцієнт стиснення C не враховує особливості змісту як відеокадрів, так і їх окремих фрагментів.

Отже, розподіл інтенсивності у межах кадру за таких умов буде здійснюватися нераціонально.

2) Обробка у базовій схемі JPEG при відносно високих коефіцієнтах стиснення не дозволяє забезпечити прийнятну візуальну якість зображення на рівні архітектури алгоритму.

2.5 Методи адаптації інтенсивності відео потоку до пропускної здатності мережі

У загальному випадку, дані методи передбачають зміну одного або кількох показників відеопотоку відповідно до зміни пропускної здатності мережі.

Слід додати, що незважаючи на нагальну потребу у таких технологіях, сьогодні більшість таких механізмів існують у формі рекомендацій або базових архітектурних рішень, відкритих для майбутніх розробників. Найбільш відомими з даної групи механізмів сьогодні є:

1) Технологія SRC (Scalable Rate Control) [9], передбачена у пакеті рішень, що відносяться до загального стандарту MPEG 4. Дана концепція передбачає можливість контролю бітової швидкості $V(t)$ відеопотоку шляхом зміни кроку квантування K_q у процесі трансляції відео потоку, що еквівалентно виразу:

$$V(t) = f(K_q, V(0)), \quad (2.2)$$

де K_q - крок квантування;

$V(0)$ - початкова бітова швидкість відеопотоку.

Метод орієнтований на забезпечення максимально можливої якості відеоданих з огляду на доступну кількість біт для їх опису. Це складається з таких кроків:

- оцінка необхідної кількості біт для опису фрагменту відеопотоку;
- пошук необхідного показника K_q кроку квантування для заданих умов обробки відеопотоку;
- застосування знайденого показника K_q до усіх кадрів групи.

Іншими словами, у ході передавання відеопотоку, коли у момент часу t виникає необхідність зміни бітової швидкості, знаходяться параметри стиснення, які буде примінено у наступний $(t+1)$ момент часу для наступної групи кадрів. Тобто, поточна група кадрів транслюється без змін.

Такий підхід дає змогу змінювати бітову швидкість відео потоку, що кодується. У той же час, SRC має ряд недоліків, серед яких:

- відсутність стандартизованих інструментів забезпечення якості відео на фоні зміни параметрів квантування;
- зміна бітової швидкості виконується на рівні групи кадрів (GOP);
- у рамках окремого кадру бітова швидкість розподіляється нераціонально;
- метод ефективний лише для певного діапазону бітових швидкостей та роздільних здатностей кадру.

До обмежень SRC також можна віднести орієнтованість на виключно MPEG 4/Part 10, оскільки дана технологія присутня у вигляді рекомендацій у рамках тільки цього кодеку. У той же час, поширеність у світі MPEG 4/Part 10 суттєво програє H.264/AVC, якій не містить у собі жодних рішень щодо забезпечення управління бітовою швидкістю відеопотоку на рівні джерела.

2) Механізм EBCOT (Embedded Block Coding With Optimal Truncation) [10].

Даний підхід існує як відкрите архітектурне рішення у рамках технології JPEG2000 та передбачає можливість регулювання бітової швидкості $V(t)$ відео потоку за рахунок маніпуляції кількістю структурних одиниць (з їх усіченням) кадру для його семантичного опису. Одночасно з

тим, механізм являє собою, скоріше, певний набір рекомендацій, при цьому ряд алгоритмів мають загальний опис та орієнтовані на майбутню реалізацію.

У базовому випадку, концептуально механізм EVCOT вміщує у собі такі кроки обробки:

- оцінка необхідної кількості біт для опису кадру (зображення);
- розрахунок приблизної величини, на яку необхідно знизити початковий об'єм кадру, щоб забезпечити вимоги щодо поточної бітової швидкості;
- обчислення необхідної кількості бітових площин молодших розрядів, які необхідно виключити з розгляду для забезпечення необхідної кількості біт для опису кадру;
- виключення визначеної кількості бітових площин.

Перевагою даного механізму, у порівнянні з SRC є робота на рівні окремих фрагментів кадру, що потенційно може забезпечити раціональний розподіл бітової швидкості у межах кадру.

Проте, як і у випадку SBR, не отримали реалізації на рівні стандарту механізми оцінки помилки, яка буде вноситися шляхом маніпулювання структурними одиницями кадру.

Окрім того, не існує чіткого механізму забезпечення відповідності ступеню стиснення та рівня помилки, що вноситься під час обробки.

У цьому випадку не ураховується те, що бітова швидкість у рамках як цілого кадру, так і окремих фрагментів, розподіляється нерівномірно. Також, залежно від особливостей змісту поточного кодованого кадру, розподіл бітової швидкості між окремими бітовими площинами теж буде мати довільний характер. У таких умовах схема розподілу інформативності між бітовими площинами та рівнями внесеної помилки для різної кількості бітових площин не буде такою, як подано у рамках опису механізму EVCOT (Рис. 2.7).

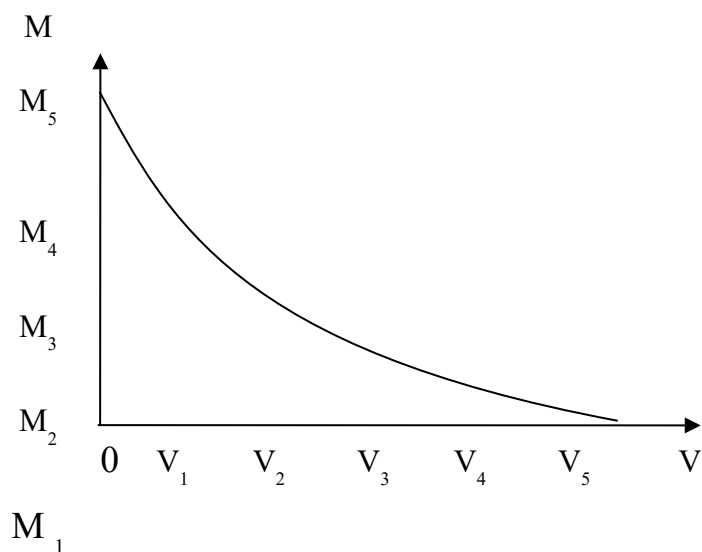


Рисунок 2.7 – Загальний характер залежності рівня внесеної помилки для різної кількості бітових площин у складі трансформанти (для випадку 5 бітових площин)

На рис. 2.7 значення від M_1 до M_5 відповідають рівню помилки, величини від V_1 до V_5 – рівню бітової швидкості.

Очевидно, що чим більша кількість бітових площин буде задіяна для опису трансформанти, тим нижчим буде рівень помилки, відповідно, вищою буде бітова швидкість, і навпаки. Якщо при цьому взяти до уваги, що у реальному випадку залежність між M та V буде мати набагато складніший характер, становиться зрозуміло, що для ефективного функціонування у складі механізму EVCOT необхідно задіяти механізм оцінки “важливості” бітових площин. Такий механізм, на відміну від початкового розподілу бітових площин за показником індексу, буде використовувати розподіл за показником рівня помилки, що може бути внесений за умови виключення з розгляду тієї чи іншої бітової площини.

Отже, методи адаптації бітової швидкості відеопотоку до пропускної здатності мережі можуть бути перспективними з позиції забезпечення стабільної передачі відеоданих та забезпечення потрібних рівнів якості за умови реалізації недостаючих механізмів.

3 КОНЦЕПТУАЛЬНА МОДЕЛЬ МЕТОДУ КОНТРОЛЮ ОБ'ЄМУ МУЛЬТИМЕДІЙНИХ ДАНИХ НА БАЗІ ТРИВИМІРНОЇ МОДЕЛІ ТРАНСФОРМОВАНОГО ФРАГМЕНТУ КАДРУ

3.1 Вимоги до методу контролю об'єму мультимедійних даних

Для того, щоб визначити ключові вимоги до методу контролю, розглянемо спочатку особливості мультимедіа, а саме – відео, як типу даних.

Перш за все – відео у багатьох випадках (інтерактивне відео, відеоконференції, дані відеонагляду) несе інформацію про певні події у реальному часі.

Таким чином, затримки трансляції відео повинні бути не вище певного порогового значення t_{th} [7]. Тут t_{th} - лімітовано незначна величина, внесення якої не вплине на обробку, передавання та відображення відео у реальному часі.

Окрім цього, відеопоток характеризується значною динамікою у часі. Це стосується як об'єму даних V , що надходить у мережу з вихідного буферу передавача за певний часовий проміжок, так і потенційно значного динамічного діапазону у відношенні до зміни об'єму відеоінформації між сусідніми кадрами F або групами GOP кадрів.

Звідси випливає, що втрата чи спотворення певних кадрів та їхніх серій може вести до зменшення інформативності відео повідомлення, як наслідок падіння достовірності прийнятих даних.

Тобто, під час трансляції відео необхідно забезпечити певний рівень достовірності, інакше кажучи – забезпечити показник якості $M(F)$ прийнятого кадру на рівні не нижчому, ніж установлений за умовами надання сервісу.

На додачу до зазначених особливостей слід сказати, що у порівнянні з іншими типами даних мультимедіа відрізняється значними об'ємами передаваних даних.

Отже, щоб забезпечити ефективне управління об'ємом відеоданих, що надходять у мережу, метод контролю повинен бути раціональним, тобто, відрізнятися простотою реалізації, не потребувати значних обчислювальних потужностей та мати певний запас ресурсів.

Тоді, у загальному вигляді, метод управління повинен відповідати наступним вимогам:

$$\begin{cases} t_{obr} \leq t_{th}; \\ M \leq M_{st}; \\ V(t) \leq V(t)_{suit}, \end{cases} \quad (3.1)$$

де t_{obr} - сукупний час обробки відеокадру та/або його фрагменту;

$V(t)_{suit}$ - прийнятний (suitable) об'єм даних, що передаються у одиницю часу, або, іншими словами, бітова швидкість відеопотоку. Тут величина $V(t)_{suit}$ еквівалентна пропускній здатності P мережі;

$V(t)$ - об'єм даних, що передаються у одиницю часу, або бітова швидкість відеопотоку;

M_{st} - показник якості кадру або його окремого фрагменту, який задано умовами надання сервісу (service terms);

M - поточний показник якості прийнятого кадру.

Показником якості прийнятого відеокадру та/або його фрагменту може слугувати одна з метрик оцінки цілісності даних.

У нашому випадку використовується показник середньоквадратичного відхилення (СКВ), Тоді вираз (1.8) у нашому випадку буде переписано наступним чином:

$$M = \frac{1}{AHW} \sum_{\alpha=1}^A \sum_{x=1, y=1}^{H, W} (I_{\alpha}(x, y) \times I'_{\alpha}(x, y))^2, \quad (3.2)$$

де A – кількість кадрів у оцінюваному відрізку відеопотоку.

Інакше кажучи, метод контролю об'єму відеоданих повинен забезпечувати кількості біт для представлення кадру на рівні, що не перевищує прийнятний, з одночасним підтримуванням потрібного показника середньоквадратичної помилки за час, що не перевищує заданий.

Вираз (3.1) визначає умови ефективної передачі відеотрафіку – тобто, такі, що забезпечують стабільну трансляцію у часі з рівнем якості, не нижче потрібного.

Проте у нашому випадку, згідно технічного завдання, необхідно забезпечити передачу відеоданих за таких умов:

- створення мінімального навантаження на мережу;
- забезпечення показника помилки на мінімальному рівні, аж до теоретичної межі;
- забезпечення часу обробки на рівні, що не перевищує пороговий.

Дані умови можемо вважати умовами оптимальної передачі. Тоді вимоги (3.1) приймуть наступний вигляд:

$$\begin{cases} t_{obr} \leq t_{th}; \\ M \rightarrow \min; \\ V(t) < P \rightarrow \min. \end{cases} \quad (3.3)$$

Дотримання таких умов обробки відеотрафіку, на відміну від традиційних технологій кодування та підходів, орієнтованих, у першу чергу, на підвищення оперативності доставки відеоданих, що показано наступним виразом:

$$\begin{cases} t_{\text{obr}} \rightarrow \min; \\ M \leq M_{\text{st}}; \\ V(t) \leq P, \end{cases} \quad (3.4)$$

забезпечує уникнення ситуацій, коли на опис високоінформативних фрагментів кадру буде відводитись недостатня кількість біт, у той час, коли для опису низько інформативних фрагментів буде відведено надмірну кількість біт.

Очевидно, що для контролю об'єму надходячих у мережу даних мультимедіа, необхідно здійснювати маніпуляцію тими чи іншими параметрами відеопотоку.

Тому, щоб визначити, які саме параметри відеопотоку можна оперативнo змінювати для забезпечення передачі відеоданих у відповідності до вимог (3.3), проаналізуємо спочатку структурний склад одного кадру, як елементу відеопотока.

3.2 Представлення фрагментів кадру у вигляді тривимірних об'єктів

Розглянемо початковий кадр F сукупністю окремих фрагментів – а саме – блоків b_i розміром 8×8 пікселів кожен.

Дані блоки є результатом рівномірного поділу початкового зображення у ході внутрішньокадрової обробки за базовим алгоритмом JPEG.

Тоді кадр може бути описаний наступним чином [8]:

$$F = \prod_{i=1}^I b_i, \quad (3.3)$$

де b_i - i -й блок кадру.

Блоки b_i відеокадру, після перетворення колірної моделі RGB у YCbCr, підлягають дискретно-косинусному перетворенню (ДКП).

У результаті утворюється множина G трансформант:

$$G = \prod_{i=1}^I \Omega_i, \quad (3.4)$$

де Ω_i - i -а трансформанта кадру F .

Отже, на даному етапі обробки об'єм даних V , що приходиться на один кадр, визначається наступним виразом:

$$V = \sum_{j=1}^{\theta_{fr}} \eta_j, \quad (3.5)$$

де η_j - об'єм даних, що відповідає одному j -му фрагменту;

θ_{fr} - кількість фрагментів кадру.

У свою чергу, для опису кожної трансформанти задіяна сукупністю компонент $\Omega(i)_{hw}$:

$$\Omega_i = \|\Omega(i)_{hw}\|, \quad (3.6)$$

де $\Omega(i)_{hw}$ - $(h; w)$ -а компонента i -ї трансформанти.

Компонента $\Omega(i)_{hw}$ трансформанти Ω_i являє собою послідовність біт, що описує представлення відповідного елемента у двійковому вигляді (рис. 3.1). Це еквівалентно перетворенню:

$$\|\Omega(i)_{hw}\| \rightarrow \left\| \left\langle \omega(i)_{hw}^{(\theta_{pl})}, \omega(i)_{hw}^{(\theta_{pl}-1)}, \dots, \omega(i)_{hw}^{(k)}, \dots, \omega(i)_{hw}^{(0)} \right\rangle^T \right\|, \quad (3.7)$$

$$\omega(i)_{hw}^{(k)} \in \{0, 1\}, \quad h = \overline{0, 7}; \quad w = \overline{0, 7}; \quad k = \overline{K, 0},$$

де $\omega(i)_{hw}^{(k)}$ – k -й біт двійкового представлення $(h; w)$ -ї компоненти p -ї трансформанти;

K – кількість двійкових розрядів для представлення трансформанти.

Сукупність усіх елементів k -го прошарку i -ї трансформанти має назву *бітова площа* $\omega(i)_k$ [8].

Тоді за допомогою сукупності двійкових представлень усіх елементів матриці Ω_i можна здійснити її опис у тривимірному просторі, як показано на рис.3.1.

При цьому верхній шар даного куба утворюють старші біти двійкового представлення. Відповідно, нижній – відповідає молодшим бітам. У загальному випадку, якщо не приймати до уваги особливості змісту фрагментів кадру, можна вважати нижні шари менш інформативними. Проте, частіше за все така залежність має випадковий характер, тобто, у ряді випадків інформативними можуть бути молодші біти.

Таке представлення трансформанти у тривимірному просторі дозволяє організувати пошарову передачу даних, основою якої може слугувати *метод послідовного наближення* (successive approximation) технології Progressive JPEG [11].

У цьому випадку забезпечується можливість управління об'ємом відеоданих, який надходить у мережу кожному одиницю часу.

Сутність управління полягає в тому, що при передачі можна відправляти в канал тільки частину бітових площин, тим самим змінюючи (у тому числі і зменшуючи) об'єм кадру [12].

Тоді формулу (3.5) може бути переписано у наступному трактуванні:

$$V = \sum_{j=1}^{\theta_{fr}} \eta_j = \sum_{j=1}^{\theta_{fr}} \sum_{\psi=1}^{\theta_{pl}} \eta_{j,\psi}, \quad (3.8)$$

де $\eta_{j,\psi}$ - об'єм даних Ψ -ї бітової площини $\omega(i)_k$, j -ї трансформанти кадру F ;

θ_{fr} - кількість трансформант Ω_i кадру;

θ_{pl} - кількість бітових площин j -ї трансформанти кадру.

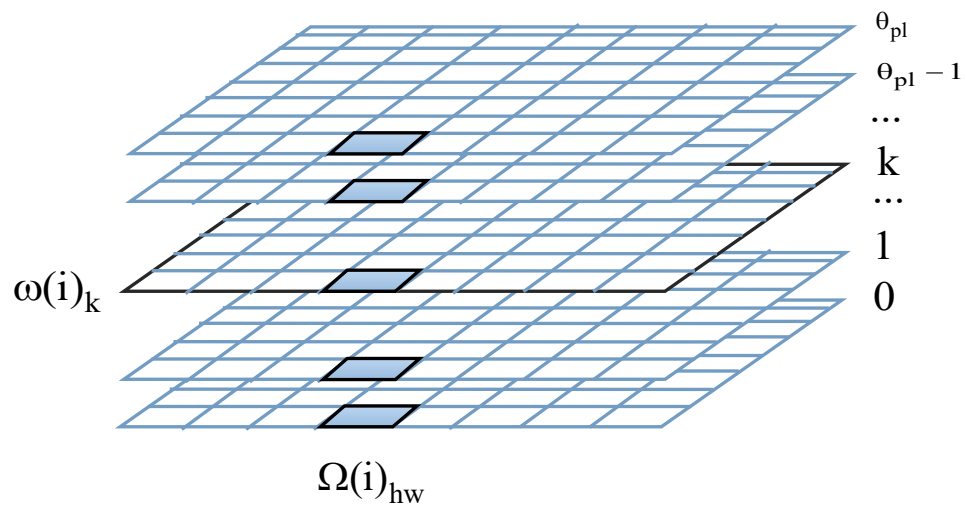


Рисунок 3.1 - Представлення трансформанти сукупністю бітових площин.

У свою чергу, для окремої трансформанти:

$$V(\Omega_i) = \sum_{k=1}^{\theta_{pl}} \omega(i)_k. \quad (3.9)$$

3.3 Реалізація алгоритму на базі метода динамічного програмування

Розглянемо підхід до організації контролю бітовою швидкістю відео потоку на рівні управління бітовою швидкістю $V(t)$ окремого кадру F . Припустимо, що вихідний відеокадр представлений m блоками.

Контроль бітової швидкості у цьому випадку можна інтерпретувати як m - кроковий процес прийняття рішень, що дає можливість реалізувати алгоритм методом динамічного програмування [13].

При цьому, тут відбувається заміна одному складного завдання з багатьма змінними - рядом завдань із меншим числом змінних. Також, у процесі обчислення, існує можливість відкидати шляхи розрахунків, на яких проміжні значення цільової функції виходять за рамки припустимих за умовами завдання.

Компоненти кожної з θ_{fr} трансформант представлені θ_{pl} бітовими площинами. При цьому, існує деяке обмеження відносно пропускну здатності P каналу.

Процес контролю розкладається на θ_{fr} етапів відповідно до кількості трансформант. при цьому, оптимальне управління будується поступово.

У процесі контролю відбувається вибір оптимальної стратегії U^* .

Далі під *оптимальною стратегією* U^* будемо розуміти сукупність бітових площин для кожної із трансформант, з урахуванням їх кількості θ й позиції P у трансформанті, яка задовольняє умовам (3.3), тобто:

$$U^* = f(p, v). \quad (3.10)$$

При цьому, оптимальним управлінням буде зватися процес пошуку сукупність бітових площин, здатних задовольнити вимоги (3.3).

Оптимальна стратегія для кадра має аддитивний характер, а саме:

$$U^* = \sum_{k=1}^{\theta_{fr}} U_k^* \quad (3.11)$$

де U_k^* - оптимальна стратегія на k -му етапі.

Аддитивний характер стратегії дає змогу формувати оптимальне рішення для кадру у вигляді сукупності рішень для окремих трансформант. Тоді принцип динамічного програмування у застосуванні до нашого випадку спочатку розглянемо для однієї трансформанти.

При виборі стратегії відбувається перебір комбінацій бітових швидкостей та СКВ для всіх θ_{fr} трансформант залежно від різної кількості θ_{pl} переданих бітових площин.

На кожному k -му етапі стратегія вибирається таким чином, щоб разом з оптимальними стратегіями на наступних кроках (починаючи з $(k + 1)$ -го), це приводило б до загальної мінімізації цільової функції $V(t)$ й діючих обмежень M відповідно до умов (3.3) [14].

Іншими словами, необхідно знайти аргумент функції $V(t) = f(U^*)$, що мінімізує значення бітової швидкості та СКВ.

Це еквівалентно виразу:

$$V(t)_k = f(V(t)_{k-1}, U_k) = \min V(t)_k, \quad (3.12)$$

де $V(t)_{k-1}$ - значення цільової функції на попередньому етапі обчислень.

При подальшому визначенні оптимальної стратегії $U^* = \{U_k^*, \dots, U_9^*\}$ відповідно до (3.12) на $\Omega_i - k + 1$ кроках, що залишилися, обчислюється величина $V(t)^* = \min V(t)$, яка залежить тільки від попереднього значення $V(F)_{k-1}$:

$$V(t)^*(V(t)_{k-1}^*) = \min_{U_k} f(V(t)_{k-1}, U_k) = f(V(t)_{k-1}, U_k^*), \quad (3.13)$$

де $V(t)^*(V(t)_{k-1}^*)$ - умовний мінімум, що є наслідком застосування оптимальної стратегії на $k - 1$ -му кроці.

Це відповідає вибору комбінації кількості бітових площин, що дають мінімальну сумарну бітову швидкість при мінімальному СКВ на даному етапі.

Відповідно до принципу оптимальності Белмана (рис.3.2), на кожному k -му кроці управління U_k вибирається таким чином, щоб у комбінації з

обраними управліннями на наступних $\Omega_i - 1 + k$ кроках, починаючи з k -го, забезпечувалася мінімізація сумарної величини цільової функції функції (показника ефективності). Даний процес у аналітичній формі може бути описаний наступним виразом:

$$V_F^*(V_{p-1}) = \min_{U_p} \{f_p(V_{p-1}, U_p)\},$$

$$V(t)^*(V(t)_{k-1}) = \min_{U_k} \{f_k(V(t)_{k-1}, U_k) + V(t)_{k+1}^*(V(t)_k)\}, \quad (3.14)$$

де U_k - управління на k -му кроці;

V_{p-1} - бітова швидкість, сформована $(p-1)$ трансформантами;

$V(t)_k$ - значення цільової функції на k -му кроці;

f_k - показник ефективності на k -му кроці, залежний від значень цільової функції на попередньому кроці й обраного управління U_k на даному кроці.

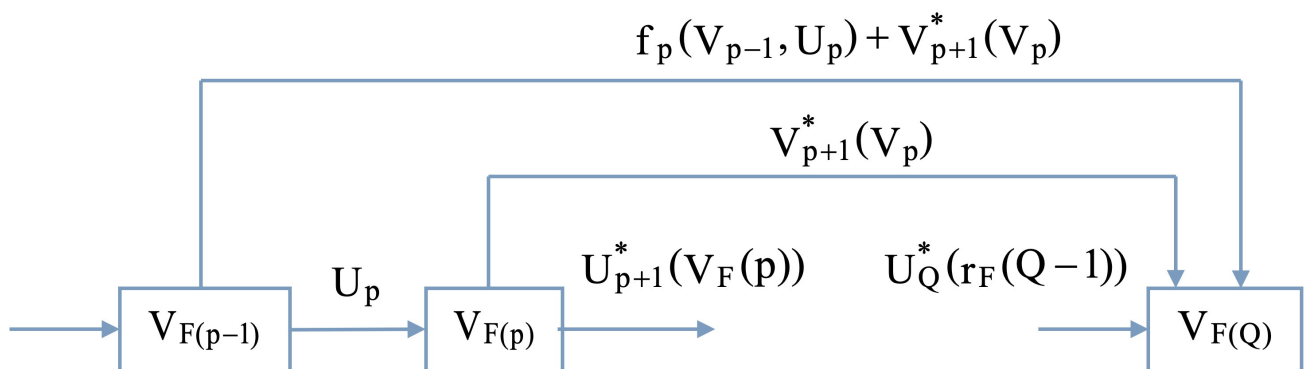


Рисунок 3.2 – Застосування принципу оптимальності Белмана у процесі управління бітовою інтенсивністю відеопотоку

На рис. 3.3 наведено приклад обробки трьох трансформант.

Після обчислення множин $V(t)$ да відповідних ним величин похибки, оптимальне управління визначається у ході наступних кроків:

1) виключення точок, що відповідають управлінням U , що приводять до переповнення буфера;

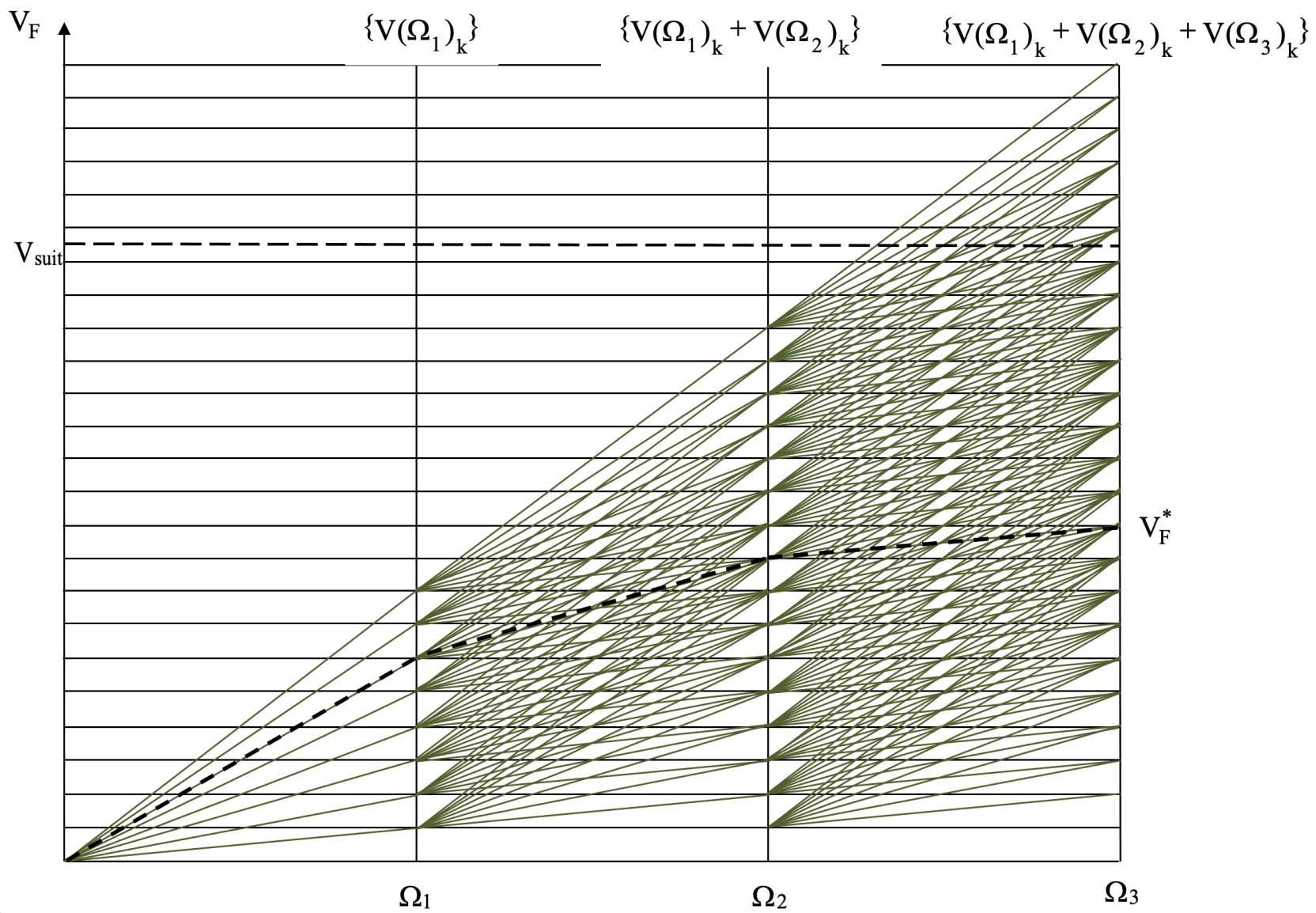


Рисунок 3.3 – Приклад обробки трьох трансформант за методом динамічного програмування

2) виключення точок з однаковими значеннями M_{\min} , за винятком таких, для яких $V(t) = V(t)_{\min}$;

3) виключення точок з однаковими значеннями $V(t)$, крім таких, для яких $M = M_{\min}$.

Швидкість роботи реалізованого у такий спосіб механізму вибору управління U залежить від розмірів оброблюваних кадрів, а також кількості розрядів θ_{pl} для опису кожного з них.

Крім розміру кадрів і їх розрядності, загальна величина затримки передачі визначається також (хоч і у меншій мірі) використовуваним методом відображення кадру [5].

У технології Progressive Jpeg використовується метод відображення "на льоту" а також метод відображення після приймання всього кадру.

Метод відображення "на льоту", передбачає послідовне виведення кадру на пристрій відображення, у міру приймання бітових площин утворюючих його трансформант. На рис. 3.4 показаний алгоритм відображення кадру "на льоту".

Для зниження обчислювального навантаження на декодувальне обладнання та зниження часу затримки, доцільно використовувати метод відображення всього кадру.

3.4 Оцінка складності алгоритму на базі динамічного програмування

Виконаємо оцінку запропонованого методу з позицій можливості реалізації у межах існуючих апаратних можливостей обчислювальних систем.

Теорема 3.1. Кількість варіантів вибору бітових площин для кадру відеопотоку визначається наступним виразом:

$$W = (2^K)^G, \quad (3.15)$$

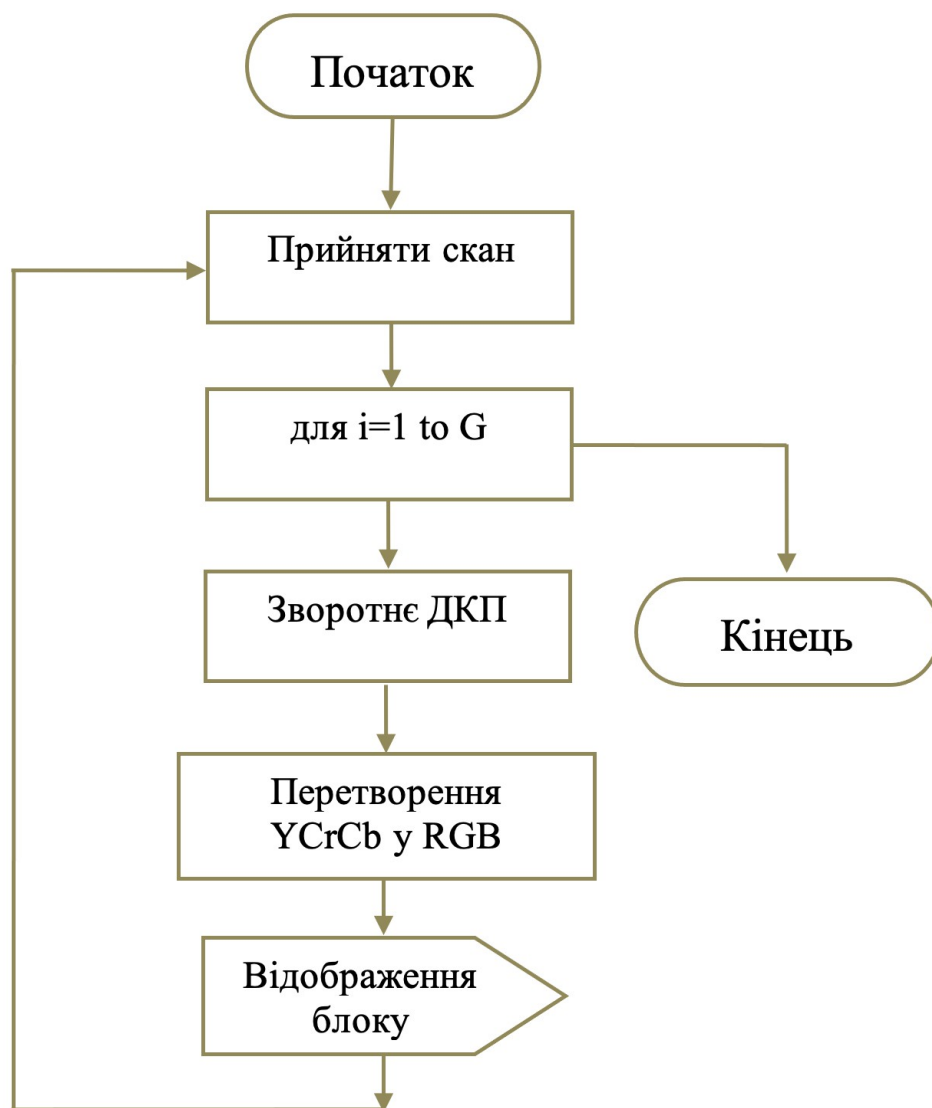


Рисунок 3.4 - Алгоритм декодування зображення “на льоту”

де K - кількість бітових площин для опису компонент однієї трансформанти;

G - кількість трансформант у кадрі.

Доказ. Нехай для опису компонент трансформант використовується K розрядів, тобто, кожна компонента представлена K бітовими площинами.

У кожній з K позицій, відповідної до кожної площини, рішення про включення даної площини у передачу може приймати 2 значення - "1" - якщо площина передається та "0" - у випадку, якщо бітова площина ігнорується.

Таким чином, маємо розміщення з повтореннями з G елементів по K - A_G^K [15].

У K -розміщенні $\{a_1, a_2 \dots a_K\}$ елемент a_1 в λ -елементній множині W можна вибрати λ способами, елемент a_2 - теж λ способами, нарешті, елемент a_K - λ способами. За правилом добутку маємо: $A_\lambda^K = \lambda^K$.

Отже, число варіантів вибору бітових площин для трансформанти дорівнює 2^K . Далі, за правилом добутку, аналогічно, для двох трансформант маємо $W = (2^K)^2$, для G трансформант, відповідно, $W = (2^K)^G$.

Твердження доведене.

Ріст величин G і K спричиняє зростання значення W та, відповідно, часу обробки відеоданих на передаючій стороні.

Число станів (варіантів прийняття рішення у ході управління) алгоритму при обробці кадрів форматів Full HD і Hd-ready, представлених 8 розрядами, представлено у таблиці 3.1.

Таблиця 3.1 – Число станів W алгоритму управління для кадрів формату Full HD та HD-ready.

Формат	Розмір р кадру	Кількість бітових площин							
		1	2	3	4	5	6	7	8
HD-ready	1280× 720	1	2^{14400}	4^{14400}	8^{14400}	16^{14400}	32^{14400}	64^{14400}	128^{14400}
Full HD	1920× 1080	1	2^{32400}	4^{32400}	8^{32400}	16^{32400}	32^{32400}	64^{32400}	128^{32400}

Аналіз даних таблиці 3.1 показує, що при таких значеннях числа станів алгоритму, розглядувана модель обробки не може бути реалізована в існуючому базисі потужностей обчислювальних пристроїв.

Відповідно, необхідно суттєво знизити обсяг станів алгоритму. У зв'язку із цим, пропонуються 2 підходи для зменшення кількості варіантів вибору управління.

3.5 Зменшення кількості варіантів вибору управління на базі звуження простору розрахунків та двокаскадного виявлення подібних фрагментів.

У рамках даного підходу, на першому етапі обробки виконується поділ вихідного кадру на ряд підкадрів F_{sub} .

Найшвидший варіант реалізації такого розділення дозволяє розподілити кадр на підкадри у напрямку рядків, як показано рис. 3.5.

Це дає змогу зменшити кількість обчислень за рахунок зменшення величини G виразу (3.15).

F_{sub}

$\Omega_{1,1}$	$\Omega_{1,2}$	$\Omega_{1,3}$	$\Omega_{1,4}$...	$\Omega_{1,\eta-3}$	$\Omega_{1,\eta-2}$	$\Omega_{1,\eta-1}$	$\Omega_{1,\eta}$
$\Omega_{2,1}$	$\Omega_{2,2}$	$\Omega_{2,3}$	$\Omega_{2,4}$...	$\Omega_{2,\eta-3}$	$\Omega_{2,\eta-2}$	$\Omega_{2,\eta-1}$	$\Omega_{2,\eta}$
...								
$\Omega_{\psi,1}$	$\Omega_{\psi,2}$	$\Omega_{\psi,3}$	$\Omega_{\psi,4}$...	$\Omega_{\psi,\eta-3}$	$\Omega_{\psi,\eta-2}$	$\Omega_{\psi,\eta-1}$	$\Omega_{\psi,\eta}$
$\Omega_{\omega,1}$	$\Omega_{\omega,2}$	$\Omega_{\omega,3}$	$\Omega_{\omega,4}$...	$\Omega_{\omega,\eta-3}$	$\Omega_{\omega,\eta-2}$	$\Omega_{\omega,\eta-1}$	$\Omega_{\omega,\eta}$

Рисунок 3.5 – Виділення підкадрів F_{sub} у структурі початкового кадру у напрямку рядків

Для кадру формату SD, у загальному випадку, таких маніпуляцій щодо зменшення об'єму обчислень може бути достатньо.

Проте кадри формату HD, та особливо UHD, потребують додаткових дій для зменшення величини W .

У такому разі, далі буде задіяно двокаскадне виявлення подібних фрагментів підкадру. Для цього виконаємо сегментацію підкадру F_{sub} за однією або декількома якісними ознаками.

Це можуть бути:

- ознаки структурної подібності S_{str} фрагментів підкадру F_{sub} між собою на рівні трансформант, або початкових фрагментів;
- ознака подібності S_{dif} за показником складності фрагментів.

Оцінка структурної подібності може бути виконана на базі порівняння максимальної $\ell(k, \Omega_i)_{\text{max}}^{(\text{bs})}$ та мінімальної $\ell(k, \Omega_i)_{\text{min}}^{(\text{bs})}$ довжин серій двійкових елементів, виявлених у межах k -х бітових площин Ω_i та Ω_{i+1} трансформант кадру. Це еквівалентно виразу:

$$S_{\text{str}} = (|\Delta\ell_{\text{max}}^{(\text{bs})}|; |\Delta\ell_{\text{min}}^{(\text{bs})}|), \quad (3.16)$$

де $|\Delta\ell_{\text{max}}^{(\text{bs})}| = |\ell(k, \Omega_i)_{\text{max}}^{(\text{bs})} - \ell(k, \Omega_{i+1})_{\text{max}}^{(\text{bs})}|$, відповідно у свою чергу, $|\Delta\ell_{\text{min}}^{(\text{bs})}| = |\ell(k, \Omega_i)_{\text{min}}^{(\text{bs})} - \ell(k, \Omega_{i+1})_{\text{min}}^{(\text{bs})}|$.

Якщо при цьому величини пари елементів $|\Delta\ell_{\text{max}}^{(\text{bs})}|$ та $|\Delta\ell_{\text{min}}^{(\text{bs})}|$ не відхиляються між собою більше, ніж на певну величину $\Delta\ell$, робиться висновок, що фрагменти Ω_i та Ω_{i+1} можуть бути подібні [8]. У такому разі, управління U , задіяне для обробки трансформанти Ω_i буде застосовуватися і для трансформанти Ω_{i+1} . Тобто, у такий спосіб може бути зменшено величину G виразу 3.15.

Розглянемо випадок, якщо за ознакою структурної подібності S_{str} фрагментів підкадру, до яких може бути застосоване єдине управління U не

виявлено, тобто $|\Delta t_{\max}^{(bs)}|$ та $|\Delta t_{\min}^{(bs)}|$, або один з цих показників суттєво відрізняються від Δt . У такому випадку додатково може бути задіяно аналіз складності S_{dif} фрагментів.

Такий процес ілюструється наступним виразом:

$$S_{\text{dif}} = \sum_{\alpha=1}^h \sum_{\beta=1}^w (\Omega(i)_{\alpha}^{(\max)} - \Omega(i)_{\alpha}^{(\min)})_{\beta} \quad (3.17)$$

де $\Omega(i)_{\alpha}^{(\max)}$ та $\Omega(i)_{\alpha}^{(\min)}$ - відповідно, максимальне та мінімальне значення компонент α -го рядку i -го блоку кадру.

Відповідно до виразу (3.17), у межах кадру визначається фрагмент, що має максимальний показник $S_{\text{dif}}^{(\max)}$. Далі уся множина фрагментів кадру розподіляються на 3 групи, а саме:

- фрагменти, що мають показник від 100% до 70% від $S_{\text{dif}}^{(\max)}$,

вважаються складними;

- від 69% до 30% $S_{\text{dif}}^{(\max)}$ - середньо-складними;

- 29% і нижче – незначно-складними.

Сенс такого розподілу полягає у наступному.

Фрагменти, які відносяться до складних, залишаються незмінними. Натомість середньо-складні фрагменти підлягають процедурі усічення бітових площин, унаслідок якої вихідний фрагмент буде представлено 3 бітовими площинами. У свою чергу, незначно-складні фрагменти буде виключено з розгляду.

Таким чином, величину W буде зменшено.

Як показує аналіз рис. 3.4. та виразів (3.15)-(3.17), у найгіршому випадку (якщо усі фрагменти визнано складними, та подібності між ними за показником S_{str} не виявлено) розглянутий підхід до зменшення об'єму обчислень дозволяє:

- за рахунок розбиття кадру на підкадри розміром рядка початкового кадру зменшити кількість варіантів W вибору управління з 128^{32400} до 128^{240} , тобто, у 128^{32160} разів, та, відповідно, зменшити час обробки;
- у випадку, коли підкадр має розмір $1/2$ рядка кадру, розмір величини W зменшується до 128^{120} , або у 128^{32280} разів. При цьому загальний процес обробки кадру замінюється кількома паралельними процесами, кожен з яких потребує значно менше обчислювальних ресурсів.

У той же час, для випадку, коли $1/3$ фрагментів визнана незначно-складною, $1/3$ – середньо-складною, між $1/2$ усіх фрагментів середньо-складних та складних є структурна подібними маємо:

- для підкадру, що має розмір рядку кадру FullHD, величина W буде оцінюватися за таким виразом:

$$W = (2^3)^{G_{md}} + (2^K)^{G_d}, \quad (3.18)$$

де G_{md} та G_d - відповідно, кількість середньо-складних та складних фрагментів.

$$\text{Тоді для даного випадку } W = (2^3)^{80} + (2^8)^{80} = 2^{240} + 2^{640};$$

Для підкадру, що має розмір половини рядку кадру FullHD, кількість станів W алгоритму, згідно з виразом (3.18), буде визначатися наступним чином: $W = (2^3)^{40} + (2^8)^{40} = 2^{120} + 2^{320}$.

Розглянемо далі розподіл початкового кадру F на підкадри F_{sub} розміром b фрагментів кожен.

Тоді кількість варіантів W вибору управління по одному підкадру у найгіршому випадку буде рівнятися 2^{48} .

Відповідно, застосування схеми двокаскадного виявлення подібних фрагментів дасть змогу зменшити величину W до $W = (2^2)^2 + (2^8)^2 = 2^4 + 2^{16}$, що є співрозмірним з наявними обчислювальними можливостями сучасних обчислювальних систем.

4 ОРГАНІЗАЦІЯ СЛУЖБОВИХ ДАНИХ ДЛЯ МЕТОДУ УПРАВЛІННЯ БІТОВОЮ ШВИДКІСТЮ

4.1 Вимоги до необхідної службової інформації

Наступним кроком після визначення бітових площин, що будуть використані для опису трансформант кадру під час передачі, необхідно визначити, як саме трансформанти будуть відновлюватися на прийомній стороні.

Іншими словами, необхідно визначити систему службових даних, на базі якої буде формуватися передаваний відеопоток.

При цьому, до службових даних можна висунути наступні вимоги:

- однозначне визначення бітових площин як значущих, так і відсічених;
- простота та тривіальність у реалізації, без помітного ускладнення загального алгоритму управління
- мінімальний об'єм.

4.2 Формування службових даних

Для того, щоб мати змогу визначити бітові площини, які необхідно передати, для кожної i -ї трансформанти задається вектор передачі $T(i)_k$ бітових площин трансформанти, який відповідає двійковому представленню чисел від K до 0 :

$$T(i)_k = \langle t(i)^K, \dots, t(i)^k, \dots, v(t)^0 \rangle \quad (4.1)$$

Кожному варіанту переданих площин буде відповідати трансформанта Ω'_i , що відповідає початковій Ω_i , але відрізняється ступінню спотворення γ_i відносно вихідної трансформанти.

Тоді кожній трансформанті Ω_i у загальному випадку буде відповідати вектор передачі $T(i)_k$ бітових площин, де кожен елемент $t(i)^k$ приймає значення 0 або 1.

Відповідно, 1 у k -му розряді вектору відповідає бітовій площині з індексом k , яка буде передаватися без змін.

У свою чергу, 0 визначає бітову площину, яка на передаючій стороні буде відсічена, тобто, виключена з розгляду, тим самим зменшуючи об'єм даних, що будуть передаватися.

Тоді на прийомному боці така площина буде формуватися з нулів.

Таким чином, після того, як вектор передачі $T(i)_k$ бітових площин $\omega(i)_{hw}^{(k)}$ заданий, можна формувати послідовність бітових площин для i -ї трансформанти, які будуть передаватися, при заданому векторі $T(i)_k$.

Процес побудови передаваної послідовності $R(i)_k$ бітових площин трансформанти Ω_i на передаючій стороні відбувається у відповідності до наступної формули:

$$S(i) = t(i)^K \& \dots t(i)^k \& \dots t(i)^0 \& (\omega(i)_{hw}^{(k)} | \omega(i)_{hw}^{(k)} = 1) \quad (4.2)$$

Іншими словами, якщо k -й біт вектора передачі рівний 1, то передаються усі 64 елемента k -ї площини $\omega(i)_{hw}^{(k)}$.

Якщо відповідний біт вектора передачі $T(i)_k$ трансформанти Ω_i рівний 0, то значення біт відповідної бітової площини не передаються, у наслідок чого зменшується об'єм даних, що надсилаються.

Отже, такий підхід дає змогу однозначно визначити, які бітові площини будуть передаватися (будуть значущими).

Алгоритм формування при заданому векторі $T(i)_k$ складається з таких кроків:

- 1) обнуління поточного значення $R(i)_k$;
- 2) обчислення вектору передачі $T(i)_k$;
- 3) вибір бітових площин у відповідності до вектору $T(i)_k$;
- 4) об'єднання вектору передачі $T(i)_k$ з вибраними бітовими площинами:

$$R(i)_k = T(i)_k \& \{\omega(i)_k\}. \quad (4.3)$$

Відповідний процес формування послідовності $R(i)_k$ для передачі i -ї трансформанти показано рис. 4.1.

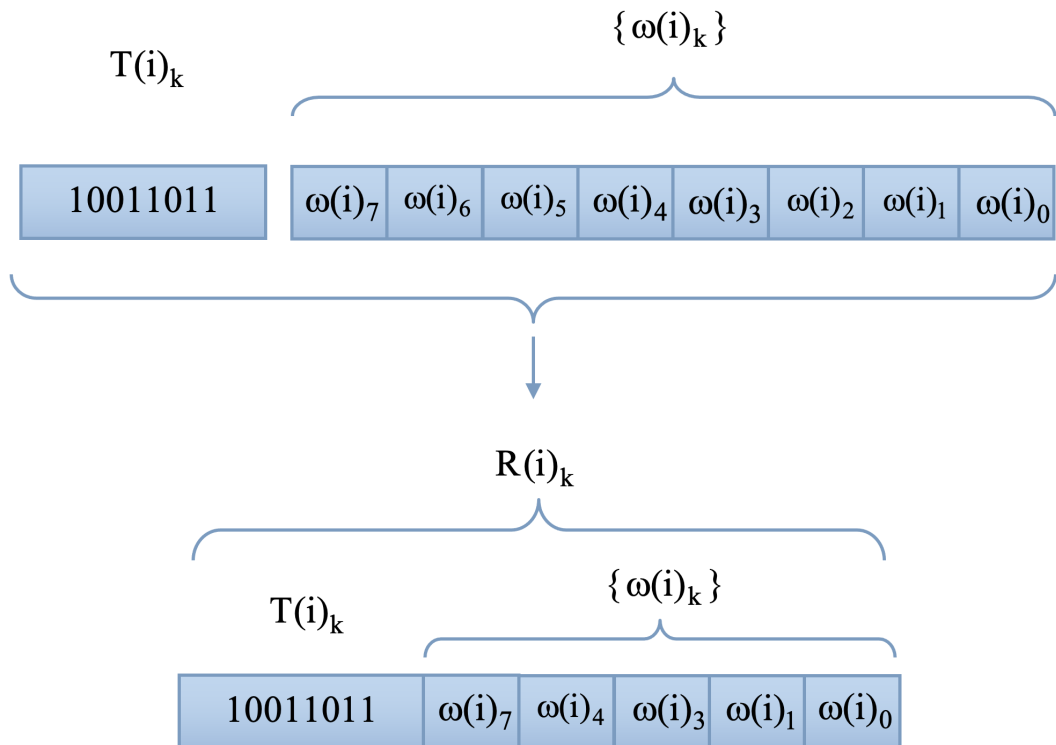


Рисунок 4.1 – Процес формування послідовності $R(i)_k$ передачі бітових площин трансформанти Ω_i .

Як бачимо з рис. 4.1, після того, як визначено одиничні елементи вектору $T(i)_k$, відбувається об'єднання з ним множини $\{\omega(i)_k\}$ бітових

площин. При цьому бітові площини, індекси яких відповідають нульовим елементам вектору $T(i)_k$, не буде включено до послідовності $R(i)_k$. Тобто, буде забезпечено зниження об'єму відеоданих.

У свою чергу, операція відновлення трансформанти Ω_i на прийомній стороні з послідовності $R(i)_k$ складається з наступних етапів:

- 1) Зчитати $T(i)_k$ (перші 8 біт послідовності);
- 2) Заповнити елементами бітові площини, які відповідають ненульовим бітам у векторі передачі: якщо біт вектора передачі рівний 1, то значення наступних 64 біт заповнюють чарунки відповідної бітової площини; якщо елемент рівний 0, то площина формується нульовими елементами (рис. 4.2).

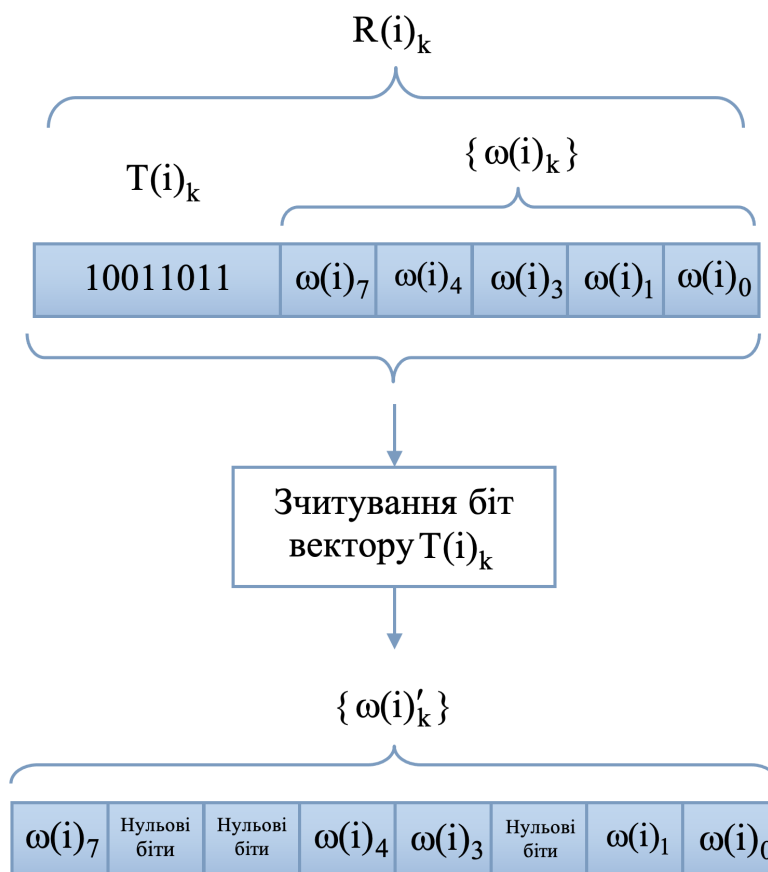


Рисунок 4.2 – Процес відновлення трансформанти Ω_i на прийомному боці з послідовності $R(i)_k$ передачі бітових площин.

4.3 Оцінка об'єму службових даних

Для визначення кількості одиниць у векторі передачі використаємо величину $K1(T(i)_k)$.

За допомогою її визначимо розмір передаваної послідовності та, відповідно, бітову швидкість $V(T(i)_k)$ і її трансформанти при векторі $T(i)_k$:

$$V(T(i)_k) = 8 + 64 \times K1(V(i)_k). \quad (4.4)$$

Таким чином, при 8-розрядному представленні трансформанти вона може містити у собі від 72 до 520 біт (від 9 до 65 байт). Зазначимо, що такий об'єм трансформанти буде за умови, що на даному етапі на застосовано ніяких технологій кодування її двійкового представлення.

Тобто, у найгіршому випадку, якщо усі елементи вектору $T(i)_k$ будуть одиницями, об'єм службової інформації буде становити 1/64 від кількості біт на представлення трансформанти.

З іншого боку, якщо трансформанта не буде передаватися взагалі ($T(i)_k = \langle 0,0,0,0,0,0,0,0 \rangle$), її об'єм становитиме 8 біт.

Окрім того, у випадку обробки певної множини трансформант з однаковими параметрами кількості бітових площин та їх розміщення, один і той же вектор може бути задіяно до ряду трансформант, що також зменшить обсяг службових даних.

Отже, об'єм службових даних у випадку використання вектору передачі не створює суттєвого впливу на загальний об'єм передаваних даних, тобто, усі вимоги щодо службових даних виконано.

5 МЕХАНІЗМ КОНТРОЛЮ ШВИДКІСТЬ ПЕРЕДАЧІ

В процесі обробки кадру відбувається його розбиття на блоки розмірністю $m \times n$. Таким чином, весь кадр являє собою безліч блоків $\{b_1, b_2, \dots, b_k\}$. Ці блоки кодуються окремо один від одного [16]. Позначимо через $d(t)_i$ та $\sigma(t)_i$ бітові витрати і середню квадратичну помилку для блоку b_i поточного кадру t . Значення $d(t)_i$ та $\sigma(t)_i$ залежать від вектора рішень Ψ^i , який використовується при кодуванні кожного блоку. Вектор рішення входить до складу безлічі рішень, яке позначимо літерою Ψ . Вектор рішень включає в себе k компонент, тобто $\Psi^i = \{\Psi_1^i, \Psi_2^i, \dots, \Psi_k^i\}$. Таким чином, вираз для $d(t)_i$ та $\sigma(t)_i$ можна представити в наступному вигляді:

$$d(t)_i = d_i(t, \Psi^i) = d_i(t, \psi_1^i, \psi_2^i, \dots, \psi_k^i);$$
$$\sigma(t)_i = \sigma_i(t, \Psi^i) = \sigma_i(t, \psi_1^i, \psi_2^i, \dots, \psi_k^i).$$

При компресії вектор рішень для методу управління бітовою швидкістю складається з двох компонентів: міра інформативності блоку і параметр якості, тобто $\Psi^i = \{\Psi_1^i, \Psi_2^i\}$. Відповідно, для кожного i -го блоку бітова швидкість і середньо квадратична помилка будуть функціями від двох змінних:

$$d(t)_i = d_i(t, \Psi^i) = d_i(t, \psi_1^i, \psi_2^i);$$
$$\sigma(t)_i = \sigma_i(t, \Psi^i) = \sigma_i(t, \psi_1^i, \psi_2^i).$$

Среднеквадратическая помилка для всього кадру визначається виразом:

$$\sigma(t, \Psi) = \prod_{i=1}^k \sigma(t, \Psi^i)$$

Аналогічно, бітові витрати на кадр визначаються як:

$$d(t, \Psi) = \prod_{i=1}^k d(t, \Psi^i)$$

Для оптимізації параметрів компресії необхідно знайти такі значення вектора рішень, які будуть відповідати таким вимогам:

$$\begin{aligned} \min_{\Psi} \sigma(t, \Psi^*) &= \min_{\Psi^i \in \Psi} \sigma(t, \Psi^i); \\ \prod d(t, \Psi^*) &\leq d_{\text{req}}. \end{aligned}$$

де d_{req} - необхідні бітові витрати на один кадр; Ψ^* - оптимальний вектор рішень.

При обробці блоків в якості параметра управління пропонується використовувати тільки фактор якості, який використовується для формування матриці квантування. Так як саме на цьому етапі відбувається корекція компонент трансформант під психовізуальні особливості зорового сприйняття. Міра інформативності блоку визначає метод формування трансформат і його вплив на бітову швидкість не значна.

При формуванні матриць квантування в стандарті JPEG використовується два підходи. Один полягає в тому, що в стандарт JPEG включені дві рекомендовані таблиці квантування: одна для яскравості (табл. 5.1), друга для кольоровості.

Другий підхід полягає в обчисленні значень таблиці квантування в реальному масштабі часу. В цьому випадку для кожного елемента матриці трансформованого зображення існує відповідний елемент матриці

квантування. Результиуюча матриця виходить розподілом кожного елемента матриці трансформант на відповідний елемент матриці квантування і наступним округленням результату до найближчого цілого числа. Як правило, значення елементів матриці квантування ростуть у напрямку зліва направо і зверху вниз.

Таблиця 5.1 - Базова матриця квантування складової яскравості стандарту JPEG

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Від вибору матриці квантування залежить баланс між ступенем стиснення зображення і його якістю після відновлення.

Розрахунок матриці квантування полягає в наступному: задається одне значення фактора якості (Quality Factor - QF) зазвичай в діапазоні [1, 25] і проводиться розрахунок значень матриці за формулою:

$$q(t)_{i,j} = 1 + (1 + i + j) \cdot QF$$

Фактор якості задає інтервал між сусідніми рівнями матриці квантування, розташованими на її діагоналях. Приклад отриманої матриці квантування представлений в табл. 5.2.

Таблиця 5.2 - Матриця квантування з фактором якості рівним 2

3	5	7	9	11	13	15	17
5	7	11	13	15	17	19	21

7	11	13	15	17	19	21	23
9	11	13	15	17	19	21	23
11	13	15	17	19	21	23	25
13	15	17	19	21	23	25	27
15	17	19	21	23	25	27	29
17	19	21	23	25	27	29	31

У даній роботі пропонується управляти не тільки кількістю біт на стиснений сегмент, але і візуальним якістю кожного сегмента. Необхідно знайти значення фактора якості QF_{opt} , при якому середньоквадратическая помилка буде мінімальна $\sigma(t, \Psi^*) = \min \sigma(t, \Psi^i)$, де $\Psi^i \in \Psi$, з урахуванням того що швидкість не буде перевищувати необхідного значення бітової швидкості $d(t, \Psi^*)_i \leq d_{req}$, яке дорівнює пропускну здатності каналу передачі.

Для пошуку оптимального коефіцієнта QF використовуємо відомий метод поділу відрізка навпіл (дихотомії), який, по відношенню до інших методів, є більш швидким, простим і забезпечує задану точність (ε).

Позначимо QF як параметр λ і відповідно QF_{opt} як параметр λ_{opt} . Алгоритм реалізації пошуку для даної задачі, можна представити у вигляді такої послідовності етапів:

Підготовчий етап.

Визначити такі значення $\lambda_1 = \min$ та $\lambda_2 = \max$, такі що свідомо справедливо

$$d(t, \Psi_{\lambda_2}) < d_{req} < d(t, \Psi_{\lambda_1})$$

Ітерація.

Знайти середнє значення λ відповідно до методу дихотомії на інтервалі $[\lambda_1; \lambda_2]$:

$$\lambda = \frac{\lambda_1 + \lambda_2}{2}$$

для наближення до значення λ_{opt} .

Для перевірки відповідності $\lambda \cong \lambda_{\text{opt}}$ необхідно обчислити: $d(t, \Psi_\lambda)$, $\sigma(t, \Psi_\lambda)$.

Якщо $d(t, \Psi_\lambda) > d_{\text{req}}$ та $\sigma(t, \Psi_\lambda) \leq \sigma_{\text{req}}$ то зміщується нижня межа інтервалу ($\lambda_1 = \lambda$);

Якщо $d(t, \Psi_\lambda) > d_{\text{req}}$ та $\sigma(t, \Psi_\lambda) > \sigma_{\text{req}}$ то дана задача не має рішення і алгоритм завершує роботу.

Якщо $d(t, \Psi_\lambda) \leq d_{\text{req}}$, то зміщується верхня межа інтервалу ($\lambda_2 = \lambda$).

Перевірка умови закінчення алгоритму.

Якщо $|\lambda_2 - \lambda_1| < \varepsilon$, то пошук рішення завершений і результат є оптимальним, тобто $\Psi^* = \Psi_\lambda$.

Якщо немає, то здійснюється перехід до наступної ітерації.

Блок-схема алгоритму реалізації пошуку оптимального параметра для управління бітової швидкістю представлена на рис. 5.1.

Розглянемо приклад роботи даного алгоритму.

Аналіз зміни фактора якості QF від мінімуму до максимуму показав, що бітова швидкість і виникає при цьому помилка відповідно буде змінюватися за графіками зазначеним на рис. 5.2.

Надаємо значення $\lambda_1 = QF_{\text{min}}$ та $\lambda_2 = QF_{\text{max}}$, такі що виконуються умови

$$d(t, \Psi_{\lambda_2}) < d_{\text{req}} < d(t, \Psi_{\lambda_1})$$

Проводимо першу ітерацію. знаходимо

$$\lambda^1 = \frac{\lambda_1 + \lambda_2}{2}$$

і обчислюємо для даного значення $d(t, \Psi_\lambda)$ та $\sigma(t, \Psi_\lambda)$. Проводимо порівняння $d(t, \Psi_\lambda) > d_{\text{req}}$. Якщо умова не виконується, то бітову швидкість необхідно збільшувати для поліпшення якості зображення.

В цьому випадку присвоюємо $\lambda_2 = \lambda$ і далі розглядається діапазон $[\lambda_1; \lambda_2]$, як показано на малюнку 5.2. якщо умова $d(t, \Psi_\lambda) > d_{req}$ виконується, проводимо перевірку по середньквдратической помилку: $\sigma(t, \Psi_\lambda) \leq \sigma_{req}$.

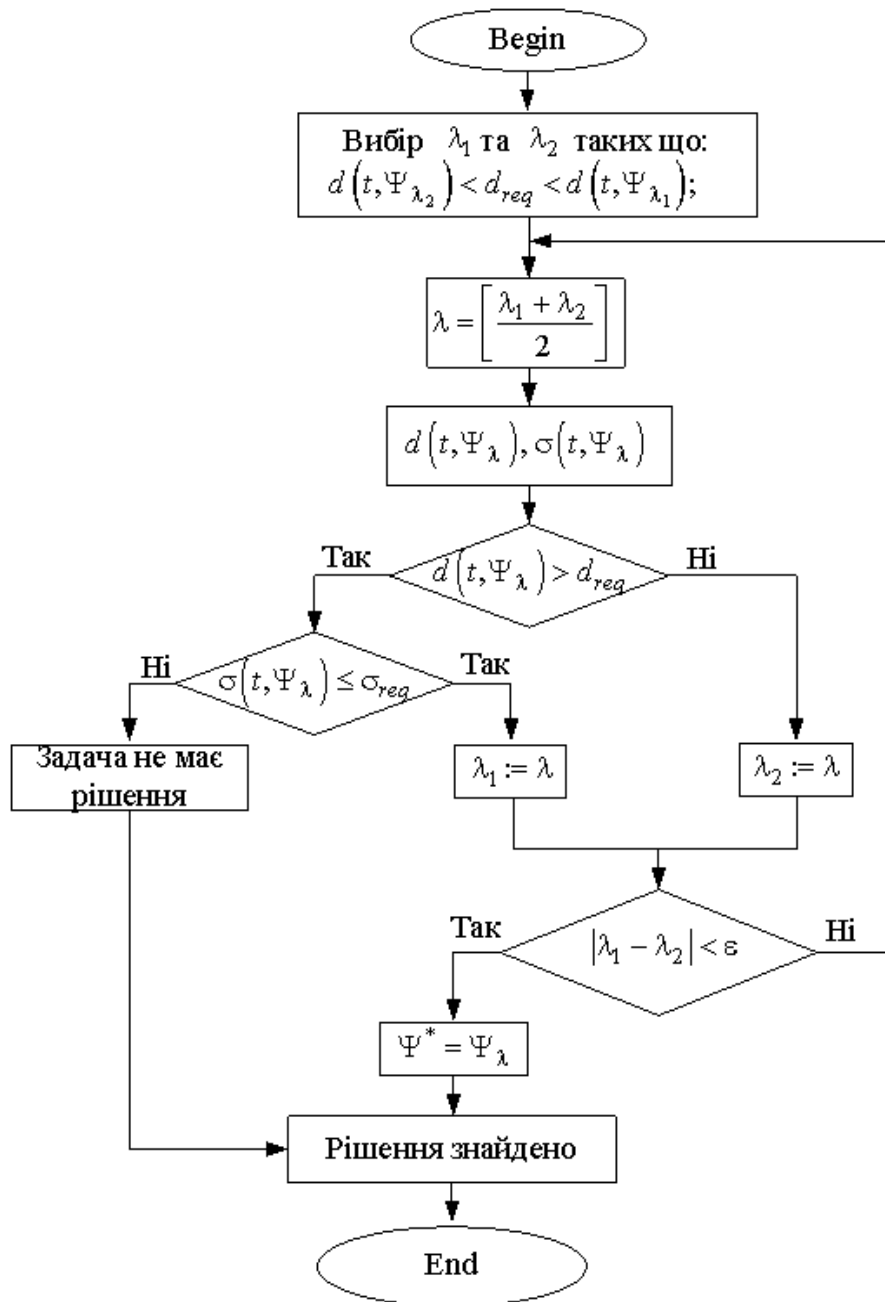


Рисунок 5.1 - Блок-схема алгоритму реалізації пошуку оптимального параметра управління

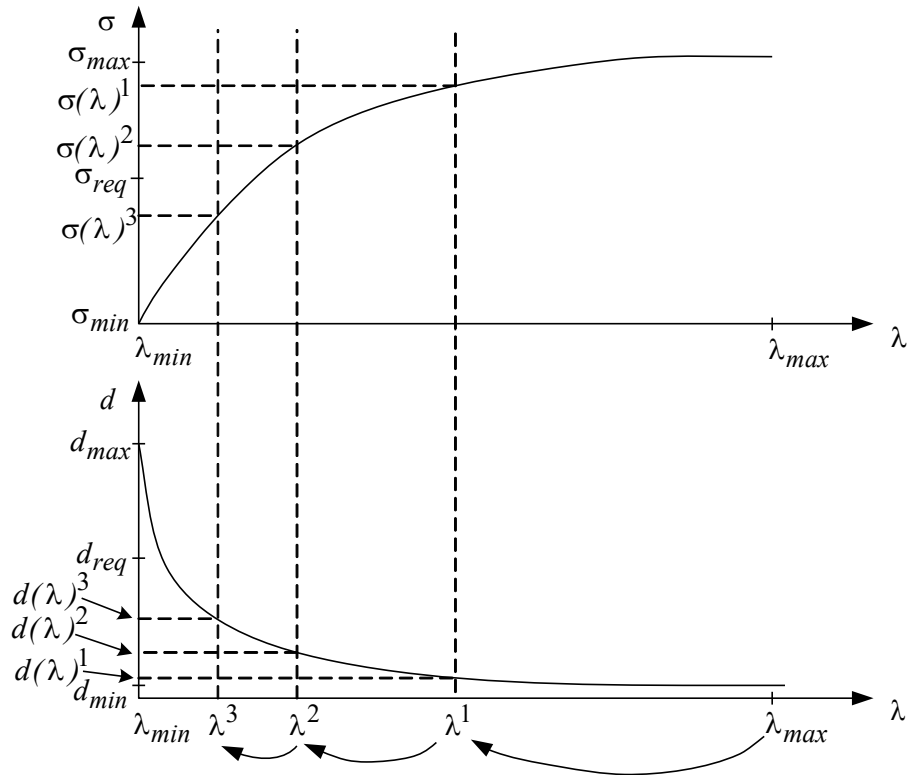


Рисунок 5.2 - Ілюстрація роботи алгоритму пошуку оптимального параметра

У разі якщо і швидкість і помилка перевищують необхідні значення не виконується жодна з умов: $d(t, \Psi_\lambda) > d_{req}$, $\sigma(t, \Psi_\lambda) > \sigma_{req}$, бачимо що оптимізація даним методом стиснення неможлива і необхідний вибір іншого методу, з тих що пропонуються стандартом JPEG. При значенні помилки менше необхідної $\sigma(t, \Psi_\lambda) \leq \sigma_{req}$ присвоюємо $\lambda_1 = \lambda$ і далі знаходження оптимального значення буде проводитися в діапазоні $[\lambda_1; \lambda_2]$. Таким чином, після першої ітерації в наведеному прикладі буде проведений зрушення $\lambda_{max} \rightarrow \lambda$.

Останнім етапом проводиться перевірка оптимальності. перевіряємо умову $|\lambda_2 - \lambda_1| < \varepsilon$, де параметр ε - показує задану точність обчислень. Якщо необхідна точність була досягнута, то оптимальним значенням вважаємо $\sigma(t, \Psi_\lambda)$. Рішення завдання знайдено. Інакше пошук триває і проводиться виконання наступної ітерації.

ВИСНОВКИ

Відповідно до технічного завдання, у рамках виконання атестаційної роботи необхідно було вирішити наступні завдання:

- проаналізувати поточний стан розвитку інфокомунікацій;
- дослідити існуючі підходи до забезпечення ефективного передавання відеоданих у мережі;
- сформувати загальну концепцію технології забезпечення ефективної трансляції відеопотоку у реальному часі з забезпеченням якості відеоданих.

Під час виконання роботи було досліджено проблематику, пов'язану з особливостями розвитку сучасних інфокомунікацій з позицій можливості забезпечення ефективної передачі відеотрафіку.

Як показав даний аналіз, головна проблема розвитку інфокомунікацій полягає у тому, що ріст об'ємів трансльованих відеоданих, на фоні вимог до їх якості, випереджає темпи збільшення пропускної здатності мереж. Тому завдання, яке слід вирішити для усунення даної проблематики, полягає у створенні та використанні методів та технологій узгодження параметрів відео потоку та мережі.

За результатом розгляду існуючих підходів по управління бітовою швидкістю відеотрафіку було зроблено висновок, що наявні метод та технології існують як системи рекомендацій та загальних розробок з відкритими технологічними концепціями, або мають суттєві обмеження до застосування.

З огляду на це, було сформовано загальні умови ефективної передачі відеоданих з позицій забезпечення максимальної якості, та запропоновано концептуальну модель управління бітовою швидкістю, яка базується на архітектурних засадах ProgressiveJPEG та включає у себе:

- механізм декомпозиції фрагментів відео потоку до рівня окремих бітових площин трансформант;

- механізм вибору оптимальних комбінацій кількості та способу розміщення бітових площин трансформант для їх синтаксичного опису на базі динамічного програмування, що забезпечує зміну бітової швидкості з внесенням мінімальних спотворень даних під час обробки;

- систему підходів до зниження об'єму обчислень, а саме – двокаскадну обробку та виявлення структурної подібності фрагментів.

Обґрунтовано, що застосування вищеназваних методів зниження об'єму обчислень у процесі динамічного програмування сприяє можливості реалізації розглянутої концепції управління на базі офісного ПК у випадках, коли мова йде про обробку відео формату FullHD та HD-ready.

Розроблено систему службових даних для організації передавання відеотрафіку на базі методу управління.

Така система не вносить суттєвого зростання об'єму даних, є тривіальною у реалізації та дозволяє однозначно визначати структурні одиниці відео потоку, які може бути відсічено з внесенням мінімальної помилки.

Розроблено метод управління бітової швидкістю роботи кодера, використовуючи в якості параметра фак-тор якості при квантуванні блоку. Для пошуку оптимального значення параметра квантування було прийнято рішення використовувати метод поділу відрізків навпіл. Основна його перевага полягає в тому, що не потрібно повний (або близький до повного) перебір безлічі рішень як, наприклад, при динамічному програмуванні. Це дозволяє знизити час обробки і передачі кадру, що необхідно при обробці відеопослідовності в реальному масштабі часу.

Простота методу дозволяє також знизити навантаження на обчислювальний апарат кодера, тому є можливість його використання в системах в обмеженими обчислювальними здібностями.

Таким чином, усі пункти технічного завдання виконано повністю.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. VNI Forecast Highlights Tool [Електронний ресурс] – Режим доступу: https://www.cisco.com/c/m/en_us/solutions/service-provider/vni-forecast-highlights.html
2. Олифер В. Г. Компьютерные сети. Принципы, технологии, протоколы : учебник для вузов / В. Г. Олифер, Н. А. Олифер. – 3-е изд. – СПб. : Питер, 2006. – 958 с.
3. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М. : Техносфера, 2005. – 1073 с.
4. Ватолин Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М. : ДИАЛОГ – МИФИ, 2003. – 384 с.
5. Миано Дж. Форматы и алгоритмы сжатия изображений в действии: учебное пособие / Дж. Миано. пер. с англ. – М. : Триумф, 2003. – 336 с.
6. Айфичер Эммануил С. Цифровая обработка сигналов: практический поход / Эммануил С. Айфичер, Барри У. Джервис. – 2-е изд. – М. : Вильямс, 2008. – 992 с
7. ITU-T Recommendation H.320. Directory services architecture for audiovisual and multimedia services.
8. Сэломон Д. Сжатие данных, изображений и звука / Д. Сэломон. – М.: Техносфера, 2004. – 368 с.
9. Ричардсон Ян. H.264 and MPEG-4 Video Compression: Video Coding for Next-Generation Multimedia / Ян Ричардсон. – Город. : Издательство, 2005. – 368 с.
10. Official JPEG2000 Homepage. [Електронний ресурс] – Режим доступу: <http://www.jpeg.org/jpeg2000.html>.
11. Shi, Yun Q. Image and video compression for multimedia engineering: fundamentals, algorithms, and standards / Yun Q Shi, Huifang Sun.

12. Беляев Е.А, Тюрликов А.М. Управление скоростью и ошибкой кодирования в системе сжатия и передачи видеоинформации с ограничениями на память передающего и принимающего устройств // Компьютерная оптика. 2007. - Т.31. №2. С.69-76.
13. Беллман Р., Енджел Э. Динамическое программирование и уравнения в частных производных. М.:Мир, 1974.
14. Зайченко О.Ю., Зайченко Ю.П., Дослідження операцій. К.: Слово, 2007. – 427 с.
15. Мартинюк О.М., Попіна С.Ю., Елементи комбінаторики й класичне означення ймовірності. Тернопіль, 2003. – 40 с.
16. Сэломон, Д. Сжатие данных, изображений и звука [Текст]/ Д. Сэломон.- Москва: Техносфера, 2004. - 368с.
17. Метод компрессии видеопотока на основе полиадического кодирования предсказываемых кадров / Н.А. Харченко, В.Н. Кривонос // Радиоэлектроника и информатика. - №1. – 2013. – С. 21 – 28.
18. Ян Ричардсон. Видеокодирование. H.264 и MPEG-4 - стандарты нового поколения. [Текст]/ Ян Ричардсон.- М.: Техносфера, 2005. - 368 с.