

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ
ФАКУЛЬТЕТ АЕРОНАВІГАЦІЇ, ЕЛЕКТРОНІКИ ТА ТЕЛЕКОМУНІКАЦІЙ
КАФЕДРА ЕЛЕКТРОНІКИ, РОБОТОТЕХНІКИ І ТЕХНОЛОГІЙ
МОНІТОРИНГУ ТА ІНТЕРНЕТУ РЕЧЕЙ

ДОПУСТИТИ ДО ЗАХИСТУ
Завідувач випускової кафедри
_____ Шутко В.М.
«___» _____ 2021 р.

КВАЛІФІКАЦІЙНА РОБОТА

ЗДОБУВАЧА ОСВІТНЬОГО СТУПЕНЯ БАКАЛАВРА
ЗІ СПЕЦІАЛЬНОСТІ 171 «ЕЛЕКТРОНІКА»
ОПП «ЕЛЕКТРОННІ СИСТЕМИ»

Тема: «Алгоритм стиснення цифрового відео з урахуванням процесорних витрат»

Виконавець

студент групи ЕС-413Б
Керівник

_____ Березівський Назарій Миколайович

д.т.н., професор

_____ Шутко Володимир Миколайович

Нормоконтролер

_____ Сініцин Р.Б.

КИЇВ 2021

НАЦІОНАЛЬНИЙ АВІАЦІЙНИЙ УНІВЕРСИТЕТ

Навчально-науковий інститут аеронавігації електроніки та телекомунікацій
Кафедра електроніки, робототехніки і технологій моніторингу та інтернету
речей

Напрямок (спеціальність) 171 «Електроніка»

(шифр, найменування)

ЗАТВЕРДЖУЮ

Завідувач кафедри

_____ Шутко В.М.

« _____ » _____ 2021р.

ЗАВДАННЯ**на виконання дипломної роботи**

Березівський Назарій Миколайович

(прізвище, ім'я, по батькові випускника в родовому відмінку)

1. Тема дипломної роботи: «Алгоритм стиснення цифрового відео з урахуванням процесорних витрат»

затверджена наказом ректора від « 1 » квітня _____ 2021_р.

№ 526/ст _____

2. Термін виконання роботи : з 17 травня 2021р по 18 червня 2021р.

3. Вихідні дані до роботи: сьогодення обумовлене тим, що більшість людей використовують свої гаджети як інструмент для роботи з відео, яке зменшує час роботи пристрою. Цей проект дасть змогу використовувати його, переглядаючи відео та витратити на 30-40% менше заряду батареї.

4. Зміст пояснювальної записки: Вступ; Розділ 1: ЗАГАЛЬНО-ТЕОРЕТИЧНА ХАРАКТЕРИСТИКА ОСОБЛИВОСТЕЙ СТИСНЕННЯ ЦИФРОВОГО ВІДЕО;

Розділ 2: АНАЛІЗ МЕТОДОЛОГІЧНОЇ БАЗИ РОЗРОБКИ АЛГОРИТМУ
СТИСНЕННЯ ВІДЕО З УРАХУВАННЯМ ПРОЦЕСОРНИХ ВТРАТ;

Розділ 3: ДОСЛІДЖЕННЯ КОДУВАННЯ ВІДЕО ЗА ДОПОМОГОЮ
ШВИДКОГО ПЕРЕТВОРЕННЯ

5. Календарний план-графік

№ з/п	Завдання	Термін виконання	Відмітка про виконання
1.	Затвердження теми кваліфікаційної роботи	15.03.2021	Виконав
2.	Затвердження завдання на кваліфікаційну роботу і плану кваліфікаційної роботи	01.04.2021	Виконав
3.	Підготовка і подання науковому керівнику вступу та I-го розділу кваліфікаційної роботи	20.04.2021	Виконав
4.	Підготовка і подання науковому керівнику II-го розділу кваліфікаційної роботи	02.05.2021	Виконав
5.	Підготовка і подання науковому керівнику III-го розділу кваліфікаційної роботи, висновків і резюме	20.05.2021	Виконав
6.	Написання висновка	01.06.2021	Виконав
7.	Оформлення роботи	05.06.21	Виконав
8.	Подання на кафедру остаточного тексту кваліфікаційної роботи в паперовому і електронному варіанті, заяви студента про допуск до захисту, висновку наукового керівника і відгуку від підприємства (організації) – бази дослідження	09.06.21	Виконав

6. Дата видачі завдання: “15” березня 2021 р.

Керівник дипломної роботи (проекту)

(підпис керівника)

Шутко В.М.

(П.І.Б.)

Завдання прийняв до виконання

(підпис випускника)

Березівський Н.М.

(П.І.Б.)

РЕФЕРАТ

Основна мета знайти можливості кодувати відео за допомогою швидкого перетворення Уолша-Адамара. Це значно зменшить навантаження на процесор телефону під час роботи з відео (відеодзвінок, запис і перегляд відео) і дасть можливість економити до 30-40% заряду батареї. Тобто, люди зможуть значно довше спілкуватися по відеозв'язку, переглядати відеоматеріали, втративши лише 3-7% від якості відео.

ЗМІСТ

ВСТУП.....	7
РОЗДІЛ 1. ЗАГАЛЬНО-ТЕОРЕТИЧНА ХАРАКТЕРИСТИКА ОСОБЛИВОСТЕЙ СТИСНЕННЯ ЦИФРОВОГО ВІДЕО	9
1.1 Поняття цифрового відео та його характеристики	9
1.2. Програмне забезпечення для роботи із цифровим відео	16
1.3. Поняття якості стиснення.....	17
1.4. Аналіз сучасних стандартів стиснення цифрового відео	19
1.5. Технології стиснення цифрового відео	26
РОЗДІЛ 2. АНАЛІЗ МЕТОДОЛОГІЧНОЇ БАЗИ РОЗРОБКИ АЛГОРИТМУ СТИСНЕННЯ ВІДЕО З УРАХУВАННЯМ ПРОЦЕСОРНИХ ВТРАТ	29
2.1. Класифікація методів стиснення цифрового відео	29
2.2. Основні алгоритми стиснення відео	36
2.3. Порівняльна характеристика алгоритмів стиснення відео та обгурнтування обраного методу	45
РОЗДІЛ 3. ДОСЛІДЖЕННЯ КОДУВАННЯ ВІДЕО ЗА ДОПОМОГОЮ ШВИДКОГО ПЕРЕТВОРЕННЯ	51
3.1. Дискретне косинус не перетворення	51
3.2. Колування за допомогою швидкого перетворення Уолша-Адамара	57
ВИСНОВКИ	60
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ТА ЛІТЕРАТУРИ	61

ВСТУП

Актуальність дослідження. Особливість сучасних інформаційно-обчислювальних систем – обробка та передача великих обсягів відеоданих, при цьому швидкість може бути обмежена. Це призводить до затримок при передачі та втрат частини інформації.

Пропускна спроможність систем передачі даних не може зростати одночасно зі збільшенням кількості інформації. Ця проблема має три рішення. Перше – обмеження кількості інформації, але, на жаль, це не завжди допустимо. Наприклад, для динамічних даних це означає зменшення розширення, що призведе до втрати, і може зробити зображення взагалі некорисним (наприклад для медичних або космічних зображень). Друге – збільшення об'єму носіїв інформації і пропускну спроможності каналів зв'язку. Це рішення пов'язано з матеріальними витратами, між іншим, деколи дуже значними. Третє рішення – використання стиснення інформації. Це рішення дає змогу в декілька разів зменшити вимоги до об'єму пристроїв зберігання інформації і пропускну спроможності каналів зв'язку без додаткових витрат (крім витрат на реалізацію алгоритмів стиснення). Умовами його застосування є надлишок інформації і можливість встановлення спеціального програмного забезпечення або апаратури як поблизу джерела, так і поблизу приймача інформації. Саме завдяки необхідності використання стиснення інформації методи стиснення є досить поширеними.

Розробляються різні проекти передачі якісного відео: через Internet у реальному часі; по мобільному зв'язку через супутник з кіностудії в кінотеатри із забезпеченням показу на великому екрані; з мультимедійних баз даних по супутникових каналах для інтерактивного телебачення.

Розробкою даної проблеми займалися багато науковців, представники різних галузей науки: О. В. Дробик, В. В. Кідалов, В. В. Коваль, Б. Я. Костік, В. С. Лазебний, Г. М. Розорінов, Г. О. Сукач та інші.

Але, незважаючи на це, сьогодні існує потреба у дослідженні, яке б узагальнило, систематизувало існуючі відомості з даної проблеми.

Враховуючи все вищесказане, нами і була обрана тема дипломної роботи: "Алгоритм стиснення цифрового відео з урахуванням процесорних витрат".

Об'єкт дослідження – сучасний стан розвитку наукових досліджень у сфері стиснення цифрового відео.

Предмет – способи та методи стиснення цифрового відео з урахуванням процесорних витрат.

Мета роботи: дослідити основні аспекти реалізації алгоритму стиснення цифрового відео з урахуванням процесорних витрат.

Відповідно до мети були визначені наступні **завдання**:

- 1) розкрити теоретичні аспекти стиснення цифрового відео;
- 2) проаналізувати методологічну базу розробки алгоритму стиснення відео з урахуванням процесорних витрат;
- 3) знайти можливість зменшити затрату на перетворення відео;

Для розв'язання поставлених завдань нами були використані такі **методи дослідження**: теоретико-критичний аналіз літератури з теми дослідження; зіставлення, узагальнення і синтезування здобутої інформації тощо.

Робота може бути використана студентами ВНЗ для підготовки до семінарських занять, також може бути використана викладачами для проведення лекції, практик тощо.

Структура роботи. Робота складається зі вступу, 3 розділів, висновків, списку використаних джерел, що містить 19 найменувань. Повний обсяг роботи: 61 сторінки.

РОЗДІЛ 1

ЗАГАЛЬНО-ТЕОРЕТИЧНА ХАРАКТЕРИСТИКА ОСОБЛИВОСТЕЙ СТИСНЕННЯ ЦИФРОВОГО ВІДЕО

1.1 Поняття цифрового відео та його характеристики

Цифрове відео являю собою перетворений у цифровий формат аналоговий сигнал. При цьому сама процедура перетворення неминуче приводила до деякої втрати якості. На сьогоднішній день відеомагнітофони й старенькі Vhs- Камери пішли в минуле, а балом правлять сучасні цифрові відеокамери, DVD і Blu-ray плеєри, які дозволяють одержувати сигнал відразу в цифровому виді. Та й аналогове телебачення поступове поступається місцем більш прогресивному цифровому. Цифрове відео має п'ять основних характеристик: екранне розширення, частота кадрів, глибина кольору, бітрейт (ширина відеопотоку) і якість зображення.

Екранне розширення (Resolution) – позначає кількість крапок (пікселів) по горизонталі й вертикалі, з яких полягає зображення (відеокадр) на екрані. При записі дозволу спочатку вказується значення кількості крапок у рядку (горизонтальний розширення), а потім число рядків, що брав участь у побудові зображення (вертикальний розширення). Наприклад, для європейського відеостандарту PAL розмір кадра становить 720x576 пікселів, для північноамериканського стандарту NTSC – 720x480, для відео високої чіткості (HD 720p) – 1280x720, а для новомодного стандарту HDTV (Full HD) - 1920x1080 крапок. Як ви, напевно, розумієте, чим вище екранне розширення, тем якість відео краще.

Частота кадрів – величина, що вказує, на те, яка кількість кадрів переміняється за секунду. Стандартною швидкістю відтворення відеосигналу

вважається величина рівна 30 кадрам/с. Для кіно цей показник трохи менше й становить 24 кадра/с.

Глибина кольору (колірний розширення) – характеристика, що вказує кількість квітів, які можуть брати участь у формуванні відеозображення. Кількість квітів у цифровому відео виміряється в бітах. Так 1 біт може ухвалювати два різні значення (0 або 1) і дозволяє відповідно закодувати тільки два кольори (звичайно чорний і білий). За допомогою двох біт можна закодувати вже 4 кольору ($2^2 = 4$), за допомогою трьох біт – 8 квітів (2^3), чотирьох – 16 (2^4) і так далі.

Як правило, колірний розширення описується за допомогою спеціальних колірних моделей. У комп'ютерній техніці застосовується модель RGB (червоний-зелений-синій), яка може бути представлена наступними найпоширенішими режимами глибини кольору: 8 біт (256 квітів), 16 біт (65 536 квітів) і 24 біта (16 777 216 квітів). До слова, людське око, по різних думках, може сприйняти від 5 до 10 мільйонів колірних відтінків.

Битрейт (ширина відеопотоку) – показує кількість оброблених біт відеоінформації за одну секунду часу. Інакше кажучи – це швидкість відеопотоку, яка виміряється в мегабітах у секунду (Мбіт/с). Чим вона вище, тем краще якість. Наприклад, для стандарту Dvd- Відео ширина потоку становить близько 5 Мбіт/с, а для формату телебачення високої чіткості HDTV – уже 10 Мбіт/с. До речі, найбільш часте значення битрейта використовується для оцінки якості переданого відео через Інтернет.

Якість зображення – характеристика покликана оцінити якість обробленого відео в порівнянні з оригіналом, що й визначається сукупністю значень дозволу, глибини кольору й швидкості відеопотоку.

Одна хвилина чистого незжатого оцифрованого звуку займає на жорсткому диску комп'ютера близько 10 Мбайт, внаслідок чого, у переважній більшості, музичні файли зберігаються в стислому виді заради економії місця. А

скільки ж займає одна хвилина незжатого відео? Наприклад, для розміщення 60-секундного ролика із частотою 30 кадрів у секунду, дозволом 720x576 пікселів і 16-бітною глибиною кольору буде потрібно близько півтора гігабайт вільного дискового простору. І це без обліку звукової доріжки. Після цих цифр, напевно, уже не потрібно пояснювати, чому цифрове відео зберігається в наших комп'ютерах винятково в стислому виді.

Існує кілька десятків популярних форматів стиснення, які використовують різні алгоритми компресії, які відповідно дають різні результати.

DV (Digital Video) – один з найперших алгоритмів стиснення для відеопотоку, розробка якого почалася в 1993 році спільно відразу декількома компаніями, що є найбільшими виробниками відеоустаткування (Sony, JVC, Panasonic, Philips і Hitachi). Формат DV забезпечує невисокий ступінь стиснення даних (5:1) і характеризується високим битрейтом, за рахунок чого вихідний відеофайл виходить досить великого розміру. Так одна хвилина Dv-Відео займає близько 200 Мб (1 година – 12 Гб) на цифрових носіях інформації.

Найбільше часто цей формат використовується для стиснення при відеозйомці за допомогою побутових цифрових камер і професійних камкордерів. При цьому через невеликий коефіцієнт стиснення зняті матеріали виходять дуже високої якості, а сама процедура компресії, яка відбувається в реальному часі, не вимагає потужних технічних компонентів.

Правда, зберігати відео на домашньому комп'ютері й тим більше оптичних дисках в Dv-форматі все-таки незручно, уже занадто багато воно займає місця. Так що фахівцям довелося задуматися про додаткові алгоритми стиснення, за допомогою яких удалося б скоротити розмір цифрового фільму ще в кілька раз.

MPEG (Moving Picture Experts Group) – ціле сімейство стандартів стиснення цифрової інформації, розроблене й стандартизоване однойменної експертною групою фахівців, сформованою організацією ISO у далекому 1988 році.

Першим плодом їх утвору став вихідний стандарт відео й аудио компресії MPEG-1, а в 1993 році при участі компаній JVC і Philips, була розроблена його специфікація Video CD (VCD), яка й відома багатьом користувачам. З назви видно, що VCD є форматом для зберігання стислого відео зі звуком на звичайних компакт дисках.

Використання для кодування алгоритмів MPEG-1 дозволяє одержувати відеопотік шириною до 1,5 Мбіт у секунду з дозволом кадра 352x288 крапок для PAL або 352x240 для NTSC, після чого на звичайному CD може вміститися 74 хвилини відео зі звуком якості VHS (як у звичайного відеомагнітофона).

В 1995 році побачив світло найпопулярніший стандарт MPEG-2, який згодом одержав широке поширення в цифрових відеодисках DVD, а так само при передачі сигналу кабельного й супутникового телебачення. Якість картинки тут значно вище, чим у попередника: при 25 кадрах у секунду, розширення становить 720x576 крапок для системи PAL, а для системи NTSC - 720x480 при 30 кадрах/с. При цьому, середня максимальна ширина потоку рівна 9,8 Мбіт/з, що практично в 7 раз вище, чим в Video CD. Ще однією незаперечною перевагою MPEG-2 є можливість збереження п'ятиканальної аудіодорожки (Dolby Digital 5.1 і DTS).

Максимальна ємність двошарового DVD диска (DVD-9) становить 8,5 Гбайт, на який можна записати до трьох годин відео з повною якістю. Якщо ж вам пропонують DVD відразу з декількома фільмами, то знайте, що, швидше за все вас очікує низькосортна картинка рівня Video CD з дуже низьким дозволом і битрейтом.

Разом з MPEG-2, приблизно в той час, початків розроблятися новий стандарт MPEG-3, призначений для кодування аудио й відеопотоків у телебаченні високої чіткості зі швидкістю передачі даних від 20 до 40 Мбіт/с. Але незабаром з'ясувалося, що для цих завдань можна використовувати трохи модифіковану версію стандарту MPEG-2, після чого всі подальші розробки

MPEG-3 були припинені й на сьогоднішній день цей стандарт не використовується.

Варто відзначити, що досить часто термін "MPEG-3" багато користувачів асоціюють із популярною технологією стиснення звуку MP3. Але це в корені не вірно, тому що її правильна назва – MPEG-1 Audio Layer 3.

Нарешті, в 1998 році з'явилося нове сімейство форматів стиснення відео - MPEG-4. Розроблялося воно з метою поліпшення якості картинки при низькій швидкості потоку. Колишній стандарт MPEG-2, розрахований на високий бітрейт, із цим завданням упоратися не міг, так що алгоритми стиснення довелося серйозно модифікувати. Так само MPEG-2 не походить і для зберігання відео високої чіткості (HD) з дозволами від 1280x720 (720p) до 1920x1080 пікселів (1080i або 1080p), яке усе більше й більше набирає популярність.

На сьогоднішній день MPEG-4 є основним стандартом стиснення мультимедиа контенту, і хоча DVD списувати з рахунків ще рано, практично всі сучасні фото й відеокамери знімають в Hd- Якості. Так що для збереження відео з таких обладнань на комп'ютер, у кожному разі прийде орієнтуватися на кодеки сімейства MPEG-4.

MPEG-4 Part 2 ASP – один з перших алгоритмів, що з'явилися в 1999 році. Кодеки, побудовані на його основі, забезпечують досить низька якість вихідних матеріалів, що й не мудро. Адже в той час ніяким відео високої чіткості ще й не пахнуло. Зате висока швидкість роботи й невимогливість до апаратних ресурсів почасти компенсує цей недолік. Саме тому й сьогодні цей алгоритм широко затребуваний при кодуванні відео для різних мобільних обладнань і компактних медиаплеєров, а так само роликів, розташовуваних у мережі.

Одними із самих яскравих представників кодеков, що базуються на основі цих алгоритмів, є знайомі багатьом користувачам, комерційний Divx і його безкоштовна альтернатива Xvid.

MPEG-4 AVC або H.264 - один із самих останніх і популярних алгоритмів, що використовуються з успіхом як для стиснення відео з низьким дозволом, так і HD контенту. До речі, більшість високоякісних фільмів на дисках Blu-ray кодується саме цим кодеком. Так само він часто використовується й у побутових Hd- Відеокамерах (AVCHD).

Як і в попередньому випадку, у цього сімейства кодеків існують як безкоштовні модифікації, наприклад, x.264, так і комерційні варіанти, що входять до складу популярних відеоредакторів (Adobe Premiere, Pinnacle Studio і інші).

VC-1 – кодек, розроблений всюдисущою компанією Microsoft і стандартизований в 2006 році. У його основу покладений власний формат відеостиснення WMV (Windows Media Video) і система кодування WMV 9. Споконвічно завданням VC-1 було кодування ігрового відео для приставок Xbox. Однак на сьогоднішній день цей кодек уже вийшов на відеоринок, і активно конкуруючи з H.264, є підтримуваним стандартом для формату Blu-ray.

Як і будь-яка інша цифрова інформація, відео зберігається на диску у вигляді файлів, або як їхнім ще називають, медиаконтейнерів відео, що містять, -, аудио й інші потоки, а так само метадані. У будь-який момент із контейнера можна вийняти, наприклад, відео або аудiodорожки, перекодувати їх, і помістити їх в інший контейнер, тобто змінити формат відеофайлу. Мультимедійні контейнери можуть бути різних типів (форматів), а на те, до якого виду вони ставиться, указує розширення файлу.

Незважаючи на те, що більшість контейнерів прив'язана до певного формату, у деяких з них може зберігатися відео в зовсім різних стандартах. Наприклад, файл із розширенням AVI здатний містити ролики як у форматі MPEG-1, так і в MPEG-2 або в MPEG-4. На що ж тоді впливає тип контейнера?

Звичайно, у більшій мері якість фільму визначається кодеком і тими параметрами, які були встановлені при стиснення. Але й від контейнера

залежить чимало. Різні види відеофайлів мають певні вимоги й обмеження по кількості звукових доріжок, каналів субтитрів, типів використовуваних кодеков, а так само сумісності з побутовими програвачами й плеєрами.

Найпопулярніші формати відеофайлів, їх переваги та недоліки

AVI (Audio Video Interleave) – найдавніший і традиційний із усіх видів медіаконтейнеров, який був уперше використаний Microsoft в 1992 році. Може містити в собі відео й аудіо інформацію, стислу різними комбінаціями кодеков. Таким чином, Avi-файлы при зовнішній подібності можуть дуже сильно відрізнятися внутрішньою «начинкою», а що б точно визначити їхній уміст, прийде скористатися спеціальними програмами (наприклад, Videotoolbox). Не варто забувати й той факт, що MP4 є «рідним» форматом для всіх продуктів компанії Apple, від iPhone до Mac. Тому якщо ви шанувальник «яблучного» заліза, те домашню відеокolleкцію краще збирати й зберігати в MP4. Варто відзначити, що контейнер MP4, на відміну від MKV, має ряд обмежень і не може містити відео стандартів MPEG-1, MPEG-2 і WMV, а так само звук у форматах AC-3 (Dolby Digital) і WMA.

VOB (Versioned Object Base) – основний контейнер, використовуваний для зберігання мультимедіа контенту на Dvd- Дисках. Може містити кілька потоків відео MPEG-2, до дев'яти аудіодорожок, до 32х каналів із субтитрами й екранне меню.

FLV (Flash Video) – медіаконтейнер, що використовується для розміщення й передачі відеороликів у глобальній мережі Інтернет. Використовується багатьма великими сервісами видеохостинга, такими як Rutube, Youtube, Vimeo, Flickr і іншими. Відеопотік в Flv-файлі як правило закодований за допомогою кодеков H.263 або H.264, а звук в MP3 або AAC.

MOV-формат файлу, розроблений компанією Apple для зберігання відео, графіки, анімації й 3D. Своєю появою зобов'язаний технології відтворення медіаконтента Quicktime.

TS і M2TS – спеціалізовані контейнери для зберігання Hd- Відео. Тсиспользується в потоковому віщанні цифрового телебачення IPTV і DVB. Правда, даний контейнер взагалі не може містити субтитри. M2TS є стандартним контейнером для Blu-ray відео, у який можуть бути включені відео й аудіопотоки, передбачені стандартом BD-ROM, а так само субтитри в графічному форматі PGS.

1.2. Програмне забезпечення для роботи із цифровим відео

Медіаплеєри – програми для декодування й відтворення відеоконтенту. Найбільш популярними з них є Windows Media Player, Kmpayer, Winamp, Quicktime, GOM Player, Powerdvd, Media Player Classic, VLC Media Player, Bsplayer, Realplayer і інші.

Як правило, після установки програвача, разом з ним у систему встановлюється певний набір кодеків, від якого буде залежати, які медіаконтейнери ви зможете відкривати. Наприклад, деякі з них не вміють за замовчуванням відтворювати DVD, а інші контейнери MKV. Тому, щоб не плодити купу непотрібних додатків на власному ПК і мати можливість переглядати всі види відеофайлів за допомогою одного улюбленого медіаплеєра, досить установити в систему набір окремих кодеків, наприклад K- Lite Codec Pack.

Конвертори – програми, що дозволяють перетворювати відеофайли з одного формату в іншій. Одні з них є мультиформатними, тобто здатні працювати відразу з багатьма видами медіаконтейнеров. Інші, вузькоспеціалізовані, розраховано на конвертацію одного певного формату. До найбільш популярних конверторів можна віднести: Total Video Converter,

Format Factory, Any Video Converter, Xilisoft Video Converter, Mediacoder, Dr.Divx, Imtoo 3GP Video Converter, Convertxtodvd та інші.

Риппери й граббери – програмні інструменти, що дозволяють копіювати фільми з DVD і Blu-ray дисків, з наступною конвертацією в різноманітні формати. Серед них: Xilisoft DVD і Blu-ray Ripper, Imtoo DVD і Blu-ray Ripper, Aleesoft DVD і Blu-ray Ripper, Kingdia DVD Ripper, Best HD Blu-ray Ripper і інші.

Відеоредактори – додатки, що мають набір інструментів для редагування (монтажу) відеофайлів на комп'ютері. Варто відзначити, що за допомогою програм подібного типу виконується ще одна дуже важлива функція – захоплення й оцифровка відео. Таким чином, за допомогою відеоредакторів здійснюється перетворення відеосигналу із зовнішніх джерел (відеокамер, відеомагнітофонів, плеєрів оптичних дисків і т.д.) у цифровий відеопотік, його стиснення і збереження в обраному медіаконтейнері, з метою наступної обробки, зберігання або відтворення.

Найбільш яскравими й професійними представниками цієї групи програм є: Adobe Premiere, Pinnacle Studio, Virtualdub, Corel Videostudio, Sony Vegas Pro, Nerovision, Ulead Videostudio Plus і інші.

1.3. Поняття якості стиснення

Стиснення відео файлів, так само як і зображень, відбувається за рахунок опущення певних деталей зображення для зменшення обсягу файлу. При цьому неминуче втрачається якість зображення, і метою алгоритмів стиснення є знаходження ідеального співвідношення якість/стиснення, такого, при якому досягається максимальна ступінь стиснення і в той самий час деградація

зображення для людського ока залишається мінімально помітною. Виходячи з цього виділяють наступні види стиснення:

1. Без втрат якості. При такому виді стиснення зображення після декомпресії в точності збігається з оригіналом, тобто втрати даних у процесі стиснення не відбуваються.

2. З втратами якості. Стиснення, при якому відбувається втрата даних зображення, тобто оригінал побітно не збігається з результатом.

3. Без втрат якості з точки зору сприйняття. При такому стисненні втрати якості, якщо і відбуваються, то залишаються непомітними для людського ока.

4. З природною втратою якості. У цьому випадку деградація зображення стає помітною, але не переходить межі впізнавання основних об'єктів зображення.

У світі поширені усі види стиснення і вибір кожного з них залежить від вимог до відео файлів, що висуваються у конкретних випадках. Художні фільми не можуть дозволити собі втрачати якість зображення, так як від цього залежить успішність кінофільму серед глядачів, але записи з камер відеоспостереження у супермаркеті не мають таких критеріїв, так як їх обсяг дуже великий і дрібні деталі на них не потрібні. Всі широко використовувані відеокомпресори використовують технології стиснення з втратами якості. При досить високих коефіцієнтах стиснення всі вони будуть стиснені з неприродною втратою якості.

Таким чином, вибираючи той чи інший компресор для стиснення цифрового відео, необхідно досягти стиснення принаймні з природними втратами якості.

1.4. Аналіз сучасних стандартів стиснення цифрового відео

Поширення технологій цифрового відео привело до необхідності розробки ефективних методів стиснення відеопослідовностей. Стиснення відео засноване на двох важливих принципах: просторової надмірності, властивій кожному кадрові відеоряду, і тимчасової надмірності, тобто подібності одного кадра на наступний

Таким чином, типовий метод стиснення полягає в кодуванні першого кадра за допомогою деякого алгоритму стиснення зображень і наступному кодуванні різниці першого кадра й наступних. Якщо черговий кадр сильно відрізняється від першого кадра послідовності, то його кодують повністю, і він стає першим кадром нової послідовності. Для збільшення ефективності усунення тимчасової надмірності використовується компенсація руху [2].

Розглянемо етапи процедури стиснення даних у загальному виді. Будь-який метод стиснення реалізує три основні етапи:

- етап попередньої обробки (фільтрація шумів);
- основне перетворення;
- кодування й упакування компонент перетворення [3].

На другому етапі виконується перетворення вихідних даних з однієї форми вистави в іншу. Зокрема, при стисненні зображень залежно від виду алгоритму стиснення може бути виконаний перехід від вихідного зображення до наступних видів:

- матриці компонентів спектра (при спектральному перетворенні);
- набори коефіцієнтів перетворення (при фрактальному стисненні);
- опис об'єктів зображення (при стисненні з розпізнаванням).

Для різних сфер використання цифрового відео висувалися різні вимоги до стиснення, які привели до формування ряду стандартів стиснення для різних областей застосування:

- для Isdn-відеоконференцій розроблений стандарт стиснення ITU H.
- для відеоконференцій у телефонних мережах - стандарт H.262 і стандарт H.263 для відеоконференцій у мережах АТМ і по широкосмугових каналах;
- необхідність стиснення відеопослідовностей для зберігання на CD-ROM (з умовою забезпечення 1.2 Мбіт/с для відео-потoku й 256 кбит/із для аудіо) привела до створення первісного стандарту ISO MPEG-1;
- для віщання й зберігання на DVD, з битрейтом від 2 до 15 Мбіт/із для відео й аудіо, був розроблений стандарт ISO MPEG-2;
- необхідність кодування окремих аудіо-візуальних об'єктів як природнього походження, так і синтезованих, привела до створення ISO MPEG-4 (цей стандарт містить у собі кілька частин, у яких розглядається, крім кодування відео, аудіокодування, кодування об'єктів і т.д., причому до відео ставляться частина 2 -ISO 14496-2 або MPEG-4 Part 2, і частина 10 - ISO 14496-10 або MPEG-4 Part 10);
- необхідність кодування метаданих, що описують властивості мультимедійних даних, привела до появи MPEG-7.

Базовий профіль націлений на кодування й декодування в реальному часі для мобільних обладнань. Він підтримує прогресивне розгорнення, використовує I- I- і P- Кадри, а також ентропийное кодування по методу CAVLC.

Основний профіль призначений в основному для використання в широкомовленні. Він підтримує чересстрочную й прогресивне розгорнення, використовує I- I-, P-, Вкадры, вагове проорокування (weighted prediction), а також ентропийное кодування по методах CAVLC і CABAC.

Розширений профіль призначений для використання в засобах передачі, підданих помилкам - наприклад, у мобільних комунікаціях. Використовує I- I-, P-, B-, SP-, Si- Кадри, підтримує як чересстрочную, так і прогресивне розгорнення, дозволяє використовувати тільки метод CAVLC для ентропийного

кодування. Профіль для відео високого дозволу призначений для ефективного кодування HDV (high definition video). Він використовує адаптивний розмір блоку (8×8 або 4×4) і дозволяє застосовувати залежні матриці квантування.

Профіль high 10 є розширенням попереднього профілю для 10 біт на відлік компонента зображення. Профіль high 4:2:2 підтримує формат YUV 4:2:2 і до 10 біт на відлік для цветоразностных компонент зображення; а профіль high 4:4:4 підтримує формат YUV 4:4:4 і до 12 біт на відлік цветоразностных компонент зображення. Крім цього він дозволяє використовувати режим кодування без втрат і пряме кодування RGB сигналу. Цей профіль призначений для кодування відео студійної якості.

Кодування здійснюється поблочно. При цьому спочатку проводиться проорокування отсчетов яркостной компоненти й цветоразностных компонент у просторовій і тимчасовій областях. Потім різниця між передвіщеними значеннями й реальними зазнає целочисленному перетворенню й квантується. Після цього результат стисненняється ентропийним кодером. Обробка кожного кадра ведеться в просторі YUV по блоках розміром 16×16 для яркостной компоненти (luma) і по 8×8 (для YUV 4:2:2) для цветоразностных компонент (chroma).

AVC/H.264 визначає два інструменти для кодування бітового потоку – контекстне адаптивне кодування зі змінною довжиною (Context-Adaptive Variable Length Coding, CAVLC) і контекстне адаптивне бінарне арифметичне кодування (Context-Adaptive Binary Arithmetic Coding, CABAC) [3]. В CAVLC кількість коефіцієнтів, не рівних 0, кодується окремо від їхнього значення й положення у векторі. Після зигзаг-сканування коефіцієнтів перетворення більші значення коефіцієнтів, відповідні до низькочастотної складової, розташовуються на початку вектора, а менші (відповідні до високочастотної

частини) – наприкінці . В [8] робота методу CAVLC розглянута на наступному типовому векторі коефіцієнтів: 7 6 -2 0 -1 0 0 1 0 0 0 0 0 0 0.

Для передачі інформації про цей вектор використовується наступні дані.

1) Кількість ненульових коефіцієнтів (N) і кількість коефіцієнтів, рівних одиниці по абсолютній величині, що розташовуються після останнього ненульового елемента (T1). Для даного прикладу $T1 = 2$ (коефіцієнти -1 і 1), а $N = 5$ (коефіцієнти 7, 6, -2, -1, 1).

2) Закодовані значення коефіцієнтів. При цьому для ненульових елементів, рівних по абсолютній величині одиниці, необхідно зберігати тільки їх знак. Коефіцієнти кодуються у зворотному порядку, тобто першим закодованим коефіцієнтом у цьому випадку буде -2. Для кодування використовуються 6 таблиць VLC (Variable Length Codes, коди змінної довжини) Голомба. Для кодування першого коефіцієнта використовується стартова таблиця. Контекстна адаптивність алгоритму полягає в можливості перемінити таблицю для кодування кожного наступного коефіцієнта залежно від попередніх коефіцієнтів.

3) Інформація про знаки. Для кодування знака використовується один біт. Для кодування одиничних коефіцієнтів крім цього біта ніякої іншої інформації не потрібно, а для інших коефіцієнтів знаковий біт включається в коди Голомба. Місцезнаходження кожного ненульового коефіцієнта кодується шляхом вказівки позицій нулів перед останнім ненульовим коефіцієнтом. Ця інформація розбивається на 2 частині:

3.1) Загальна кількість нулів. Це число визначає кількість нулів між останнім ненульовим коефіцієнтом вектора і його початковим елементом. Для даного прикладу це число рівне 3. Тому що вже відомо, що число ненульових коефіцієнтів у векторі (N) рівно 5, те це число повинне бути в межах [0,11]. Для N у межах від 1 до 15 доступно 15 таблиць. N, рівне 16, означає, що у векторі немає нульових коефіцієнтів.

3.2) Розташування нулів у векторі. У даному прикладі потрібно позначити положення 3 нулів. Спочатку кодується кількість нулів перед останнім ненульовим коефіцієнтом (у прикладі - 2). Це число повинне перебуває в межах $[0,3]$, тому використовується відповідна таблиця VLC. Залишилося закодувати положення останнього нуля. Кількість нулів перед передостаннім ненульовим коефіцієнтом повинне бути в межах $[0,1]$. У прикладі це число рівне 1. Більше нулів ні, тому кодування закінчується.

Ефективність кодування може бути підвищена шляхом використання контекстно-адаптивного двійкового арифметичного кодування (САВАС). Використання арифметичного кодування допускає використання нецілої кількості біт на символ алфавіту, що особливо добре для символів з імовірністю більше 0,5. Важливою особливістю САВАС є контекстне моделювання. Статистичні дані про вже закодовані символи використовуються для оцінки ймовірностей символів. Ці дані використовуються для перемикання між декількома моделями для оцінки ймовірності. В H.264/AVC арифметичний кодер являє собою сукупність процедур низької складності, у яких відсутні операції множення. Процедури містять у собі зрушення й звертання до таблиць. Використання САВАС дозволяє зменшити в середньому битрейт на 10--15%. Найбільший вигреш виходить, звичайно, при обробці чересстрочних сигналів ТВ. Порівняння стандартів стиснення показано в таблиці 1.1.

Таблиця 1.1.

Порівняння стандартів стиснення й використання у них алгоритмів стиснення

<i>Назва станда рту</i>	<i>Область застосування</i>	<i>Первинний алгоритм</i>	<i>Вторинний алгоритм</i>	<i>Дозволу</i>	<i>Потік</i>	<i>Недоліки та переваги</i>

H.261	ISDN відеоконференції, апаратні кодеки	DCT (блоки 8x8) і квантування міжкадрової різниці	VLC (метод Хаффмана)	352x288x30 176x144x30	0,04-2 Мбіт/с (рх64 Кбит/с, де р від 1 до 30)	<i>Переваги:</i> простий в апп. реалізації. <i>Недоліки:</i> низький ступінь стиснення, погана компенсація руху (алгоритм розрахований на відео з невеликою кількістю руху, напр. голова співрозмовника)
H.263	Для відеоконференцій у мережах АТМ і по широкосмугових каналах	DCT (блоки 8x8) і квантування	Арифметичне кодування	Sub-QCIF, QCIF, CIF, 4CIF, 16CIF, дозволи, що окремо настраюються	0,04-2 Мбіт/с (рх64 Кбит/с, де р від 1 до 30)	<i>Переваги:</i> поліпшений алгоритм компенсації руху в порівнянні з H.261, більш ефективний вторинний алгоритм. <i>Недоліки:</i> перебуває між MPEG-2 і MPEG-4 по кількості закладених ідей
ISO MPEG-1	Зберігання відеопослідовностей на CD-ROM	DCT (блоки 8x8) і квантування	Коди зміною довжини (Хаффман)	352x240x30 352x288x25	1,5 Мбіт/с	<i>Переваги:</i> простий в апп. реалізації. <i>Недоліки:</i> низький ступінь стиснення,

						недостатня гнучкість формату.
ISO MPEG-2	Зберігання на DVD			Універсальний	3-15 Мбіт/с	<i>Переваги:</i> порівняльна простота реалізації. <i>Недоліки:</i> недостатній ступінь стиснення, мала гнучкість.
MPEG-4 Part 10 (AVC, H.264)	Кодування окремих аудіовізуальних об'єктів, як природнього походження, так і синтезованого	Цілісне DCT (блоки 4x4), або DWT і квантування міжкадрової різниці	CAVLC чи CABAC	Універсальний	0,0048-20 Мбіт/с	<i>Переваги:</i> високий ступінь універсальності та об'єктноорієнтована робота з потоком даних. <i>Недоліки:</i> висока складність реалізації

MPEG-4, будучи останнім за часом появи стандартом, використовує найбільш ефективні алгоритми стиснення даних, наприклад, алгоритм арифметичного кодування для ентропійного кодування й вейвлет-перетворення в якості первичного алгоритму. Найбільш перспективними зараз представляються наступні напрямки розвитку: використання альтернативних методів стиснення без втрат, таких як різні реалізації арифметичного кодування (напр., CABAC), ланцюгові коди й т.п.; використання альтернативних

перетворень – вейвлет-перетворень (DWT), фрактальних перетворень і виділення об'єктів з їхнім наступним описом.

1.5. Технології стиснення цифрового відео

Існує безліч технологій стиснення цифрового відео, однак зупинимось на тих з них, які лягли в основу найбільш популярних компресорів відео. Деякі з розглянутих компресорів використовують не одну технологію стиснення, а деяку їхню сукупність. Наприклад, і Indeo 3.2, і Сінерак використовують векторну квантизацію. Міжнародні стандарти MPEG-1, MPEG-2, MPEG-4, H.261 і H.263 використовують комбіновану технологію БДКП і компенсацію руху. Деякі сучасні алгоритми використовують технологію ДВП (Discrete Wavelet Transform, або DWT). Інші технології включають фрактальное стиснення зображень (Fractal Image Compression).

Для початку давайте визначимо поняття «якість стиснення відео». Незважаючи на всю суб'єктивність такого поняття, все-таки спробуємо умовно розділити якість стиснення на наступні рівні (ступені стиснення).

Стиснення без втрат якості.

Стиснення зображень може здійснюватися без втрат якості лише в тому випадку, якщо в процесі стиснення не було втрат даних. У результаті отримане після декомпресії зображення буде в точності (побитно) збігатися з оригіналом. Прикладом такого стиснення може служити формат GIF для статичної графіки й GIF89а для відео.

Стиснення із втратами якості.

Стиснення може відбуватися із втратами якості, якщо в процесі стиснення інформація була загублена. Однак з погляду людського сприйняття стисненням із втратами слід уважати лише такий стиснення, при якому

можливо на око відрізнити результат стиснення від оригіналу. Таким чином, незважаючи на те що два зображення - оригінал і результат стиснення з використанням того або іншого компресора - побитно можуть не збігатися, проте різниця між ними може бути зовсім непомітною. Прикладом може служити алгоритм JPEG для стиснення статичної графіки й алгоритм М-JPEG для стиснення відео.

Стиснення без втрат з погляду сприйняття.

Формально будучи стисненням із втратами якості, схема стиснення може в той же час видатися стисненням без втрат з погляду сприйняття її людиною. Більшість технологій стиснення з формальною втратою якості мають так званий Фактор Якості Стиснення (ФКС), що характеризує саме сприйману сторону якості, що й варіюється в межах від 0 до 100. При факторі якості стиснення рівному 100 сприймані характеристики якості стислого відео неотличимі від оригіналу.

MPEG і інші технології стиснення із втратою якості іноді стисненняють, без втрат переступаючи за грань стиснення з погляду сприйняття відеоінформації. Проте стислі відео й статичні зображення цілком прийнятні для адекватного сприйняття їх людиною. Іншими словами, у цьому випадку спостерігається так звана природня деградація зображення, при якій губляться деякі дрібні деталі сцени. Схоже може відбуватися й у природніх умовах, наприклад при дощі або тумані. Зображення в таких умовах, як правило, помітне, однак деталізація його зменшується.

Низька якість стиснення, у значній мірі, що спотворює зображення, що й вносить у нього штучні (не існуючі в оригіналі) деталі сцени, називається неприродним стисненням із втратою якості. Прикладом тому може служити деяка «блочність» у сильно стислому Mpeg-E й в інших компресорах, що використовують технологію БДКП. Неприродність полягає в першу чергу в порушенні найважливіших з погляду сприйняття людиною характеристик

зображення – контурів. Досвід показує, що саме контури дозволяють сприймаючому апарату людини правильно ідентифікувати той або інший візуальний об'єкт.

Отже, відзначимо, що всі широко використовувані відеокомпресори використовують технології стиснення із втратами якості. При досить високих коефіцієнтах стиснення всі вони будуть стисненняти з неприродною втратою якості. Таким чином, вибираючи той або інший компресор для стиснення цифрового відео, необхідно досягти стиснення, принаймні із природніми втратами якості.

РОЗДІЛ 2

АНАЛІЗ МЕТОДОЛОГІЧНОЇ БАЗИ РОЗРОБКИ АЛГОРИТМУ СТИСНЕННЯ ВІДЕО З УРАХУВАННЯМ ПРОЦЕСОРНИХ ВТРАТ

2.1. Класифікація методів стиснення цифрового відео

Число різноманітних форматів цифрового відео величезне.

AVI (Audio Video Interlive)

Avi-файли - особливий випадок файлів RIFF. (скорочено від Resource Interchange File Format). Цей формат, що споконвічно призначав для обміну мультимедійними даними, був Microsoft разом з IBM. Даний формат є найпоширенішою формою вистави відео на персональних комп'ютерах. Залежно від форми вистави відеоданих файли AVI бувають різних стандартів.

Стандарти Avi- Файлов

Відеодані в AVI можуть бути сформовані й стиснуті різними способами. Наприклад, Video for Windows 1.1поставляється зі стандартним набором компресорів. А саме:

- Intel Indeo (version 3.2);
- Microsoft Video 1;
- Microsoft RLE (Run Length Encoding);
- Cinepak;

Для кожного стандарту передбачений так званий 4- символний код The Four Character Code (FOURCC). Аматорам стандартів і екзотики пропонуємо відвідати сторінку в Інтернеті FOURCC Microsoft за посиланням: <http://www.microsoft.com/hwdev/devdes/fourcc.htm>.

Можна тримати Avi-файли у формі незжатих кадрів. Для цього не потрібно ніяких компресорів.

4-символьний код цього стандарту - 'DIB', (Microsoft Device Independent Bitmap).

Проте існує принаймні ще три FOURCC для незжатих Avi-файлів:

'RGB '; 'RAW '; 0x00000000 (FOURCC, чие шістнадцятиричне значення рівне 0).

Кодування змінної довжини (MRLE)

Особливістю даної схеми кодування відео зображень є те, що MRLE здатна кодувати відео тільки в 8-бітній палітрі (256 кольорів). Вона не підтримує так звані режими "High Color" (216 кольорів) або "True Color" (224 або 232 кольорів).

Intel Real Time Video 2.1 (Indeo 2.1) (RT21)

Включений до складу стандартної поставки Microsoft Video For Windows 1.1.

Intel Indeo 3.1/3.2 (IV31 і IV32)

Особливістю методу є використання алгоритму векторної квантизації зображень (див. нижче).

Intel® Indeo® Media Kit 5.0

Cinepak (CVID)

Один з найпоширеніших і використовуваних компресорів Video for Windows. Забезпечує найбільш швидке відтворення відео. На відміну від Indeo 32, яка забезпечує ледве краща якість, однак помітно обтяжує процесор при декомпресії, Cinepak максимально розвантажує процесорний ресурс.

На сьогоднішній день існує принаймні три стандарти для Cinepak.

- Cinepak Supermac (споконвічний, 16- бітний компресор);
- Cinepak Radius (новий, поліпшений 16- бітний компресор);
- Cinepak Radius (32-бітна версія Radius Cinepak, що поставляється

разом з Windows 95).

Особливістю методу є використання алгоритму векторної квантизації зображень разом з алгоритмом різниці кадрів (див. нижче).

MPEG (Motion Pictures Expert Group). Один з найбільш популярних форматів вистави цифрового відео. У цей час існують три його специфікації: MPEG-1, MPEG-2 і MPEG-4. Незважаючи на більшу гнучкість стандарту, що дозволяє для різних додатків міняти в широкий межах значення більшості його параметрів (такі як розширення зображень, аспектне відношення, частота кадрів), розроблювачі його спочатку були орієнтовані на використання як основний носія кодової інформації компакт дисків (CD-ROM) зі швидкістю передачі даних 50 Кбіт/с. У результаті базовий алгоритм (MPEG 1) обмежує швидкість передачі діапазоном 150-225 Кбіт/с, розширення зображень (кадрів) як 352*288(Pal) або 320*240 (NTSC), частоту їх зміни 25 (PAL) або 30 (NTSC). Далі для простоти обмежимося розглядом Pal-системи.

Mpeg-Компресія використовує наступні основні ідеї:

1. Усунення тимчасової надмірності відео, що враховує той факт, що в межах коротких інтервалів часу більшість фрагментів сцени виявляються нерухливими або незначно зміщаються по полю.
2. Усунення просторової надмірності зображень придушенням дрібних деталей сцени, несуттєвих для візуального сприйняття людиною;
3. Використання більш низького колірною дозволу при uiv- виставі зображень (u – яскравість, i і v – сигнали різниці кольорів) – установлене, що око менш чутливе до просторових змін відтінків кольору в порівнянні зі змінами яскравості.
4. Підвищення інформаційної щільності результуючого цифрового потоку шляхом вибору оптимального математичного коду для його опису (наприклад, використання більш коротких кодових слів для найбільше часто повторюваних значень).

Зображення в Mpeg-послідовності підрозділяються на наступні типи:

- I (intra), що відіграють роль опорних при відновленні інших зображень по їхніх різницях;
- P (predicted), що містять різницю поточного зображення з попереднім I або P з урахуванням зсувів окремих фрагментів;
- B (bidirectionally predicted), що містять різницю поточного зображення з попереднім і наступним зображеннями типів I або P з урахуванням зсувів окремих фрагментів.

Зображення поєднуються в групи (GOP – Group Of Pictures), що представляють собою мінімальний повторюваний набір послідовних зображень, які можуть бути декодовані незалежно від інших зображень у послідовності. Типової є група виду (I0 B1 B2 P3 B4 B5 P6 B7 B8 P9 B10 B11) (I12 B13 B14 P15 B16 B17 P18...), у якій I тип повторюється кожні полсекунди. Оборотною увагою, що в зображенні P3 основна частина фрагментів сцени передвіщається на підставі відповідних зміщених фрагментів зображення I0. Властиво кодуванню зазнають тільки різниці цих пар фрагментів. Аналогічно P6 «будується» на базі P3, P9 – на базі P6 і т.д. У той же час більшість фрагментів B1 і B2 передвіщаються як напівсума зміщених фрагментів з I0 і P3, B4 і B5 – з P3 і P6, B7 і B8 – з P6 і P9 і т.д. Поряд із цим B-зображення не використовуються для пророкування ніяких інших зображень. У силу залежності зображень у процесі їх кодування міняється порядок проходження. Для вищенаведеної послідовності він буде наступним: I0 P3 B1 B2 P6 B4 B5 P9 B7 B8 I12 B10 B11 P15 B13 B14 P18 B16 B17....

Ясно, що точність кодування повинна бути максимальної для I, нижче – для P, мінімальної – для B. Установлене, що для типових сцен гарні результати досягаються при відведенні числа біт для I в 3 рази більше, чим для P, і для P в 2-5 раз більше, чим для B. Ці відносини зменшуються для динамічних сцен і збільшуються для статичних.

Окремі зображення складаються з макроблоків. Макроблок - це основна структурна одиниця фрагментації зображення. Він відповідає ділянці зображення розміром 16x16 пікселів. Саме для них визначаються вектора зсуву відносно I- I- або P- зображень. Загальне число макроблоків у зображенні – 396. Для підвищення стійкості процесу відновлення зображень до можливих помилок передачі даних послідовні макроблоки поєднують у незалежні один від одного розділи (slices), максимальним числом 396. У граничному випадку «чистої» передачі на зображення доводиться всього один розділ з 396 макроблоків. У свою чергу кожний макроблок складається із шести блоків, чотири з яких несуть інформацію про яскравість Y, а по одному визначають колірні U- і V-компоненти. Кожний блок являє собою матрицю 8*8 елементів. Блоки є базовими структурними одиницями, над якими здійснюються основні операції кодування, у тому числі виконується дискретне косинусне перетворення (DCT – Discrete Cosine Transform) і квантування отриманих коефіцієнтів.

Motion JPEG или M-JPEG (MJPEG). Більшість настільних систем відеозаписи (video capture) і відеоредагування (video editing) використовують саме цей метод при записі відео в Avi-файли. В Motion JPEG кожний відеокадр зображення стисненняється окремо з використанням стандарту JPEG. Ніяких інших додаткових алгоритмів при цьому не використовується. Безумовною гідністю цього методу є можливість редагування відео без втрат якості, тому що кадри є незалежними. Цим, по суті справи, і визначається використання даного методу саме як механізму зберігання відео, що служить для його редагування, а не для поширення. Motion JPEG використовує алгоритм блокового ДКП (Block Discrete Cosine Transform (DCT)) для стиснення зображень.

Editable MPEG (XMPG. MPGI). Editable MPEG, так само як і M-JPEG, використовується для редагування цифрового відео являє собою Avi-файл, що полягає тільки з кадрів MPEG типу Й. Однак усі інші механізми стиснення

MPEG отут задіяні. Входить у стандартну поставку Microsoft Video for Windows 1.1. і використовується такою настільною системою редагування цифрового відео, як, наприклад, Adobe Premiere.

Хоча перераховані вище компресори досить популярні, проте це далеко не всі стандарти стиснення Avi-файлів. Характеризуючи цю групу компресорів, можна відзначити, що вони проектували й створювалися в першу чергу як засобу стиснення відео- і аудіоданих компакт-дисках, що зберігаються на жорстких дисках і, а це, у свою чергу, свідчить про їхні невеликі можливості при стиснення й відносно високій якості при відтворенні.

Із приходом Інтернету все більшу популярність одержують методи й засобу стиснення відео- і аудіоданих, що дозволяють, застосовуючи передові технології (sophisticated motion estimation and compensation, wavelets, fractals і інші), досягти найменших співвідношень «кілобіт/секунда», що дозволили проводити, наприклад, сеанси відеоконференцій засобами Інтернету. Ясно, що такі методи стиснення забезпечують суттєво більший ступінь стиснення, при відносно низькій якості.

Покоління компресорів – H.XXX. Зовсім недавно була офіційно зареєстрована нова серія компресорів цифрового відео, що визначає тенденції розвитку механізмів стиснення відео. Деякі компресори сімейства H.XXX, наприклад H.261, досить популярні, інші маловідомі й використовують такі передові й поліпшені технології, як wavelets.

Відмінною рисою компресорів сімейства H.XXX є їхня націленість на зменшення потоку цифрового відео через Інтернет, що природно приводить до відходу фактора якості на другий план. Деякі продукти сімейства H.XXX Microsoft удосконалює, а деякі вже входять до складу таких пакетів телеконференцій, як Netshow і Netmeeting.

Vdowave або Vdolive від Vdonet (VDOM, VDOW)

Vdonet випускає wavelet – відеокомпресор, включений у комплект реалізації 32- бітної версії Video for Windows. Microsoft використовує Vdowave як частина Netshow. У цей час існують дві версії компресора Vdowave:

- Vdowave 2.0 is a fixed rate video codec.
- Vdowave 3.0 is a "scalable" video codec.

Стандартний набір Netshow 2.0 установлює тільки декодер Vdowave. Засіб розробки Netshow 2.0 установлює як кодер, так і декодер Vdonet Vdowave. По деяких тестах, Vdowave суттєво перевершує по стисненню MPEG-1 і інші компресори, що базуються на алгоритмі блокового ДКП (block Discrete Cosine Transform), але лише при низьких відносинах «кілобіт/секунда».

MPEG-4 (MPG4).

Microsoft's Netshow 2.0 установлює компресор MPEG-4, що є новим міжнародним стандартом, який, однак, офіційно ще не визнаний.

Продуктивність компресорів AVI. Для перевірки продуктивності різних компресорів була зроблена серія експериментів на десятисекундному відеофрагменті, записаному при 30 кадрах у секунду, 320x240, у колірній палітрі 24 біта на пиксел зображення.

Далі даний відеофрагмент послідовно стисненнявся різними відеокомпресорами засобами Microsoft Video for Windows 1.1. За винятком відзначених випадків, отримана якість істотна не відрізнялося від оригінального.

Умовні позначки: ФКС, що коливається в інтервалі 0.. 100 фактор, що й наструюється, якості стиснення. ТС: технологія, що використовується при стиснення. КК Х: ключовий кадр через кожні Х.

2.2. Основні алгоритми стиснення відео

Відео – це сукупність зображень, зроблених у часі. Тому в більшості випадків різниця між сусідніми зображеннями невелика. Стиснення відео техніки використовують перевагу повторення частин малюнка з одного зображення до іншого, концентруючись на змінах між сусідніми зображеннями. Іншими словами, у відеокадрах є велика надмірність. Існує два типи надмірності:

- просторова надмірність: піксельна або спектральна кореляція в межах одного кадру;
- часова надмірність: подібність двох або більше різних кадрів.
- статистична: нерівномірний розподіл даних.

Стиснення відео на основі компенсації руху.

Алгоритм стиснення відео MPEG спирається на дві основні техніки: компенсація руху для зменшення тимчасової надмірності та перетворення стиснення на основі домену (DCT) для зменшення просторової надмірності. Техніка з компенсацією руху – техніка, яка використовує часову надмірність відеосигналів. Концепція компенсації руху базується на оцінці рух між відеокадрами, тобто якщо всі елементи у відеосцені приблизно просторово зміщений, рух між кадрами може бути описаний обмеженою кількістю параметрів руху (за допомогою векторів руху для поступального руху пікселів). Залишковий сигнал (помилка передбачення) додатково стискається з просторовою надмірністю зменшення (DCT). Інформація щодо руху базується на 16x16 блоках і передається разом з просторовою інформацією. Інформація про рух стискається за допомогою кодів змінної довжини для досягнення максимальної ефективності.

Алгоритми стиснення відео поділяють багато методів стиснення, що використовуються при стисненні нерухомих зображень. Однак ключова відмінність полягає в тому, що стиснення відео може скористатися подібністю

між послідовними відеокадрами для досягнення ще кращих коефіцієнтів стиснення. Використовуючи вищезгадані методи, алгоритми стиснення нерухомих зображень, такі як JPEG, можуть досягти гарної якості зображення при ступені стиснення 10:1. Найдосконаліші кодеки нерухомих зображень можуть досягти гарної якості зображення при ступенях стиснення до 30:1. На відміну від цього, алгоритми стиснення відео можуть забезпечити хорошу якість відео у співвідношенні до 200:1. Ця підвищена ефективність досягається додаванням специфічних для відео методів стиснення, таких як оцінка руху та компенсація руху.



Рис.2.1. Компенсація руху

Для кожного блоку макросів у поточному кадрі оцінка руху намагається знайти область у попередньо закодованому кадрі (так званий «опорний кадр»). Просторове зміщення між поточним блоком та вибраним блоком з опорного кадру називається «вектором руху» (рис. 2.1.). Кодер обчислює різницю в пікселях між виділеним блоком з опорного кадру та поточним блоком і передає

цю «помилку передбачення» разом із вектором руху. Більшість стандартів стиснення відео дозволяють обходити прогнозування на основі руху, якщо кодеру не вдається знайти відповідність блоку макросів. У цьому випадку замість помилки передбачення кодується сам блок макросів.

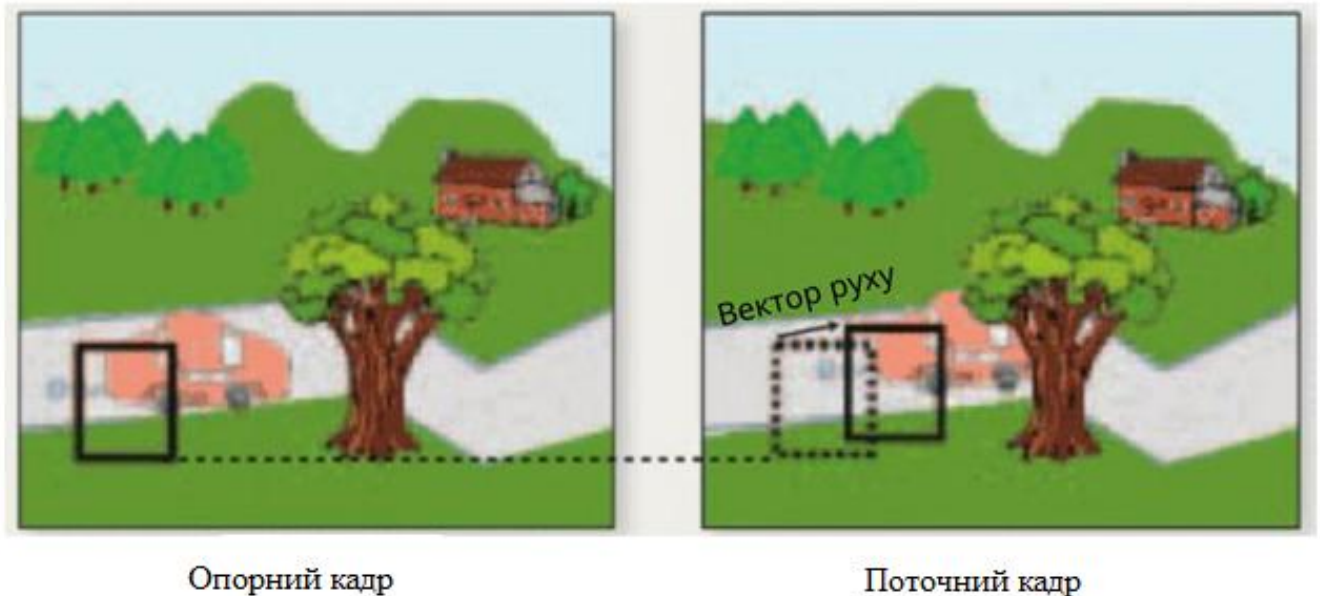


Рис.2.2. Опорні кадри

Зверніть увагу, що контрольний (опорний) кадр не завжди є попередньо відображеним кадром у послідовність відеокадрів. Алгоритми стиснення відео зазвичай кодують кадри в відрізняються від порядку, в якому вони відображаються. Кодер може пропустити кілька кадрів вперед і кодувати майбутній відеокадр, а потім перейде назад і кодує наступний кадр у послідовності. Це здійснюється для того, щоб можна було виконати оцінку руху назад у часі, використовуючи закодований майбутній кадр як опорний кадр.

Відео-алгоритми стиснення також дозволяють використовувати два опорні кадри – один попередньо відображений кадр і один раніше закодований майбутній кадр. Алгоритми стиснення відео періодично кодують один відеокадр, використовуючи лише методи кодування нерухомого зображення, не

покладаючись на раніше закодовані кадри. Ці кадри називаються «внутрішні кадри» або «I-кадри». Якщо кадр у стислому потоці бітів пошкоджений через помилки, відеодекодер може «перезапустити» при наступному I-кадрі, який не потребує опорного кадру для реконструкції.

Як показано на рисунку 2.3, кадри, кодовані з використанням лише попередньо відображеного опорного кадру, називаються «P-кадрами», а кадри, кодовані як з використанням майбутніх, так і раніше відображених опорних кадрів, називаються «B-кадрами». Типова послідовність кадрів проілюстрована на рисунку 2.3.

Через важливість довільного доступу до збереженого відео та значне зменшення швидкості передачі даних за рахунок інтерполяції з компенсацією руху, у MPEG визначено чотири типи кадрів:

- внутрішні кадри (I-кадри);
- передбачені кадри (P-кадри);
- інтерпольовані кадри (B-frames);
- DC-кадри (D-кадри).

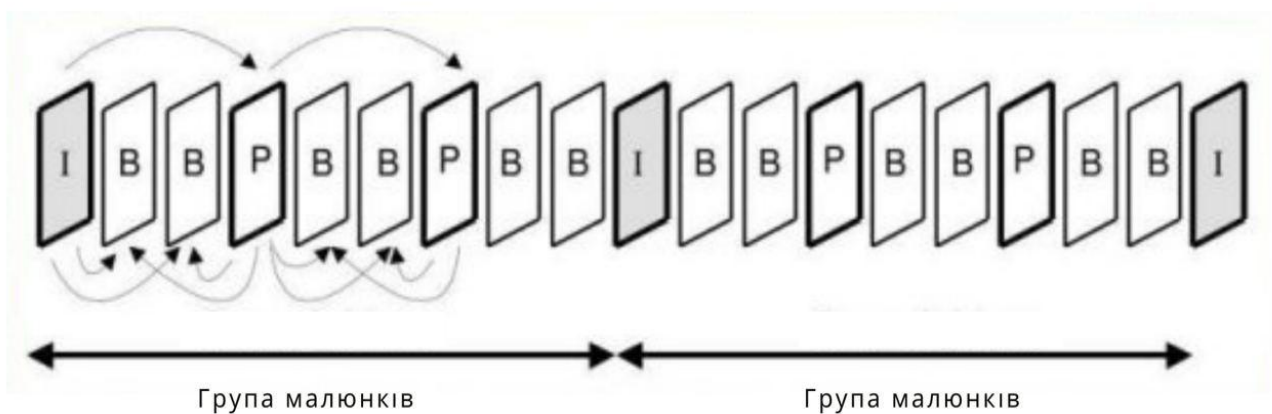


Рис.2.3. Типи кадрів у MPEG

I-кадри.

I-кадри (внутрішньокодовані кадри) кодуються незалежно, не посилаючись на інші кадри. I-кадри надають точки довільного доступу в стиснених відеоданих, оскільки I-кадри можуть бути декодовані самостійно, не посилаючись на інші кадри. З I-кадрами, бітовий потік MPEG є більш редагованим. Крім того, розповсюдження помилок через помилки передачі в попередніх кадрах буде припинено I-кадром, тому що I-кадр не має посилання на попередні кадри. Оскільки I-кадри використовують лише кодування перетворення без передбачуваного кодування з компенсацією руху, воно забезпечує лише помірне стиснення.

P-кадри.

P-кадри (індикативно кодовані кадри) кодуються з використанням прогнозу з компенсацією руху вперед із попереднього I- або P-кадру. P-кадри забезпечують більше стиснення, ніж I-кадри завдяки передбаченню компенсації руху. Вони також служать еталонами для B-кадрів та майбутніх P-кадрів. Помилки передачі в I-кадрах та P-кадри можуть поширюватися на наступні кадри, оскільки I-кадри та P-кадри використовуються для прогнозування наступних кадрів.

B-кадри.

B-кадри (двонаправлені кодовані кадри) дозволяють кодувати макроблоки, використовуючи двонаправлене прогнозування з компенсацією руху як з попередніх, так і з майбутніх еталонних кадрів I-кадрів або P-кадрів. У B-кадрах кожен двонаправлений макроблок з компенсацією руху може мати два вектори руху: прямий вектор руху, який посилається на найкращий відповідний блок у попередніх I-кадрах або P-кадрах, і вектор руху назад, який посилається на найкращий відповідний блок у наступних I-кадрах або P-кадрах. Прогноз із компенсацією руху може бути сформований середнім значенням двох опорних блоків з компенсацією руху. Завдяки усередненню між попередніми та

майбутніми еталонними блоками, ефект шуму можна зменшити. В-кадри забезпечують найкраще стиснення порівняно з I- та P-кадрами – вони використовуються як опорні кадри для прогнозування V-кадрів. Щоб спростити структуру і, оскільки немає очевидної переваги використовувати Vframes для прогнозування інших V-кадрів, V-кадри не використовуються як опорні кадри. Отже, V-кадри не поширюють помилок.

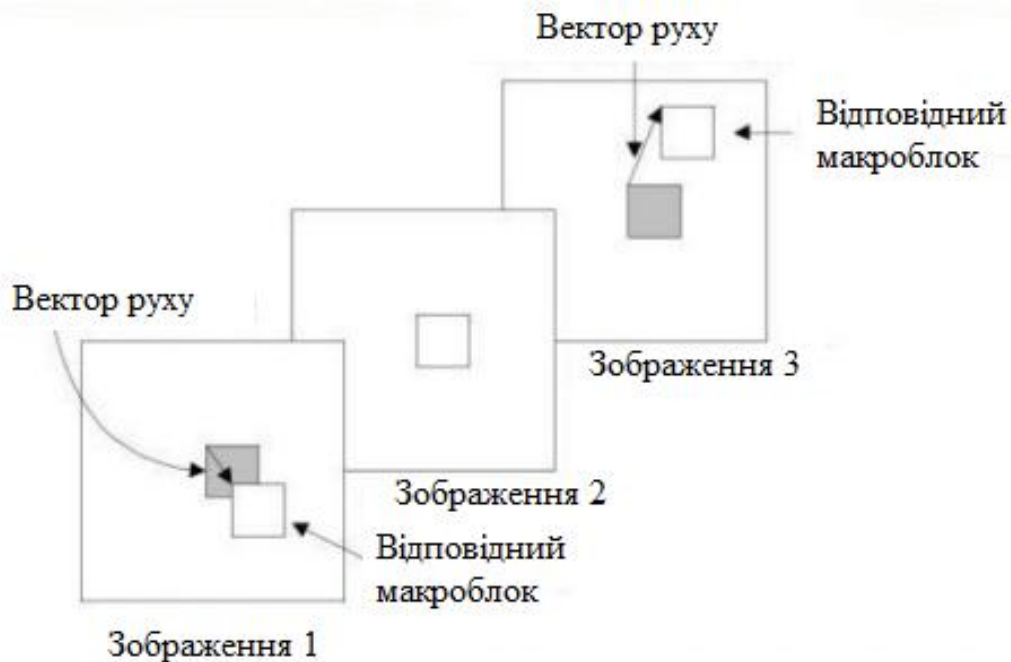


Рис.2.4. Оцінка двонаправленого руху

D-кадри.

D-кадри (DC-кадри) – це кадри з низькою роздільною здатністю, отримані шляхом декодування лише DC коефіцієнта косинусного перетворення кожного макроблока. Вони не використовуються в поєднанні з I-, P- або V-кадрами. D-кадри використовуються рідко, але визначені для швидкого пошуку на послідовних цифрових носіях інформації.

Одним із факторів, який ускладнює оцінку руху, є те, що зміщення об'єкта від опорного кадру до поточного кадру може бути нецілим числом пікселів. Наприклад, припустимо, що об'єкт перемістився на 22,5 пікселя вправо та на 17,25 пікселя вгору. Для вирішення таких ситуацій сучасні стандарти стиснення відео дозволяють векторам руху мати нецілі значення; часто роздільна здатність векторів руху половини або чверті пікселя. Для підтримки пошуку збігів при зміщеннях часткових пікселів, кодер повинен використовувати інтерполяцію для оцінки значень пікселів опорного кадру в нецілих місцях.

Найпростіший і найретельніший спосіб оцінки руху – це оцінка кожної можливої області 16×16 і вибору найкращої. Як правило, обчислення «суми абсолютних різниць» (SAD) або «суми квадратних різниць» (SSD) використовується для визначення того, наскільки близька область 16×16 кандидатів відповідає макроблоку. SAD або SSD часто обчислюється лише для площини яскравості, але може також включати площину кольоровості. Але такий підхід може бути надто вимогливим до процесорів: для вичерпного пошуку області 48×24 пікселів потрібно понад 8 мільярдів арифметичних операцій в секунду при роздільній здатності відео QVGA (640×480) та частоті кадрів 30 кадрів в секунду.

Через це велике обчислювальне навантаження практичні реалізації оцінки руху не використовують вичерпний пошук. Натомість алгоритми оцінки руху використовують різні методи для вибору обмеженої кількості перспективних кандидатів векторів руху (в більшості випадків приблизно від 10 до 100 векторів) та оцінки лише 16×16 областей, що відповідають цим векторам-кандидатам. Один із підходів полягає у виборі кандидатів векторів руху у кілька етапів. Наприклад, п'ять початкових векторів-кандидатів можуть бути обрані та оцінені. Результати використовуються для усунення малоймовірних частин області пошуку і до нульової точки в найбільш перспективній частині області пошуку.

Вибирається п'ять нових векторів і процес повторюється. Через кілька таких етапів вибирається найкращий вектор руху, знайдений на сьогодні.

Інший підхід аналізує вектори руху, раніше вибрані для оточуючих макроблоків у поточному та попередньому кадрах, намагаючись передбачити рух у поточному макроблоці. На основі цього аналізу вибирається кілька кандидатів векторів руху, і оцінюються лише ці вектори.

Вибравши невелику кількість кандидатів-векторів, замість того, щоб вичерпно сканувати область пошуку, обчислювальну потребу оцінки руху можна значно зменшити - іноді більш ніж на два порядки. Але існує компроміс між навантаженням на обробку та якістю зображення або ефективністю стиснення; загалом, пошук більшої кількості кандидатів векторів руху дозволяє кодеру знайти блок в опорному кадрі, який краще відповідає кожному блоку в поточному кадрі, зменшуючи тим самим помилку передбачення. Чим менша помилка предикації, тим менше бітів потрібно для кодування зображення. Отже, збільшення кількості кандидатів-векторів дозволяє зменшити стиснуту швидкість передачі даних за рахунок виконання більших обчислень SAD (або SSD). Або, як альтернатива, збільшення кількості кандидатів-векторів при утриманні константи стислої швидкості передачі дозволяє кодувати помилку передбачення з більшою точністю, покращуючи якість зображення.

Деякі кодеки (включаючи H.264) дозволяють розділити макроблок 16x16 на менші блоки (наприклад, різні комбінації блоків 8x8, 4x8, 8x4 та 4x4), щоб зменшити помилку передбачення. Кожен із цих менших блоків може мати свій власний вектор руху. Пошук оцінки руху для такої схеми починається з пошуку хорошого положення для всього блоку 16x16. Якщо матч досить близький, немає необхідності додатково розподіляти. Але якщо збіг поганий, тоді алгоритм починається з найкращої позиції, знайденої на даний момент, і далі поділяє вихідний блок на блоки 8x8. Для кожного блоку 8x8 алгоритм шукає найкращу позицію поблизу позиції, обраної пошуком 16x16.

Залежно від того, наскільки швидко знайдений хороший збіг, алгоритм може продовжувати процес, використовуючи менші блоки 8x4, 4x8 тощо.

Навіть за різкого зменшення кількості кандидатів векторів руху, оцінка руху є найбільш вимогливою до обчислень завданням у програмах стиснення відео. Включення оцінки руху робить відеокодування набагато більш вимогливим до обчислень, ніж декодування. Для оцінки руху може знадобитися до 80% циклів процесора, проведених у відеокодері. Тому багато процесорів, орієнтованих на мультимедійні програми, надають спеціалізовану інструкцію для прискорення обчислень SAD або спеціальний співпроцесор SAD для вивантаження цього завдання з центрального процесора. Наприклад, ядро

ARM11 ARM надає інструкцію для прискорення обчислення SAD, а деякі члени сімейства DSP Texas Instruments TMS320C55x забезпечують співпроцесор SAD.

Зверніть увагу, що для виконання оцінки руху кодер повинен зберігати в пам'яті один або два опорні кадри на додаток до поточного кадру. Необхідні буфери кадрів дуже часто перевищують наявну вбудовану пам'ять, що вимагає додаткових мікросхем пам'яті у багатьох додатках. Зберігання опорних кадрів у позачиповій пам'яті призводить до дуже високої пропускну здатності зовнішньої пам'яті в кодері, хоча великі кеші на мікросхемі можуть допомогти значно зменшити необхідну пропускну здатність.

2.3. Порівняльна характеристика алгоритмів стиснення відео та обґрунтування обраного методу

Тепер, проаналізувавши вищевказані алгоритми стиснення окремо, варто порівняти їх переваги, недоліки і області застосування (рис. 2.5).

<i>Алгоритм</i>	<i>Переваги</i>	<i>Недоліки</i>	<i>Застосування</i>
Кодування довжин серій	Простий алгоритм	Не підходить для стиснення реальних відеозаписів	Цифрові 8-бітні зображення, <u>цифрові анімації</u> з малою кількістю кольорів
Різниця кадрів	Краще стиснення, ніж незалежне стиснення окремих кадрів	Необхідність введення ключового кадру	Відео з <u>великою</u> кількістю повторюваних кадрів
Компенсація руху	Краще стиснення ніж у різниці кадрів	Кодування вельми трудомістке і вимагає спеціальну апаратуру	Відео з переміщенням одного об'єкту відносно нерухомого фону
Векторне квантування	Процес декодування дуже швидкий	Кодування вельми трудомістке і вимагає спеціальну апаратуру; Блочні спотворення при <u>сильному</u> стисненні	Відео високої якості, яке має небагато повторюваних елементів
Дискретне косинус перетворення	Краще стиснення, ніж при векторному квантуванні	Блочні спотворення при <u>сильному</u> стисненні; Закруглення гострих кутів зображення	Відео високої якості, яке має небагато повторюваних елементів

Рис. 2.5. Порівняльна характеристика методів стиснення відео

Секретні кодери.

На додаток до двох підходів, описаних вище, існує багато інших методів вибору відповідних варіантів векторів руху, включаючи широкий спектр власних рішень. Більшість стандартів стиснення відео визначають лише формат потоку стиснених відео-бітів та кроки декодування, а процес кодування залишають невизначеним, щоб кодери могли використовувати різні підходи до оцінки руху. Підхід до оцінки руху є найбільшим диференціатором серед реалізацій відеокoderів, які відповідають загальноприйнятому стандарту. Вибір методики оцінки руху суттєво впливає на обчислювальні вимоги та якість відео; тому деталі підходу до оцінки руху в комерційних кодерах часто пильно охороняються комерційною таємницею.

Компенсація руху.

У відеодекодері компенсація руху використовує вектори руху, кодовані у стислому бітовому потоці, для прогнозування пікселів у кожному блоці макросів. Якщо горизонтальна та вертикальна складові вектора руху є цілочисельними значеннями, то передбачений макроблок є просто копією 16-піксельної на 16-піксельної області опорного кадру. Якщо будь-який компонент вектора руху має неціле значення, для оцінки зображення використовується неціле піксельне розташування. Далі помилка передбачення декодується і додається до передбачуваного блоку макросів для реконструкції фактичних пікселів блоку макросів. Як згадувалося раніше, макроблок 16x16 може бути розділений на менші ділянки з незалежними векторами руху.

Порівняно з оцінкою руху, компенсація руху є набагато менш вимогливою до обчислень. Хоча оцінка руху повинна виконувати SAD або SSD-обчислення на кількості 16-піксельних на 16-піксельних областей, намагаючись знайти найкращий збіг для кожного макроблоку, компенсація руху просто копіює або інтерполює одну таку область для кожного макроблоку. Тим не менш, компенсація руху може споживати до 40% циклів процесора у відеодекодері,

хоча це число сильно варіюється в залежності від вмісту відеопослідовності, стандарту стиснення відео та реалізації декодера. Наприклад, робоче навантаження з компенсацією руху може складати лише 5% циклів процесора, проведених у декодері для кадру, який мало використовує інтерполяцію.

Як і оцінка руху, компенсація руху вимагає від відеодекодера збереження в пам'яті одного або двох опорних кадрів, для цього часто потрібні зовнішні мікросхеми пам'яті. Однак компенсація руху робить менше звернень до буферів опорного кадру, ніж оцінка руху. Отже, вимоги до пропускної здатності пам'яті менш жорсткі для компенсації руху в порівнянні з оцінкою руху, хоча висока пропускна здатність пам'яті все-таки бажана для кращої продуктивності процесора.

Полірування зображення: розблокування та затухання.

В ідеалі алгоритми стиснення зображень та відео з втратою відкидають лише незначну для сприйняття інформацію, так що для людського ока реконструйоване зображення або відеопослідовність здається ідентичним оригінальному нестисненому зображенню чи відео. Однак на практиці деякі артефакти можуть бути видимими. Це може статися через погану реалізацію кодера, відеовмісту, кодування якого є особливо складним, або вибраної бітової швидкості, яка є занадто низькою для відеопослідовності, роздільної здатності та частоти кадрів. Останній випадок є особливо розповсюдженим, оскільки багато програм обмінюють якість відео для зменшення вимог до зберігання та / або пропускної здатності.

Два типи артефактів – «блокування» та «дзвінок» – поширені в програмах стиснення відео. Блокування артефактів пов'язано з тим, що алгоритми стиснення ділять кожен кадр на блоки 8x8. Кожен блок реконструюється з невеликими помилками, і помилки на краях блоку часто контрастують з помилками на краях сусідніх блоків, роблячи видимими межі блоків. На відміну від цього, дзвінки артефакти виглядають як спотворення по краях елементів

зображення. Артефакти дзвінка зумовлені тим, що кодер відкидає занадто багато інформації при квантуванні високочастотних коефіцієнтів DCT.

Програми для стиснення відео часто використовують фільтри після декомпресії, щоб зменшити артефакти блокування та дзвінка. Ці етапи фільтрації відомі як «деблокування» та «усунення» відповідно. І для розблокування, і для зменшення використовуються низькочастотні фільтри FIR (кінцева імпульсна характеристика), щоб приховати ці видимі артефакти.

Деблокувальні та деринг-фільтри досить вимогливі до обчислень. У поєднанні ці фільтри можуть споживати більше циклів процесора, ніж сам відеодекодер. Наприклад, декодер MPEG-4 простого профілю, рівень 1 (176x144 пікселі, 15 кадрів в секунду), оптимізований для ядра процесора загального призначення ARM9E, вимагає, щоб процесор працював із швидкістю циклу інструкцій близько 14 МГц при декодуванні помірно складного відеопотік. Якщо додається деблокування, процесор повинен працювати на частоті 33 МГц. Якщо одночасно додано дедеринг та деблокування, процесор повинен працювати на частоті приблизно 39 МГц – майже втричі більше тактової частоти, необхідної лише для алгоритму декомпресії відео.

Пост-обробка або в рядку.

Фільтри для деблокування та дерингування можуть застосовуватися до відеокадрів як окремий крок після обробки, який не залежить від декомпресії відео. Цей підхід надає розробникам систем гнучкість у виборі найкращих фільтрів розблокування та / або дедерингу для їх застосування або повністю відмови від цих фільтрів з метою зменшення обчислювальних вимог. При такому підході відеодекодер використовує кожен нефільтрований реконструйований кадр як еталонний кадр для декодування майбутніх відеокадрів, а додатковий буфер кадру потрібен для остаточного відфільтрованого відеовиходу.

Альтернативно, деблокування та / або деринг можуть бути інтегровані в алгоритм декомпресії відео. Цей підхід, який іноді називають "фільтрацією циклу", використовує відфільтрований реконструйований кадр як еталонний кадр для декодування майбутніх відеокадрів. Цей підхід вимагає, щоб відеокодер виконував ті самі етапи деблокування та / або стримуючого фільтрування, що і декодер, щоб зберегти кожен опорний кадр, що використовується в кодуванні, ідентичним тому, який використовується при декодуванні. Необхідність виконувати фільтрацію в кодері збільшує вимоги до продуктивності процесора для кодування, але може покращити якість зображення, особливо при дуже низьких бітових швидкостях. Крім того, додатковий буфер кадру, який необхідний при деблокуванні та / або дерингуванні, реалізовані як окремі етапи подальшої обробки, не потрібен, коли деблокування та деринг інтегровані в алгоритм декомпресії. Фігура 6 ілюструє деблокування та усунення, застосовувані "в циклі" або як крок після обробки. Наприклад, H.264 включає фільтр розблокування "в циклі", який іноді називають «фільтром циклу».

Перетворення кольорового простору.

Однією з ускладнень стиснення відеопотоків є те, що спосіб представлення кольору в кодах відрізняється від способу, яким його представляють відеокамери та дисплеї. Як зазначалося вище, алгоритми стиснення відео зазвичай представляють кольорові зображення з використанням площин яскравості та кольоровості. На відміну від них, відеокамери та дисплеї зазвичай поєднують червоне, зелене та синє світло для представлення різних кольорів. Отже, червоні, зелені та сині пікселі, зняті камерою, повинні бути перетворені у значення яскравості та кольоровості для кодування відео, а пікселі яскравості та кольоровості, що виводяться відеодекодером, повинні бути перетворені на певні рівні червоного, зеленого та синього для відображення. Алгоритм цього перетворення вимагає близько 12 арифметичних операцій на піксель

зображення, не включаючи інтерполяцію, необхідну для компенсації того факту, що площини кольоровості мають нижчу роздільну здатність, ніж площина яскравості, на вході та виході алгоритму стиснення відео. Для роздільної здатності VGA (640x480 пікселів) з роздільною здатністю 30 кадрів в секунду для перетворення (без будь-якої інтерполяції) потрібно понад 110 мільйонів операцій в секунду. При реалізації в програмному забезпеченні це обчислювальне навантаження може бути досить значним. Однак перетворення кольорів для відтворення часто підтримується апаратним забезпеченням дисплея, тому, можливо, це не потрібно робити в програмному забезпеченні.

РОЗДІЛ 3

ДОСЛІДЖЕННЯ КОДУВАННЯ ВІДЕО ЗА ДОПОМОГОЮ ШВИДКОГО ПЕРЕТВОРЕННЯ

3.1. Дискретне косинус не перетворення

На Землі зростає попит на всі види енергії, паливо, воду. Проблема раціонального використання ресурсів набуває все більшої актуальності, а її рішення стає стратегічним завданням кожного жителя Землі. У зв'язку зі збільшенням витрат енергоресурсів, зменшення їх запасів на Землі, постійно зростають тарифи на електроенергію та інші паливно-енергетичні ресурси. [17]

Весь ХХ ст. характеризувався швидким зростанням споживання первинних енергоресурсів і електричної енергії – сумарне світове споживання енергії збільшилася в 15 разів, а споживання на душу – в 4,4 рази. (Різниця обумовлено збільшенням чисельності населення від 1,6 до 7 млрд чоловік).

Перше десятиліття нового століття не внесло радикальних змін в «енергетичну картину» світу та в тенденції, що сформувалися. [18]

Споживання енергії продовжує зростати, незважаючи на економічні кризи, що періодично трапляються, і викликані ними короточасні зниження енергоспоживання (рис. 3.1). [19]

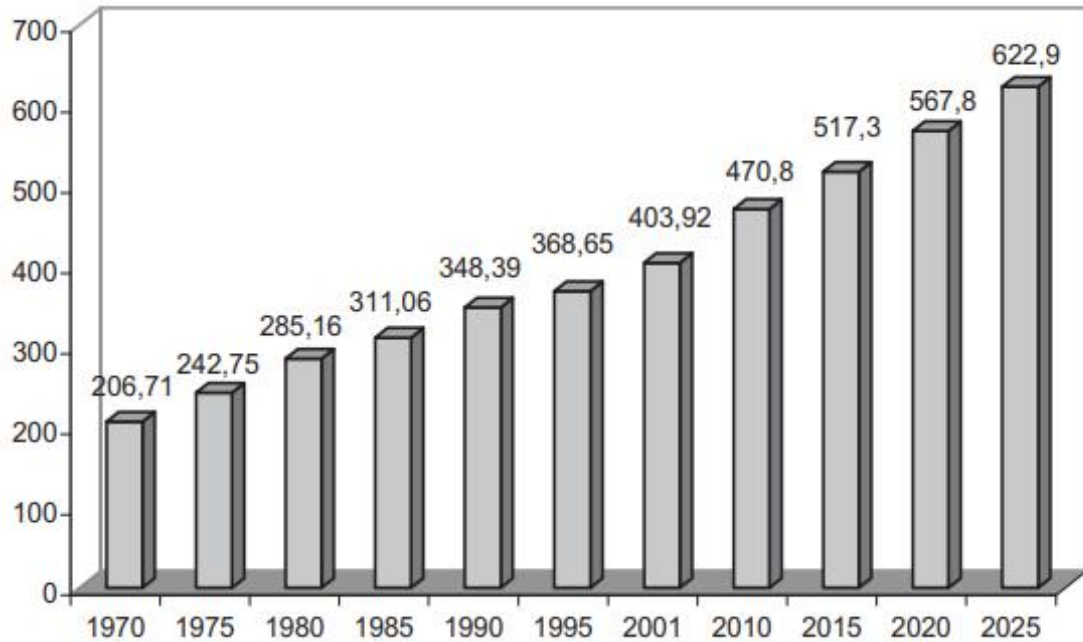


Рис. 3.1. Динаміка світового споживання первинної енергії (в британських теплових одиницях – BTU. 1 BTU = 252 кал.) за 1970-2025 рр.

Основні проблеми суспільства, пов'язані з енергетикою.

Можна говорити про «тріаді енергетичних проблем», в найбільшій мірі впливають на всі сторони життя людини і які зачіпають самі основи сталого розвитку цивілізації. Цю тріаду складають:

- дефіцит енергоресурсів та електроенергії (журналісти назвали цю проблему як «енергетичний голод»);
- загроза благополуччю навколишнього середовища внаслідок техногенного впливу об'єктів енергетики (загроза «екологічного інфаркту»);
- геополітичні та соціальні загрози

На основі аналізу «енергетичної картини» сучасного світу показано, що основним способом вирішення цих проблем є реалізація концепцій енергозбереження.

Сьогодні в сучасному світі енергозбереження – це невід'ємна частина життя цивілізованого суспільства. Це і турбота про здоров'я, і економія грошей, і комфорт проживання. Але одна з найголовніших (глобальних) характеристик енергозбереження – це захист навколишнього середовища від негативних впливів. Енергозбереження це технології і уклад життя, які допомагають нам зменшити споживання енергії за рахунок її раціонального використання. У більш вузькому сенсі, енергозбереження має на увазі ведення екологічно-дружнього способу життя за рахунок використання нових технологій, які допомагають заощадити наші з вами гроші і природні ресурси нашої планети, останнім часом так бездумно витрачаються людством, що в кінцевому підсумку може призвести до катастрофічних змін клімату . Коли ми зменшуємо кількість споживаної енергії, ми автоматично намагаємося знизити підвищення температури нашої атмосфери (цей процес відомий як глобальне потепління).

Енергозбереження в побуті, в кінцевому підсумку залежить від споживача. Населення на побутовому рівні може підтримати енергозбереження та зайнятися підвищенням енергоефективності в рамках окремо взятої квартири або будинку. На жаль, ощадливість не є характерною рисою більшості жителів на сьогодні в нашій країні та у світі загалом. Культура споживання енергії у населення дуже низька.

Економити ресурси в побуті – це означає з розумом витрачати електроенергію. Причому в більшості випадків така економія не викликає незручності, а навпаки, дає додаткові переваги у вигляді для власного або сімейного бюджету, а також економії енергоресурсів Землі (на глобальному рівні).

У своїй роботі ми намагалися знайти способи економії електроенергії: дотримуватися елементарних правил культури енергоспоживання, які не потребують великих витрат і спеціальних знань. Виходячи з цього, нашою

розробкою є додаток, який може бути використаним для автономних та стаціонарних пристроїв, таких як:

- мобільні пристрої,
- комп'ютери,
- ноутбуки,
- планшети тощо.

Як відомо, акумулятор – слабе місце більшості сучасних смартфонів. Виробники, з одного боку, борються за потужність процесора, розширення екрана та інші характеристики, що вимагають потужного акумулятора, а з іншого боку, зазвичай намагаються робити смартфони як можна тонше, в результаті чого в смартфонах використовуються невеликі акумулятори, чиєю ємності нерідко буває недостатньо навіть для того, щоб апарат простягнув до кінця одного робочого дня. Аналогічна проблема є у планшетів та ноутбуків.

Основними причинами цього є:

- перегляд високоякісних відео- та фото-матеріалів;
- спілкування за допомогою відеозв'язку;
- енергозатратні додатки та програми;
- зв'язок та бездротові мережі;
- робота ОС;
- постійний моніторинг свого будинку за технологією «розумний дім».

У зв'язку з пандемією, більшість населення України та інших країн у цей період перебували у статусі карантину та дотримувалися всіх норм та вимог. Через це була обмежена робота сфери послуг, навчальних закладів, деяких промислових підприємств та інших закладів, громадського транспорту, що стало основною причиною перебування вдома. Це спричинило збільшення використання мобільних телефонів та ноутбуків основними соціальними групами населення: дорослим потрібно виконувати свою роботу, а дітям та

студентам – навчатися. Тому основним способом комунікації став саме відеозв'язок – для спілкування, а відео- та фото-ресурси – для навчання. Проте такі інструменти взаємодії потребують великої кількості енергії електропристрою, що призводить до швидкого зниження заряду цього пристрою. Це стало ще одним обмеженням, тому що потрібно часто заряджати свої гаджети, розраховувати час дзвінка або перегляду відео, час прогулянки або походу до супермаркету.

В основі цього лежить дискретне косинусне перетворення. У відео зйомці, відео дзвінках та відеоспостереженні – незмінно присутня деяка надмірність. Якщо в межах поля знаходиться об'єкт достатнього розміру, всі пікселі, що представляють цей об'єкт, мають приблизно одні й ті ж значення яскравості і кольору. Це і є надмірність, що дозволяє зменшити кількість інформації в кожному окремому пікселі, певним чином описавши контури об'єкта і вказавши середні значення яскравості і кольору в межах цього контуру. Великі об'єкти відповідають низьким просторовим частотам, а дрібні – високим. На верхньому рівні ці частоти одночасно не присутні. У цифровому відеосигналі може бути переданий весь спектр просторових частот, однак якщо провести частотний аналіз зображення, то можна лишити в сигналі лише ті частоти, що дійсно в ньому присутні. Отже, важливим кроком в процесі стиснення зображень є аналіз просторових частот.

Кадри оброблюються по блоках, як правило, квадратами 8×8 пікселів. В результаті перетворення виникає матриця з 64 коефіцієнтів. Коефіцієнт – число, яке визначає зміст в блоці певної просторової частоти.

Дискретне косинусне перетворення (DCT) представляє зображення у вигляді суми синусоїд різної величини та частоти. Функція $dct2$ обчислює двовимірне дискретне косинусне перетворення (DCT) зображення. DCT має властивість, що для типового зображення більша частина візуально значущої інформації про зображення концентрується лише у кількох коефіцієнтах DCT. З

цієї причини DCT часто використовується в програмах стиснення зображень. Наприклад, DCT лежить в основі міжнародного стандартного алгоритму стиснення зображень із втратами, відомого як JPEG. (Назва походить від робочої групи, яка розробила стандарт: формат графічних даних JPEG).

Двовимірний DCT матриці M -by- N A визначається наступним чином.

$$B_{pq} = a_p a_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad \begin{array}{l} 0 \leq p \leq M-1 \\ 0 \leq q \leq N-1 \end{array}$$

$$a_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq p \leq M-1 \end{cases} \quad a_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq q \leq N-1 \end{cases}$$

Значення B_{pq} називаються коефіцієнтами DCT A . DCT є оборотним перетворенням.

$$A_{mn} = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} a_p a_q B_{pq} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad \begin{array}{l} 0 \leq m \leq M-1 \\ 0 \leq n \leq N-1 \end{array}$$

$$a_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, & 1 \leq p \leq M-1 \end{cases} \quad a_q = \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, & 1 \leq q \leq N-1 \end{cases}$$

Обернене рівняння DCT можна інтерпретувати як таке, що будь-яку матрицю M-N-N можна записати як суму функцій MN виду

$$a_p a_q \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad \begin{matrix} 0 \leq p \leq M-1 \\ 0 \leq q \leq N-1 \end{matrix}$$

$$T_{pq} = \begin{cases} \frac{1}{\sqrt{M}} & p = 0, & 0 \leq q \leq M-1 \\ \sqrt{\frac{2}{M}} \cos \frac{\pi(2q+1)p}{2M} & 1 \leq p \leq M-1, & 0 \leq q \leq M-1 \end{cases}$$

Ці функції називаються базовими функціями DCT.

3.2. Колування за допомогою швидкого перетворення Уолша-Адамара

Перетворення Уолша-Адамара – це несинусоїдальний, ортогональний метод перетворення, який розкладає сигнал на набір базових функцій. Ці базові функції є функціями Уолша, які являють собою прямокутні або квадратні хвилі зі значеннями +1 або -1.

Перетворення Уолша-Адамара повертає значення послідовності. Послідовність є більш узагальненим поняттям частоти і визначається як половина середньої кількості нульових перетинів за одиницю інтервалу часу. Кожна функція Уолша має унікальне значення послідовності. Ви можете використовувати повернені значення послідовності для оцінки частот сигналу в вихідному сигналі.

Перетворення Уолша-Адамара застосується в обробці зображень, обробці мови, фільтрації та аналізу спектра потужностей. Це дуже корисно для зменшення вимог до пропускну здатності та аналізу розширеного спектру.

Перетворення Уолша-Адамара має швидку версію, швидке перетворення Уолша-Адамара (FWHT). Порівняно з DCT, FWHT вимагає менше місця для зберігання і швидше обчислюється, оскільки використовує лише додавання та віднімання, тоді як DCT вимагає складних значень. FWHT здатний подавати сигнали з різкими розривами більш точно, використовуючи меншу кількість коефіцієнтів. Як FWHT, так і зворотний FWHT є симетричними і, отже, використовують однакові процеси обчислення. FWHT та IFWHT для сигналу $x(t)$ довжиною N визначаються як:

$$y_n = \frac{1}{N} \sum_{i=0}^{N-1} x_i \text{WAL}(n, i),$$

$$x_i = \sum_{n=0}^{N-1} y_n \text{WAL}(n, i),$$

Всі коефіцієнти FWHT рівні $+1$ або -1 . FWHT може бути зведене до операцій додавання і віднімання (без ділення або множення).

У ході розробки додатку ми дійшли висновку, що за допомогою швидкого перетворення Уолша-Адамара є можливість значно зменшити енергозатрати для пристроїв при здійсненні наступних дій:

- перегляд відеозаписів та стрімів;
- відеодзвіник та відео конференції;
- запис відео.

Таким чином, у користувачів буде можливість або відтворювати відео в стандартному режимі у коротший проміжок часу, оскільки енергозатратність

батареї буде більшою, або економити до 40% енергозатратності свого пристрою, втративши лише 3-7% якості відео.

У наш час це дуже актуальне питання, оскільки більшість пристроїв є автономними, а отже потребують постійної підзарядки. Наша розробка дасть можливість при незначній втраті якості, яка може бути непомітною для людського ока та не є критичною, набагато довше бути на зв'язку під час роботи чи навчання, під час подорожі чи поїздки за місто. Для батареї це дає також позитивний ефект, тому що кількість циклів заряду зменшується, тому термін експлуатації збільшується.

ВИСНОВКИ

На сьогоднішній день люди все частіше і частіше використовують відеоматеріали. Через карантинні обмеження у житті людей збільшилося використання гаджетів для роботи та навчання.

У дипломному проекті ми розглянули способи перетворення відео. Це потрібно для того, щоб мати можливість відображати відео. Без зжимання, перетворення відео- та фото-матеріалів займали б у 10 разів більше пам'яті.

Основним об'єктом дослідження є швидке перетворення Уолшера-Адамара. Ключовою перевагою над іншими способами перетворення можна вважати зменшення витрат електроенергії.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ТА ЛІТЕРАТУРИ

1. Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М.: ДИАЛОГ-МИФИ, 2003. 384 с.
2. Сэломон Д. Сжатие данных, изображений и звука. М.: Техносфера, 2004. 368 с.
3. Семенюк В.В. Экономное кодирование дискретной информации. СПб: СПбГИТМО (ТУ), 2001. 115 с.
4. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. М.: Техносфера, 2005. 368 с.
5. Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC), 2003.
6. Image and Video Coding – Emerging Standards and Beyond. // IEEE Transactions on Circuits and Systems for Video Technology. 1997. V. 8. № 7. P. 814–837.
7. ITU-T Recommendation H.264 (03/2005) / Prepublished version.
8. Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, Ajay Luthra. Overview of the H.264/AVC Video Coding Standard. // IEEE Transactions on Circuits and Systems for Video Technology, July 2003.
9. Ожиганов А.А., Тропченко А.А., Тропченко А.Ю. Модифицированный фрактальный метод сжатия многоуровневых изображений.// Информационные технологии. 2003. № 4.
10. Білінський Й. Й., Огородник К. В., Юкиш М. Й. Електронні системи : навч. посіб. / Він. держ. техн. ун-т. Вінниця : ВДТУ, 2011. 208 с.
11. Loeffler C., Ligtenberg A., Moschytz G. Practical Fast 1-D DCT Algorithms with 11 Multiplications. Proc. Int'l. Conf. on Acoustics, Speech, and Signal Processing 1989 (ICASSP '89). P. 991.

12. Hongqing Zhu, Huazhong S., Liang J., Lumin L., Coatrieux J. Image analysis by discrete orthogonal Racah moments. Signal Processing. April, 2007. P. 708.
13. Цифрова обробка аудіо- та відеоінформації у мультимедійних системах : навч. посіб. / О. В. Дробик, В. В. Кідалов, В. В. Коваль, Б. Я. Костік, В. С. Лазебний, Г. М. Розорінов, Г. О. Сукач. Київ : Наукова думка, 2008. 144 с.
14. Михалевський Д. В. Мельник В. М., Є. С. Наугольних Оцінка параметрів відеозображення в телекомунікаційних системах. Вимірювальна та обчислювальна техніка в технологічних процесах. № 1. 2013. С. 4.
15. Клименко Л. А. Козелков О.А. Анализ методов сжатия цифрового изображения. Системы обработки информации. Харьков : 2004. Вып. 2. С. 191-194.
16. Айфичер Э., Джервис Б. Цифровая обработка сигналов. Практический подход. Пер. с англ.: 2-ое издание. – Москва-Санкт-Петербург/Киев: Вильямс, 2004.
17. Кавкаев Артем, Пискунов Денис Сергеевич, Козырева Лариса Леонидовна. URL: <https://rosuchebnik.ru/material/issledovatel'skiy-proekt-sokhranyaem-prirodu-na-ekonomii-elektroenergii-7455/>
18. Ушаков В.Я. Современная и перспективная энергетика: технологические, социально-экономические и экологические аспекты. – Томск: Издво ТПУ, 2008. – 469 с
19. Лукутин Б.В., Суржикова О.А., Шандарова Е.Б. Возобновляемая энергетика в децентрализованном электроснабжении. – М.: Энергоатомиздат, 2008. – 231 с.