Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет _____Кібербезпеки_____
(повна назва)

Кафедра _____Інформаційно-мережна інженерія_____
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

рівень вищої освіти_____перший(бакалаврський)_____
____Аналіз ефективності застосування індексу структурної подібності SSIM___
_____для оцінювання стиснених із втратами зображень_____
(тема)

Виконав:
здобувач _____4____ року навчання,
групи _ТРІМІ-22-2_____
_____Іван Солдатенко_____
(власне ім'я, прізвище)

Спеціальність __172 Телекомунікації та_____
_радіотехніка_____
(код і повна назва спеціальності)
Тип програми ____освітньо-професійна_____
Освітня програма ___Інформаційно-мережна інженерія_____
(повна назва освітньої програми)

Керівник _ст._викл._Олексій Федоров_____
(посада, власне ім'я, прізвище)

Допускається до захисту

В. о. зав. кафедри _____ __Дарія Чеботарьова
(підпис) (власне ім'я, прізвище)

2026 р.

Харківський національний університет радіоелектроніки

Факультет Кібербезпеки

Кафедра Інформаційно-мережна інженерія

Рівень вищої освіти перший (бакалаврський)

Спеціальність 172 Телекомунікації та радіотехніка
(код і повна назва)

Тип програми освітньо-професійна

Освітня програма "Інформаційно-мережна інженерія"
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

«____»_____ 20 ___ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Солдатенко Іван Андрійович
(прізвище, ім'я, по батькові)

1. Тема роботи Аналіз ефективності застосування індексу структурної подібності SSIM для оцінювання стиснених із втратами зображень

затверджена наказом університету від 02 06 2026 р. № 466Ст

2. Термін подання здобувачем роботи до екзаменаційної комісії 16 07 2026 р.

3. Вихідні дані до роботи

1. Провести аналіз цифрових зображень, способів їх отримання, подання та зберігання.
2. Дослідити ортогональні перетворення, що застосовуються при стисненні зображень.
3. Розглянути дискретне косинусне перетворення як основу JPEG-стиснення.
4. Дослідити принципи стиснення зображень із втратами та формат JPEG.
5. Проаналізувати традиційні метрики оцінювання якості зображень MSE та PSNR.
6. Дослідити індекс структурної подібності SSIM.

Виконати аналіз ефективності SSIM для оцінювання якості стиснених із втратами зображень.

4. Перелік питань, що потрібно опрацювати в роботі

ВСТУП.

1 ЗОБРАЖЕННЯ, СПОСОБИ ОТРИМАННЯ ТА ЗБЕРІГАННЯ ЗОБРАЖЕНЬ.

2 ОРТОГОНАЛЬНІ ПЕРЕТВОРЕННЯ, ЩО ЗАСТОСОВУЮТЬСЯ ПРИ СТИСНЕННІ ЗОБРАЖЕНЬ.

3 СТИСНЕННЯ ЗОБРАЖЕНЬ ІЗ ВТРАТАМИ. ФОРМАТ JPEG.

4 ОСНОВИ ОЦІНЮВАННЯ ЯКОСТІ ЗОБРАЖЕНЬ.

5 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ.

ВИСНОВКИ

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ.

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій (п.5 включається до завдання за рішенням випускової кафедри) _____
слайди презентації в форматі MS Power Point _____

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

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

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1.	Ознайомлення із завданням. Уточнення ТЗ.	09.06.-11.06.2026	виконано
2.	Підбір та ознайомлення з літературою за темою роботи.	12.06.-16.06.2026	виконано
3.	Виконання розділу 1-2.	18.06.-26.06.2026	виконано
4.	Виконання розділу 3-4.	28.06.-02.06.2026	виконано
5.	Виконання розділу 5.	10.06.-11.07.2026	виконано
6.	Оформлення роботи та презентації.	11.06.-13.07.2026	виконано
7.	Перевірка на плагіат.	14.07.-15.07.2026	виконано

Дата видачі завдання 08 06 2026 р.

Здобувач _____
(підпис)

Керівник роботи _____ ст. викл. Олексій Федоров
(підпис) (посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка до кваліфікаційної роботи бакалавра містить: 59 сторінок, 13 рисунків, 13 таблиць, 10 джерел.

Об’єкт дослідження – процес оцінювання якості цифрових зображень після їх стиснення із втратами.

Предмет дослідження – індекс структурної подібності SSIM та його застосування для оцінювання якості стиснених цифрових зображень.

Мета роботи – аналіз ефективності застосування індексу структурної подібності SSIM для оцінювання якості цифрових зображень після стиснення із втратами.

Об'єктом дослідження є процес оцінювання якості цифрових зображень після їх стиснення із втратами.Предметом дослідження є індекс структурної подібності SSIM та його застосування для оцінювання якості стиснених цифрових зображень. Методами дослідження є методи цифрової обробки зображень, математичного моделювання, статистичного аналізу, методи порівняльного аналізу та експериментальні дослідження. Практичне значення отриманих результатів полягає у дослідженні можливостей застосування індексу SSIM під час оцінювання ефективності алгоритмів стиснення цифрових зображень, а також у порівнянні його характеристик із традиційними метриками оцінювання якості.

Отримані результати показують, що SSIM є корисним інструментом для аналізу структурних змін у зображеннях, однак для оцінювання якості JPEG-стиснених зображень доцільно застосовувати комплексний підхід із використанням SSIM, MSE, PSNR та візуального аналізу артефактів.

ЦИФРОВЕ ЗОБРАЖЕННЯ, СТИСНЕННЯ З ВТРАТАМИ, JPEG, ДИСКРЕТНЕ КОСИНУСНЕ ПЕРЕТВОРЕННЯ, MSE, PSNR, SSIM, ОЦІНЮВАННЯ ЯКОСТІ ЗОБРАЖЕНЬ

ABSTRACT

The explanatory note to the bachelor's qualification work contains: 59 pages, 13 figures, 13 tables, 10 sources.

The object of the study is the process of assessing the quality of digital images after lossy compression.

The subject of the study is the Structural Similarity Index SSIM and its application for assessing the quality of compressed digital images.

The purpose of the work is to analyze the effectiveness of using the Structural Similarity Index SSIM for evaluating the quality of digital images after lossy compression.

The object of this study is the process of evaluating the quality of digital images after lossy compression. The subject of this study is the Structural Similarity Index (SSIM) and its application in evaluating the quality of compressed digital images. The research methods include digital image processing, mathematical modeling, statistical analysis, comparative analysis, and experimental studies. The practical significance of the results lies in investigating the potential applications of the SSIM index in evaluating the effectiveness of digital image compression algorithms, as well as in comparing its characteristics with traditional quality assessment metrics.

The obtained results show that SSIM is a useful tool for analyzing structural changes in images. However, for assessing the quality of JPEG-compressed images, it is advisable to use a combined approach involving SSIM, MSE, PSNR and visual analysis of compression artifacts.

DIGITAL IMAGE, LOSSY COMPRESSION, JPEG, DISCRETE COSINE TRANSFORM, MSE, PSNR, SSIM, IMAGE QUALITY ASSESSMENT

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАК, СИМВОЛІВ, ОДИНИЦЬ І ТЕРМІНІВ……8
ВСТУП……11
1 ЗОБРАЖЕННЯ, СПОСОБИ ОТРИМАННЯ ТА ЗБЕРІГАННЯ ЗОБРАЖЕНЬ..13
1.1 Поняття цифрового зображення……13
1.2 Формування цифрових зображень, дискретизація та квантування……16
1.3 Подання, колірні моделі та зберігання цифрових зображень……..19
1.4 Особливості зорової системи людини……...22
2 ОРТОГОНАЛЬНІ ПЕРЕТВОРЕННЯ, ЩО ЗАСТОСОВУЮТЬСЯ ПРИ СТИСНЕННІ ЗОБРАЖЕНЬ……..24
2.1 Загальна ідея перетворень у цифровій обробці зображень……24
2.2 Дискретне косинусне перетворення……..28
2.3 Значення ортогональних перетворень для стиснення зображень……...33
3 СТИСНЕННЯ ЗОБРАЖЕНЬ ІЗ ВТРАТАМИ. ФОРМАТ JPEG……..36
3.1 Основні принципи стиснення цифрових зображень……36
3.2 Стиснення з втратами та джерела спотворень……..38
3.3 Формат JPEG та базовий алгоритм стиснення……..39
3.4 Артефакти JPEG-стиснення та значення для оцінювання SSIM……..….44
4 ОСНОВИ ОЦІНКИ ЯКОСТІ ЗОБРАЖЕНЬ……47
4.1 Традиційні метрики оцінювання якості……47
4.2 Недоліки піксельних метрик……..48
4.3 Індекс структурної подібності SSIM……48
4.4 Локальне обчислення SSIM та карта якості……..49
4.5 Обмеження SSIM при оцінюванні стиснених зображень……50
5 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ……52
5.1 Мета експериментального аналізу……52
5.2 Методика експериментів Дозельман і Яанг……..53

5.3 Результати рангового порівняння SSIM і MSE…….56
5.4 Формальний зв’язок між SSIM та MSE…….57
5.5 Значення результатів для оцінювання JPEG-зображень……..59
ВИСНОВКИ……...61
ПЕРЕЛІК ПОСИЛАНЬ…….63
Додаток А СЛАЙДИ ПЕРЕЗЕНТАЦІЙ……...64

ПЕРЕЛІК СКОРОЧЕНЬ, УМОВНИХ ПОЗНАК, СИМВОЛІВ, ОДИНИЦЬ І ТЕРМІНІВ

SSIM (Structural Similarity Index) – індекс структурної подібності зображень.

MSSIM (Mean Structural Similarity Index) – середнє значення локальних оцінок SSIM для всього зображення.

MSE (Mean Squared Error) – середньоквадратична помилка.

PSNR (Peak Signal-to-Noise Ratio) – пікове відношення сигналу до шуму.

JPEG (Joint Photographic Experts Group) – стандарт стиснення нерухомих цифрових зображень із втратами.

JPEG2000 – стандарт стиснення зображень, що використовує вейвлет-перетворення.

DCT (Discrete Cosine Transform) – дискретне косинусне перетворення.

FDCT (Forward Discrete Cosine Transform) – пряме дискретне косинусне перетворення.

IDCT (Inverse Discrete Cosine Transform) – обернене дискретне косинусне перетворення.

DFT (Discrete Fourier Transform) – дискретне перетворення Фур’є.

FFT (Fast Fourier Transform) – швидке перетворення Фур’є.

WHT (Walsh-Hadamard Transform) – перетворення Уолша-Адамара.

KLT (Karhunen-Loeve Transform) – перетворення Карунена-Лоева.

PCA (Principal Component Analysis) – метод головних компонент.

RGB (Red, Green, Blue) – колірна модель, що використовує червону, зелену та синю компоненти.

YCbCr – колірна модель, у якій Y відповідає яскравісній складовій, а Cb і Cr – кольоророзницевим складовим.

DC (Direct Current coefficient) – коефіцієнт нульової частоти у блоці DCT, що відповідає середньому рівню яскравості блока.

AC (Alternating Current coefficient) – коефіцієнти ненульових частот у блоці DCT, що описують просторові зміни яскравості.

PSF (Point Spread Function) – функція розсіювання точки.

HVS (Human Visual System) – зорова система людини.

IQM (Image Quality Metric) – метрика якості зображення.

BPP (Bits Per Pixel) – кількість бітів на один піксель.

Cb – синя кольорорізницева компонента у просторі YCbCr.

Cr – червона кольорорізницева компонента у просторі YCbCr.

Y – яскравісна компонента зображення.

R – червона компонента кольору у моделі RGB.

G – зелена компонента кольору у моделі RGB.

B – синя компонента кольору у моделі RGB.

Q – матриця або коефіцієнт квантування.

Q-JPEG – параметр якості JPEG-стиснення.

f(x,y) – значення інтенсивності або яскравості зображення в точці з координатами xxx, yyy.

F(u,v) – коефіцієнт зображення у частотній області після перетворення.

x,y – просторові координати пікселя.

u,v – частотні координати або індекси коефіцієнтів.

M,N – розміри зображення або блока зображення за вертикаллю та горизонталлю.

L – максимальне можливе значення пікселя, наприклад 255 для 8-бітного зображення.

μ_x, μ_y – середні значення яскравості відповідних областей зображень.

σ_x, σ_y – середньоквадратичні відхилення відповідних областей зображень.

σ_{xy} – коваріація між відповідними областями двох зображень.

C_1, C_2, C_3 – стабілізаційні константи у формулі SSIM.

l(x,y) – складова SSIM, що відповідає порівнянню яскравості.

c(x,y) – складова SSIM, що відповідає порівнянню контрасту.

s(x,y) – складова SSIM, що відповідає порівнянню структури.

r_s – коефіцієнт рангової кореляції Спірмена.

Zig-zag ordering – zig-zag впорядкування DCT-коефіцієнтів у JPEG.

Квантування – процес зменшення точності подання коефіцієнтів або значень пікселів.

Ентропійне кодування – спосіб кодування даних з урахуванням частоти появи символів.

Блокові артефакти – спотворення у вигляді видимих меж між блоками зображення, характерні для сильного JPEG-стиснення.

ВСТУП

У сучасному інформаційному суспільстві цифрові зображення є одним із найважливіших видів даних. Вони широко застосовуються в системах комп'ютерного зору, телекомунікаціях, медицині, дистанційному зондуванні Землі, відеоспостереженні, мультимедійних технологіях, цифровій фотографії та багатьох інших галузях. Постійне збільшення обсягів графічної інформації обумовлює необхідність ефективного зберігання та передавання зображень, що досягається шляхом використання алгоритмів стиснення.

Найбільшого поширення набули методи стиснення із втратами, які дозволяють значно зменшити обсяг даних за рахунок незначного погіршення візуальної якості зображення. Одним із найвідоміших стандартів такого стиснення є JPEG, що базується на дискретному косинусному перетворенні. Незважаючи на високу ефективність, застосування алгоритмів стиснення призводить до виникнення різноманітних спотворень, тому важливим завданням є об'єктивне оцінювання якості отриманих зображень.

Традиційно для оцінювання якості використовують середньоквадратичну помилку (Mean Squared Error, MSE) та пікове відношення сигнал/шум (Peak Signal-to-Noise Ratio, PSNR). Однак зазначені метрики ґрунтуються лише на порівнянні значень окремих пікселів і недостатньо враховують особливості сприйняття зображень людиною. Через це результати оцінювання нерідко не відповідають суб'єктивній візуальній оцінці. Одним із сучасних підходів до оцінювання якості цифрових зображень є індекс структурної подібності Structural Similarity Index (SSIM). На відміну від традиційних методів, SSIM враховує зміни яскравості, контрасту та структури зображення, що дозволяє точніше моделювати особливості людської зорової системи. Саме тому ця метрика отримала широке застосування у задачах стиснення, передавання, відновлення та покращення цифрових зображень. Разом із тим ефективність застосування SSIM для оцінювання якості зображень після стиснення із

втратами залишається предметом наукових досліджень. Аналіз переваг і недоліків цього індексу, а також його порівняння з традиційними методами оцінювання є актуальним завданням сучасної цифрової обробки сигналів.

Метою кваліфікаційної роботи є аналіз ефективності застосування індексу структурної подібності SSIM для оцінювання якості цифрових зображень після стиснення із втратами. Для досягнення поставленої мети необхідно вирішити такі задачі:

– дослідити принципи формування, подання та зберігання цифрових зображень;

– розглянути ортогональні перетворення, що застосовуються під час стиснення зображень;

– дослідити принципи роботи алгоритму JPEG;

– проаналізувати існуючі методи оцінювання якості цифрових зображень;

– дослідити принцип роботи індексу структурної подібності SSIM;

– виконати порівняльний аналіз ефективності SSIM та традиційних метрик оцінювання якості зображень.

1 ЗОБРАЖЕННЯ, СПОСОБИ ОТРИМАННЯ ТА ЗБЕРІГАННЯ ЗОБРАЖЕНЬ

1.1 Поняття цифрового зображення

Цифрове зображення є однією з основних форм подання візуальної інформації в сучасних інформаційних, телекомунікаційних та мультимедійних системах. На відміну від аналогового зображення, яке є неперервним, цифрове зображення складається з окремих елементів — пікселів. Кожен піксель має певне числове значення, що характеризує яскравість або колір відповідної ділянки зображення. У загальному випадку зображення можна розглядати як двовимірну функцію просторових координат. Значення цієї функції визначає інтенсивність світла у відповідній точці площини. Для монохромного зображення кожній точці відповідає одне значення яскравості, а для кольорового – кілька компонентів кольору, наприклад червона, зелена та синя складові у моделі RGB.

У [1] зазначено, що двовимірний сигнал, отриманий сенсором, є відображенням тривимірної сцени на площину. Автори [1] підкреслюють: “sampled and quantized to yield digital images”. Це означає, що для формування цифрового зображення необхідно виконати дискретизацію за просторовими координатами та квантування за рівнями яскравості або кольору. Формула (1.1), [1] описує неперервне зображення як двовимірну функцію інтенсивності

$$s(x, y) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \delta(x - j\Delta x, y - k\Delta y), \qquad (1.1)$$

де x — горизонтальна просторова координата; y — вертикальна просторова координата; f(x, y) — значення яскравості або інтенсивності зображення в точці з координатами x та y.

Після дискретизації зображення подається у вигляді матриці. Кожен елемент цієї матриці відповідає окремому пікселю. Якщо зображення має розмір M × N, то воно містить M рядків і N стовпців, а загальна кількість пікселів дорівнює M · N. Чим більшою є кількість пікселів, тим вищою є просторова роздільна здатність зображення.

Важливою характеристикою цифрового зображення є глибина квантування. Вона визначає кількість бітів, що використовуються для подання значення одного пікселя. Для бінарного зображення достатньо одного біта, оскільки піксель може мати лише два значення: 0 або 1. Для зображень у відтінках сірого зазвичай використовують 8 бітів на піксель, що дає 256 рівнів яскравості – від чорного до білого. Формула (1.2), [1], показує зв'язок між кількістю бітів і кількістю можливих рівнів квантування

$$M = 2^n, \tag{1.2}$$

де M – кількість рівнів квантування; n – кількість бітів, що використовуються для кодування одного відліку або одного компонента кольору.

Залежно від способу подання пікселів цифрові зображення поділяють на бінарні, напівтонові та кольорові. Бінарні зображення використовуються для тексту, схем і силуетів. Напівтонові зображення містять кілька рівнів яскравості та застосовуються, наприклад, у медичних знімках або монохромній фотографії. Кольорові зображення містять кілька колірних компонентів і використовуються у цифровій фотографії, мультимедіа та комп'ютерній графіці таблиця (1.1).

Таблиця 1.1 — Основні типи цифрових зображень за способом подання пікселя

Тип зображення	Подання пікселя	Основні особливості	Приклад застосування
Бінарне	1 біт: 0 або 1	Містить лише два рівні яскравості	Текстові документи, маски, технічні схеми
Напівтонове	Один канал яскравості	Кожен піксель задає рівень сірого	Медичні знімки, чорно-білі фотографії
Кольорове RGB	Три канали: R, G, B	Колір формується комбінацією червоного, зеленого та синього каналів	Фотографії, веб-графіка
Індексоване	Піксель містить номер кольору з палітри	Дає змогу зменшити обсяг даних	GIF-зображення, прості ілюстрації
Багатоканальне	Декілька спектральних каналів	Містить інформацію не лише у видимому діапазоні	Супутникові та наукові зображення

Отже, цифрове зображення можна визначити як дискретне двовимірне подання візуальної інформації, сформоване шляхом дискретизації та квантування аналогового сигналу. Воно складається з пікселів, кожен з яких має числове значення або набір значень. Основними характеристиками цифрового зображення є просторова роздільна здатність, глибина квантування, кількість колірних компонентів та обсяг даних.

1.2 Формування цифрових зображень, дискретизація та квантування

Формування цифрового зображення починається з реальної сцени, яка освітлюється джерелом світла. Світлові промені відбиваються від поверхонь об'єктів і потрапляють на оптичну систему пристрою реєстрації, наприклад цифрової камери. Після проходження через об'єктив світло фокусується на сенсорній площині, де оптичний сигнал перетворюється на електричний. У джерелі [1] зазначено, що процес формування зображення пов'язаний із трьома основними складовими: освітленням, відбивними властивостями поверхні та формуванням зображення на сітківці ока або сенсорній площині камери. Тому якість отриманого цифрового зображення залежить не тільки від характеристик камери, але й від умов освітлення, властивостей об'єкта та оптичної системи.

Одним із факторів, що впливає на якість зображення, є точність фокусування оптичної системи. Якщо об'єктив неправильно сфокусований або має оптичні недоліки, точковий об'єкт не відображається як ідеальна точка. Замість цього формується розмита пляма. Для опису такого явища використовується функція розсіювання точки – Point Spread Function (PSF). Вона показує, як система формування зображення відтворює окрему світлову точку. Формула (1.3), [1] описує отримане зображення як результат згортки ідеального зображення з функцією розсіювання точки

$$I_{res}(x, y) = I_{id}(x, y) \otimes P(x, y), \qquad (1.3)$$

де $I_{res}(x, y)$ – результуюче зображення; $I_{id}(x, y)$ – ідеальне зображення; P(x, y) – функція розсіювання точки; $\otimes$ – операція згортки.

Практично це означає, що якість цифрового зображення залежить не тільки від кількості пікселів, але й від фізичних властивостей системи отримання зображення. Навіть зображення з високою роздільною здатністю може бути нечітким, якщо воно отримане за умов неправильного фокусування,

руху камери або об'єкта, недостатнього освітлення чи наявності атмосферних спотворень.

Після формування аналогового зображення виконується його дискретизація. Дискретизація – це процес перетворення неперервного просторового сигналу в набір окремих відліків. У цифровому зображенні такими відліками є пікселі. Частота дискретизації визначає, наскільки детально буде представлено зображення. Якщо крок дискретизації занадто великий, дрібні деталі можуть бути втрачені або спотворені. Формула (1.4), [1] описує процес отримання дискретизованого зображення шляхом множення неперервної функції зображення на функцію дискретизації

$$f_s(x,\ y) = f(x,y)s(x,y) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(j\Delta x, k\Delta y) \cdot \delta(x - j\Delta x, y - k\Delta y), \quad (1.4)$$

де f(x, y) - неперервна функція інтенсивності; s(x, y) - функція дискретизації; fs(x, y) - дискретизоване зображення; Δx - крок дискретизації за горизонталлю; Δy - крок дискретизації за вертикаллю.

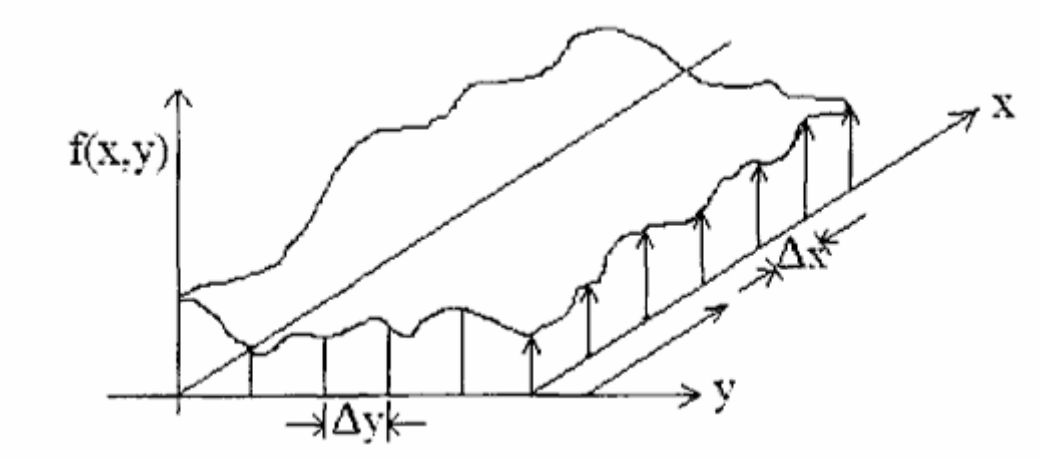


Рисунок 1.2 – Двовимірна сітка дискретизації цифрового зображення

Важливим правилом під час дискретизації є критерій Найквіста. Згідно з ним, частота дискретизації повинна бути не меншою ніж удвічі більшою за

максимальну просторову частоту сигналу. Для зображень це означає, що дрібні деталі сцени можуть бути правильно відтворені лише за умови достатньої кількості відліків. Якщо дискретизація виконується з недостатньою частотою, виникає ефект накладання спектрів, або aliasing. Візуально це може проявлятися у вигляді помилкових контурів, “сходинок”, муару або втрати дрібних деталей. Формула (1.5), [1] описує умову дискретизації за критерієм Найквіста для двовимірного зображення

$$w_{xs} \geq 2w_x^m,\ w_{ys} \geq 2w_y^m, \tag{1.5}$$

w_{xs} - частота дискретизації за горизонтальною координатою;

w_{ys} - частота дискретизації за вертикальною координатою;

w_x^m - максимальна просторова частота зображення за горизонталлю;

w_y^m - максимальна просторова частота зображення за вертикаллю.

Після дискретизації виконується квантування. Квантування – це процес перетворення неперервних значень яскравості в обмежену кількість дискретних рівнів. Наприклад, якщо використовується 8-бітне квантування, кожне значення яскравості може бути подане одним із 256 рівнів. Збільшення кількості рівнів квантування дозволяє точніше передати плавні переходи яскравості, але водночас збільшує обсяг даних. У разі недостатньої кількості рівнів квантування можуть виникати помітні спотворення. Одним із таких спотворень є контуризація, коли плавний перехід яскравості замінюється різкими ступінчастими переходами. Це особливо помітно на ділянках зображення, де яскравість змінюється повільно, наприклад на небі, тінях або рівномірно освітлених поверхнях. Таким чином, формування цифрового зображення є багатоступеневим процесом. Він включає освітлення сцени, відбиття світла, фокусування оптичною системою, реєстрацію сенсором, дискретизацію та

квантування. Кожен із цих етапів впливає на кінцеву якість зображення, а також на можливість його подальшого стиснення, зберігання та оцінювання.

1.3 Подання, колірні моделі та зберігання цифрових зображень

Після отримання цифрове зображення повинно бути подане у формі, придатній для обробки, зберігання та передавання. Найчастіше цифрове зображення розглядають як матрицю числових значень. Для зображення у відтінках сірого використовується одна матриця, елементи якої відповідають рівням яскравості пікселів. Для кольорового зображення використовується кілька матриць, кожна з яких відповідає окремій колірній компоненті. Найпоширенішою моделлю подання кольору є RGB. У цій моделі кожен піксель описується трьома складовими: R - червоною, G - зеленою та B - синьою. Якщо кожна складова кодується 8 бітами, то один піксель потребує 24 біти пам'яті. Таке подання широко використовується у цифровій фотографії, комп'ютерній графіці, дисплеях і відеокамерах.

Однак модель RGB не завжди є найзручнішою для обробки та стиснення зображень. У багатьох задачах доцільно відокремлювати яскравісну складову від колірної інформації. Це пов'язано з особливостями людського сприйняття: людина більш чутлива до змін яскравості, ніж до змін кольору. Тому в системах стиснення зображень часто використовуються колірні моделі, у яких яскравість подається окремо від кольорових компонентів. Однією з таких моделей є YCbCr. У цій моделі Y відповідає яскравісній складовій, а Cb і Cr містять колірнорізницеву інформацію. Таке подання широко використовується в цифровому відео та стандартах стиснення зображень. Саме перехід від RGB до YCbCr дозволяє ефективніше використовувати особливості зорової системи людини та зменшувати обсяг даних під час стиснення. Формула (1.6), [1] описує перетворення колірних компонентів з простору RGB у простір YCbCr

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.16874 & -0.33126 & 0.50000 \\ 0.50000 & -0.41869 & -0.08131 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix}, \quad (1.6)$$

де R - червона складова кольору; G - зелена складова кольору; B - синя складова кольору; Y - яскравісна складова; Cb - синя колірнорізницева складова; Cr - червона колірнорізницева складова.

Для оцінювання обсягу нестисненого цифрового зображення можна використати залежність між кількістю пікселів і кількістю бітів на один піксель. Формула (1.7), визначає обсяг пам'яті, необхідний для зберігання нестисненого цифрового зображення

$$V = M \cdot N \cdot b, \quad (1.7)$$

де V - обсяг зображення в бітах; M - кількість рядків; N - кількість стовпців; b - кількість бітів на один піксель.

Наприклад, кольорове RGB-зображення розміром 1920 × 1080 за умови 24-бітного подання одного пікселя потребує значного обсягу пам'яті. Якщо таких зображень багато або якщо йдеться про відеопослідовність, вимоги до пам'яті та пропускної здатності каналу передавання зростають дуже швидко. Саме тому важливу роль відіграють формати зберігання та методи стиснення зображень. У джерелі [1] розглянуто різні формати файлів для зберігання зображень, зокрема TIFF, PostScript, PBM, PGM, PPM та інші. Формат файлу визначає спосіб організації даних, наявність заголовка, спосіб кодування пікселів, підтримку кольору, прозорості та стиснення. Одні формати орієнтовані на просте зберігання растрових даних, інші — на ефективне стиснення або поліграфічне використання.

Таблиця 1.2 — Основні формати зберігання цифрових зображень

Формат	Тип стиснення	Основні переваги	Основні недоліки	Типове застосування
BMP	Без стиснення або просте стиснення	Простота структури, легкість обробки	Великий розмір файлу	Технічні та проміжні файли
TIFF	Без втрат або без стиснення	Висока якість, підтримка різних режимів	Значний розмір файлу	Поліграфія, архівування
PNG	Стиснення без втрат	Збереження якості, підтримка прозорості	Менш ефективний для фотографій	Веб-графіка, схеми
GIF	Стиснення без втрат для палітрових зображень	Підтримка анімації	Обмежена кількість кольорів	Прості анімації, іконки
JPEG	Стиснення з втратами	Малий розмір файлу, зручність для фотографій	Втрата частини інформації, артефакти	Цифрові фотографії, веб
WebP	З втратами або без втрат	Висока ефективність стиснення	Не завжди зручний для старих систем	Веб-зображення

У межах цього розділу формат JPEG розглядається лише як один із поширених форматів зберігання цифрових зображень. Детальний алгоритм JPEG-стиснення, включаючи перетворення кольору, дискретне косинусне перетворення, квантування та ентропійне кодування, буде розглянуто в окремому розділі роботи. Отже, спосіб подання і зберігання цифрового зображення визначає не тільки його обсяг, але й можливості подальшої обробки. Вибір колірної моделі, глибини квантування та формату файлу впливає на якість зображення, швидкість обробки, ефективність стиснення та зручність передавання каналами зв'язку.

1.4 Особливості зорової системи людини

Під час оцінювання якості цифрових зображень важливо враховувати не лише технічні параметри файлу або числові значення пікселів, але й особливості сприйняття зображення людиною. Людська зорова система не сприймає зображення як набір незалежних точок. Вона аналізує структуру сцени, контури, контраст, текстури, співвідношення між сусідніми ділянками та просторове розташування об'єктів. У джерелі [1] зазначено, що зорове сприйняття людини пов'язане з обробкою сигналів, які надходять від сусідніх рецепторів сітківки. Це означає, що око й мозок реагують не тільки на абсолютну яскравість, але й на різницю між сусідніми ділянками зображення. Саме тому контури та межі об'єктів мають особливо важливе значення для розпізнавання сцени.

Людська сітківка містить два основних типи фоторецепторів: палички та колбочки. Палички відповідають переважно за сприйняття яскравості при низькому рівні освітлення, а колбочки — за кольоровий зір при достатньому освітленні. Колбочки мають різну спектральну чутливість і забезпечують сприйняття кольорів. У джерелі [1] зазначено, що існують три типи колбочок, які пов'язані зі сприйняттям різних діапазонів довжин хвиль.

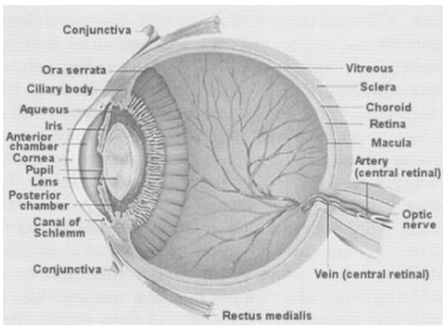


Рисунок 1.3 – Будова людського ока

З погляду цифрової обробки зображень особливо важливим є те, що людина по-різному сприймає різні типи спотворень. Наприклад, невелика зміна

яскравості може бути менш помітною, ніж руйнування контурів або текстур. Так само зображення з однаковою середньою числовою помилкою можуть сприйматися людиною як зображення різної якості. Це пояснюється тим, що візуальне сприйняття залежить не тільки від окремих пікселів, але й від структури всієї сцени.

У джерелі [2] підкреслюється, що об'єктивне оцінювання якості зображень залишається складною задачею, оскільки потрібно враховувати “quality as perceived by the human visual system”. Це твердження важливе для даної дипломної роботи, оскільки традиційні метрики, які ґрунтуються лише на різниці між значеннями пікселів, не завжди відповідають реальному зоровому сприйняттю. Саме ця особливість пояснює необхідність використання більш складних метрик оцінювання якості зображень. Якщо метрика враховує лише числову різницю між відповідними пікселями оригінального та стисненого зображення, вона може не відображати реальну помітність спотворень. Наприклад, шум, розмиття, блокові артефакти JPEG-стиснення та втрата дрібних текстур можуть мати різний вплив на сприйняття, навіть якщо числові значення помилки є близькими. У зв'язку з цим у подальших розділах роботи буде розглянуто індекс структурної подібності SSIM. Його ідея полягає в тому, що якість зображення доцільно оцінювати не тільки через різницю між пікселями, а й через зміну яскравості, контрасту та структури. У цьому розділі SSIM не розглядається детально, оскільки його математична модель буде наведена окремо в розділі, присвяченому оцінюванню якості зображень.

Таким чином, зорове сприйняття людини є важливою основою для аналізу якості цифрових зображень. Людина сприймає не лише окремі значення яскравості, а передусім контури, текстури, контраст і структуру сцени. Тому під час дослідження стиснення зображень із втратами необхідно враховувати не тільки технічні параметри файлу, але й те, як спотворення впливають на візуальне сприйняття.

2 ОРТОГОНАЛЬНІ ПЕРЕТВОРЕННЯ, ЩО ЗАСТОСОВУЮТЬСЯ ПРИ СТИСНЕННІ ЗОБРАЖЕНЬ

2.1 Загальна ідея перетворень у цифровій обробці зображень

Цифрове зображення можна розглядати як двовимірний дискретний сигнал, значення якого відповідають яскравості або колірним компонентам окремих пікселів. У просторовій області таке зображення подається у вигляді матриці, елементи якої безпосередньо відповідають пікселям. Однак для багатьох задач цифрової обробки зображень, зокрема для фільтрації, аналізу частотного складу та стиснення, зручніше перейти від просторового подання до іншого простору ознак або коефіцієнтів. Перетворення зображень використовуються для того, щоб подати початкові дані у формі, у якій певні властивості сигналу стають більш очевидними. Наприклад, у просторовій області важко безпосередньо визначити, які ділянки зображення містять повільні зміни яскравості, а які — різкі переходи, контури або дрібні текстури. У частотній області ці властивості описуються через низько- та високочастотні складові. Низькі частоти відповідають плавним змінам яскравості, а високі частоти — різким переходам, межам об'єктів і деталям.

У джерелі [1] зазначено, що різні ортогональні перетворення дозволяють аналізувати зображення як двовимірний сигнал. Зокрема, перетворення Фур'є дає змогу подати зображення як сукупність просторових синусоїд різних напрямів і частот. Такий підхід є важливим для розуміння структури зображення, оскільки він показує, які частотні компоненти формують візуальний вміст сцени. Найбільш відомим інструментом частотного аналізу є перетворення Фур'є. Для цифрових зображень використовується дискретне перетворення Фур'є, яке працює з дискретними відліками сигналу. Воно дозволяє перейти від матриці значень яскравості до матриці частотних

коефіцієнтів. Формула (2.1), [1] описує двовимірне дискретне перетворення Фур'є для зображення розміром M × N

$$F(u,v) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)exp\left[-j2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)\right], \quad (2.1)$$

де f(x, y) — значення пікселя зображення в просторовій області; F(u, v) — коефіцієнт зображення в частотній області; x, y — просторові координати пікселя; u, v — частотні координати; M, N — розміри зображення за вертикаллю та горизонталлю.

Двовимірне перетворення дає змогу отримати частотне представлення всього зображення. У цьому представленні коефіцієнти з малими значеннями u та v відповідають низькочастотним складовим, а коефіцієнти з великими значеннями u та v — високочастотним деталям. Для задач стиснення це має важливе значення, оскільки енергія природних зображень зазвичай зосереджена переважно в низькочастотній області.

Окрему роль у цифровій обробці відіграють ортогональні перетворення. Їх особливість полягає в тому, що початковий сигнал розкладається за системою взаємно ортогональних базисних функцій. Завдяки цьому зменшується залежність між коефіцієнтами, а інформація про зображення може бути подана більш компактно. Така властивість є основою перетворювального кодування, яке широко використовується у стисненні зображень. Для квадратного зображення або блока зображення перетворення може бути записане у матричній формі. Це зручно для пояснення практичної реалізації алгоритмів обробки, оскільки зображення вже подається як матриця. Формула (2.2), [1] описує матричне подання двовимірного перетворення зображення

$$F = Af(x,y)A, \quad (2.2)$$

де f(x, y) – початкова матриця зображення або блока зображення; F – матриця коефіцієнтів після перетворення; A – матриця перетворення; A^{-1} або обернена матриця — використовується для відновлення початкового зображення.

Матричне подання показує, що перетворення зображення можна розглядати як зміну базису. У початковому просторі зображення описується значеннями пікселів, а після перетворення — набором коефіцієнтів. Якщо перетворення є оборотним, то за всіма коефіцієнтами можна відновити початкове зображення без втрат. Однак у задачах стиснення часто використовують не всі коефіцієнти або зберігають їх із різною точністю, що дозволяє зменшити обсяг даних. У цифровій обробці зображень застосовують різні перетворення: дискретне перетворення Фур'є, дискретне косинусне перетворення, перетворення Уолша-Адамара, перетворення Карунена-Лоева та інші. Кожне з них має власні переваги та обмеження.

Дискретне перетворення Фур'є добре підходить для спектрального аналізу, але його коефіцієнти є комплексними, що ускладнює практичне використання у деяких алгоритмах стиснення. Перетворення Уолша-Адамара має просту структуру, оскільки його базисні функції містять лише значення зі знаками плюс і мінус, але воно не завжди забезпечує достатньо ефективну концентрацію енергії зображення. Перетворення Карунена-Лоева може забезпечувати дуже ефективне подання даних, але воно залежить від статистичних властивостей конкретного набору зображень, тому є складнішим для практичного використання. Найбільш важливим для стиснення фотографічних зображень є дискретне косинусне перетворення. Воно використовує лише дійсні коефіцієнти, має хорошу властивість концентрації енергії та може ефективно реалізовуватися для невеликих блоків зображення. Саме тому DCT стало основою стандарту JPEG, який буде розглянуто в наступному розділі.

Таблиця 2.1 – Основні ортогональні перетворення, що застосовуються в цифровій обробці зображень

Назва перетворення	Основна ідея	Переваги	Обмеження	Застосування
DFT	Подання зображення через комплексні синусоїдальні складові	Точний частотний аналіз	Комплексні коефіцієнти	Спектральний аналіз, фільтрація
DCT	Подання зображення через косинусні базисні функції	Висока концентрація енергії, дійсні коефіцієнти	Можливі блокові артефакти при стисненні	JPEG, MPEG
WHT	Подання через функції Уолша-Адамара	Простота обчислень	Нижча ефективність для фотографічних зображень	Швидкі алгоритми аналізу
KLT/PCA	Статистично оптимальне подання даних	Ефективне зменшення розмірності	Залежність від статистики даних	Розпізнавання образів, аналіз даних

Отже, ортогональні перетворення є важливим інструментом цифрової обробки зображень. Вони дозволяють перейти від просторового подання до подання через коефіцієнти, зменшити надлишковість даних і підготувати зображення до ефективного стиснення. Найбільше практичне значення для даної роботи має дискретне косинусне перетворення, оскільки саме воно використовується в алгоритмі JPEG-стиснення із втратами.

2.2 Дискретне косинусне перетворення

Дискретне косинусне перетворення, або DCT, є одним із найважливіших перетворень у цифровій обробці зображень. Його широке використання пояснюється тим, що воно дозволяє ефективно перерозподіляти інформацію зображення між коефіцієнтами. У результаті більша частина енергії зображення зосереджується в невеликій кількості низькочастотних коефіцієнтів, тоді як багато високочастотних коефіцієнтів мають малі значення.

У джерелі [1] прямо зазначено важливість DCT для стиснення: [1] “Discrete Cosine Transform (DCT) is the basis for many image and video compression algorithms”. Це твердження є принциповим для даної роботи, оскільки JPEG-стиснення базується саме на використанні DCT для перетворення блоків зображення перед квантуванням і кодуванням. DCT можна розглядати як перетворення, у якому сигнал подається через суму косинусних функцій різної частоти. На відміну від дискретного перетворення Фур'є, DCT використовує тільки дійсні косинусні базисні функції. Це спрощує обчислення та зберігання коефіцієнтів. Крім того, для природних зображень DCT добре концентрує енергію в низькочастотній області, що робить його зручним для стиснення.

Спочатку розглядають одновимірне пряме дискретне косинусне перетворення. Воно застосовується до послідовності відліків, наприклад до одного рядка зображення або одного стовпця. Результатом такого перетворення є набір DCT-коефіцієнтів. Формула (2.3), [1] описує одновимірне пряме дискретне косинусне перетворення

$$F(u) = \sqrt{\frac{2}{N}} C(u) \sum_{x=0}^{N-1} f(x) cos\left[\frac{\pi(2x+1)u}{2N}\right], \quad (2.3)$$

де f(x) — значення x-го відліку вхідного сигналу; F(u) — DCT-коефіцієнт

з індексом u; N — кількість відліків сигналу; x — індекс відліку у просторовій області; u — індекс коефіцієнта у частотній області; C(u) — нормувальний коефіцієнт, значення якого залежить від індексу u.

Нормувальний коефіцієнт C(u) використовується для коректного масштабування результатів перетворення. Для нульового коефіцієнта він має одне значення, а для інших коефіцієнтів — інше. Завдяки цьому забезпечується правильне енергетичне співвідношення між початковим сигналом і його перетвореним поданням.

Обернене дискретне косинусне перетворення дозволяє відновити початкову послідовність відліків із набору DCT-коефіцієнтів. Якщо всі коефіцієнти збережено без змін, відновлення є точним. Якщо частина коефіцієнтів була відкинута або квантизована з втратами, відновлений сигнал буде наближеним до початкового. Формула (2.4), [1] описує одновимірне обернене дискретне косинусне перетворення

$$f(x) = \sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} C(u)F(u)cos\left[\frac{\pi(2x+1)u}{2N}\right], \quad (2.4)$$

де f(x) - відновлене значення x-го відліку сигналу; F(u) - DCT-коефіцієнт; C(u) - нормувальний коефіцієнт; N - кількість відліків; x - індекс відліку у просторовій області; u - індекс коефіцієнта у частотній області.

Для цифрових зображень більш важливим є двовимірне DCT, оскільки зображення має дві просторові координати. Двовимірне пряме DCT застосовується до блока або всього зображення і перетворює матрицю значень пікселів у матрицю DCT-коефіцієнтів. Формула (2.5), [1] описує двовимірне пряме дискретне косинусне перетворення для блока розміром M × N

$$F(u,v) = \frac{2}{\sqrt{MN}} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) cos\left[\frac{\pi(2x+1)u}{2N}\right] cos\left[\frac{\pi(2y+1)v}{2M}\right] \quad , \tag{2.5}$$

де f(x, y) - значення пікселя або відліку в точці з координатами x, y; F(u, v) - DCT-коефіцієнт з індексами u, v; M, N - розміри блока або зображення; x, y - просторові координати; u, v - частотні координати;C(u), C(v) - нормувальні коефіцієнти.

У матриці DCT-коефіцієнтів коефіцієнт з індексами u = 0 та v = 0 називається DC-коефіцієнтом. Він характеризує середній рівень яскравості блока. Інші коефіцієнти називаються AC-коефіцієнтами і описують зміни яскравості різної просторової частоти. Коефіцієнти, розташовані ближче до початку матриці, відповідають нижчим частотам, а коефіцієнти, розташовані далі, — вищим частотам.

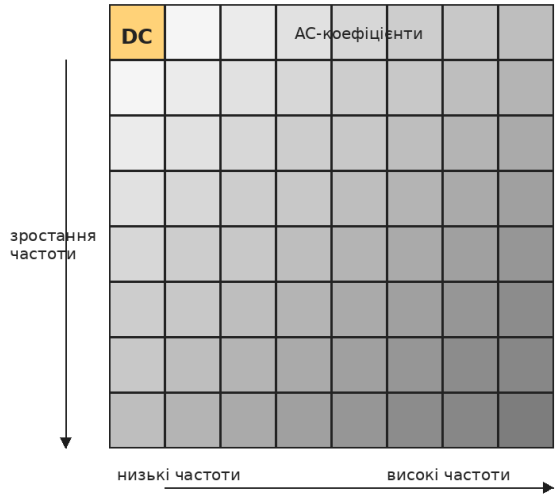


Рисунок 2.1 – Розташування DC- та AC-коефіцієнтів у блоці DCT

Однією з найважливіших властивостей DCT є сепарабельність. Це означає, що двовимірне перетворення можна виконати у два етапи: спочатку застосувати одновимірне DCT до рядків, а потім – до стовпців. Такий підхід значно спрощує реалізацію алгоритму та зменшує обчислювальну складність.

Формула (2.6), [1] описує перший етап обчислення двовимірного DCT – застосування одновимірного DCT до рядків блока

$$F(u,y) = \sqrt{\frac{2}{N}}C(u)\sum_{x=0}^{N-1} f(x,y)cos\left[\frac{\pi(2x+1)u}{2N}\right], \quad (2.6)$$

де f(x, y) - початкове значення пікселя в рядку y; F(u, y) - проміжний результат після перетворення рядків; x - індекс пікселя в рядку; u - індекс частотного коефіцієнта за горизонталлю; N - кількість відліків у рядку; C(u) - нормувальний коефіцієнт.

Формула (2.7), [1] описує другий етап обчислення двовимірного DCT – застосування одновимірного DCT до стовпців проміжної матриці

$$F(u,v) = \sqrt{\frac{2}{N}}C(v)\sum_{y=0}^{M-1} f(u,y)cos\left[\frac{\pi(2y+1)v}{2M}\right], \quad (2.7)$$

де F(u, y) - проміжний результат після перетворення рядків; F(u, v) - кінцевий DCT-коефіцієнт; y - індекс рядка; v - індекс частотного коефіцієнта за вертикаллю; M - кількість відліків у стовпці; C(v) - нормувальний коефіцієнт.

Сепарабельність є дуже важливою для практичного використання DCT у стандартах стиснення. Замість безпосереднього обчислення двовимірної формули для кожного коефіцієнта можна послідовно виконати набір одновимірних перетворень. Це особливо зручно для блоків 8 × 8 пікселів, які використовуються у базовому алгоритмі JPEG.

Обернене двовимірне DCT використовується для відновлення блока зображення за його DCT-коефіцієнтами. У системах стиснення цей етап виконується під час декодування. Якщо коефіцієнти були змінені внаслідок квантування, відновлене зображення буде відрізнятися від початкового.

Формула (2.8), [1] описує двовимірне обернене дискретне косинусне перетворення

$$f(x,y) = \frac{2}{\sqrt{MN}} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} C(u)C(v)F(u,v)cos\left[\frac{\pi(2x+1)u}{2N}\right]cos\left[\frac{\pi(2y+1)v}{2M}\right], \quad (2.8)$$

де f(x, y) - відновлене значення пікселя; F(u, v) - DCT-коефіцієнт; M, N - розміри блока або зображення; x, y - просторові координати; u, v - частотні координати; C(u), C(v) - нормувальні коефіцієнти.

З погляду стиснення зображень головною перевагою DCT є концентрація енергії. У більшості природних зображень сусідні пікселі мають близькі значення яскравості, тобто між ними існує сильна кореляція. Після DCT ця кореляція зменшується, а основна інформація переходить у невелику кількість коефіцієнтів, переважно в низькочастотній області. Це дозволяє застосовувати подальше квантування. Низькочастотні коефіцієнти, які найбільше впливають на візуальне сприйняття, зберігаються точніше. Високочастотні коефіцієнти, які часто відповідають дрібним деталям і шуму, можуть бути збережені менш точно або частково відкинуті. Саме на цьому принципі базується JPEG-стиснення із втратами.

Таблиця 2.2 — Інтерпретація коефіцієнтів DCT у блоці зображення

Область коефіцієнтів	Тип частот	Що описує	Значення для стиснення
Лівий верхній кут	Нульова частота	Середню яскравість блока	Зберігається з високою точністю
Верхня ліва область	Низькі частоти	Плавні зміни яскравості	Найбільш важлива частина інформації
Середня область	Середні частоти	Контури та переходи	Частково зберігається
Права нижня область	Високі частоти	Дрібні деталі та шум	Може квантуватися грубіше

Таким чином, DCT не стискає зображення саме по собі, але створює умови для ефективного стиснення. Воно перетворює блок пікселів у блок коефіцієнтів, де інформація розподілена нерівномірно. Саме ця нерівномірність дозволяє зменшити кількість даних без різкого погіршення візуальної якості.

2.3 Значення ортогональних перетворень для стиснення зображень

Стиснення цифрових зображень ґрунтується на усуненні надлишковості. У природних зображеннях сусідні пікселі часто мають близькі значення, особливо в однорідних областях. Це означає, що між пікселями існує просторова кореляція. Якщо зберігати кожен піксель незалежно, то частина інформації дублюється. Ортогональні перетворення дозволяють зменшити таку надлишковість шляхом переходу до нового подання даних. У перетвореній області інформація про зображення подається не через окремі пікселі, а через коефіцієнти. Кожен коефіцієнт відповідає певній базисній функції. Якщо перетворення вибрано вдало, більша частина важливої інформації зосереджується у відносно невеликій кількості коефіцієнтів. Це дозволяє зберігати або передавати не всі дані з однаковою точністю. Перетворювальне кодування зазвичай складається з кількох етапів. Спочатку зображення або його блоки перетворюються з просторової області в область коефіцієнтів. Потім коефіцієнти квантуються. Після цього вони кодуються з використанням ентропійного кодування або інших методів зменшення обсягу даних. У разі стиснення із втратами саме квантування є основним джерелом втрати інформації.

DCT є найбільш поширеним перетворенням для стиснення фотографічних зображень, оскільки воно має декілька важливих властивостей. По-перше, воно працює з дійсними коефіцієнтами, що спрощує реалізацію. По-друге, воно є сепарабельним, тому двовимірне перетворення можна

виконувати через послідовність одновимірних операцій. По-третє, воно добре концентрує енергію природних зображень у низькочастотній області.

У JPEG-стисненні зображення поділяється на блоки, найчастіше розміром 8 × 8 пікселів. До кожного блока застосовується DCT, після чого отримується блок із 64 коефіцієнтів. Перший коефіцієнт описує середню яскравість блока, а решта коефіцієнтів описують різні просторові зміни. Далі коефіцієнти квантуються: менш важливі високочастотні коефіцієнти можуть бути збережені з меншою точністю. У цьому розділі наведена лише загальна роль DCT у процесі стиснення. Детальний опис JPEG-стиснення, включаючи перетворення кольору, поділ на блоки, квантування DCT-коефіцієнтів та ентропійне кодування, буде розглянуто в наступному розділі роботи. Варто зазначити, що DCT не є єдиним можливим перетворенням для стиснення. Перетворення Карунена-Лоева теоретично може забезпечувати дуже ефективне компактне подання даних, оскільки воно враховує статистичні властивості конкретного набору зображень. Однак така залежність від даних ускладнює практичне використання цього перетворення у масових стандартах стиснення. Перетворення Уолша-Адамара простіше для обчислення, але не забезпечує такої ефективної концентрації енергії для природних фотографічних зображень, як DCT.

Перевага DCT полягає у компромісі між якістю, ефективністю та складністю реалізації. Воно достатньо добре наближається до оптимального перетворення для багатьох природних зображень, але при цьому не потребує попереднього статистичного аналізу кожного набору даних. Саме тому DCT стало основою багатьох практичних алгоритмів стиснення. Разом із перевагами DCT має й певні обмеження. Оскільки у JPEG воно застосовується до окремих блоків, при сильному стисненні можуть виникати блокові артефакти. Вони проявляються у вигляді помітних меж між сусідніми блоками. Крім того, грубе квантування високочастотних коефіцієнтів може призводити до втрати дрібних деталей, розмиття текстур і появи спотворень біля контурів.

Ці спотворення мають безпосереднє значення для теми дипломної роботи. Оскільки DCT і подальше квантування змінюють структуру зображення, виникає необхідність оцінювати, наскільки стиснене зображення залишається подібним до оригінального. Для цього можуть використовуватися традиційні метрики, такі як MSE та PSNR, а також більш складні підходи, зокрема індекс структурної подібності SSIM.

Таким чином, ортогональні перетворення відіграють ключову роль у стисненні цифрових зображень. Вони дозволяють перейти від подання через пікселі до подання через коефіцієнти, зменшити кореляцію між даними та підготувати зображення до ефективного квантування і кодування. Серед таких перетворень найбільше практичне значення має дискретне косинусне перетворення, оскільки воно поєднує обчислювальну ефективність, хорошу концентрацію енергії та широке застосування у стандартах JPEG і MPEG.

3 СТИСНЕННЯ ЗОБРАЖЕНЬ ІЗ ВТРАТАМИ. ФОРМАТ JPEG

3.1 Основні принципи стиснення цифрових зображень

Цифрові зображення потребують значного обсягу пам’яті для зберігання та передавання. У книзі Acharya і Ray [1] зазначено, що цифрове подання сигналів має багато переваг порівняно з аналоговим, однак воно потребує великої кількості бітів. Наприклад, навіть зображення або відео середньої роздільної здатності можуть займати дуже великий обсяг пам’яті. Через це стиснення зображень є важливою задачею цифрової обробки сигналів, мультимедійних систем і телекомунікацій. Основна ідея стиснення полягає у зменшенні надлишковості даних. У природних зображеннях сусідні пікселі часто мають близькі значення яскравості або кольору. Це означає, що між ними існує просторова залежність. Якщо зберігати кожен піксель незалежно, частина інформації дублюється. Методи стиснення намагаються подати зображення так, щоб зменшити цю надлишковість.

У джерелі [1] також підкреслюється, що зображення містять не тільки інформаційну, але й надлишкову частину. Інформація повинна бути збережена для правильного відновлення змісту зображення, а надлишковість може бути зменшена або вилучена. Саме тому стиснення можна розглядати як процес зменшення надлишковості у представленні даних. Важливим поняттям теорії інформації є ентропія. Вона характеризує середню кількість інформації, що припадає на один символ джерела. Якщо деякі символи трапляються частіше, а інші рідше, то такі дані можна закодувати ефективніше. Формула (3.1), [1] описує ентропію дискретного джерела інформації

$$H(A) = -\sum_{i=1}^{N} p(a_i) log_2 p(a_i), \qquad (3.1)$$

де A – алфавіт джерела; ai – окремий символ алфавіту; p(ai) – ймовірність появи символу ai; N – кількість символів алфавіту; H(A) – ентропія джерела.

Чим менша ентропія джерела, тим менше в середньому бітів потрібно для кодування його символів. Це є основою ентропійного кодування, яке використовується у багатьох методах стиснення. У JPEG, зокрема, застосовується кодування Хаффмана, яке дозволяє коротше кодувати частіші символи. Стиснення зображень поділяють на стиснення без втрат і стиснення з втратами. При стисненні без втрат після декодування можна повністю відновити початкове зображення. При стисненні з втратами відновлене зображення вже не є точною копією оригіналу, але може залишатися візуально дуже схожим на нього. Саме такий підхід використовується у стандартному JPEG-стисненні фотографічних зображень.

Таблиця 3.1 — Основні види надлишковості у цифрових зображеннях

Вид надлишковості	Сутність	Приклад	Спосіб зменшення
Просторова	Сусідні пікселі часто мають близькі значення	Однорідні ділянки неба або стіни	Перетворення, прогнозування
Спектральна	Колірні компоненти можуть бути взаємопов’язані	Подібність каналів RGB	Перехід до YCbCr
Психовізуальна	Людське око не однаково сприймає всі деталі	Менша чутливість до високих частот	Квантування
Кодова	Символи мають різну ймовірність появи	Часті нулі після квантування DCT	Ентропійне кодування

Отже, стиснення зображень ґрунтується на усуненні різних видів надлишковості. Для стиснення з втратами особливо важливо, що частину даних можна вилучити так, щоб зображення залишалося прийнятним для людського сприйняття.

3.2 Стиснення з втратами та джерела спотворень

Стиснення з втратами передбачає навмисне вилучення частини інформації із зображення. Такий підхід дозволяє отримати значно менший розмір файлу, ніж стиснення без втрат. Однак після декодування зображення вже не збігається повністю з оригіналом. У JPEG основними джерелами втрат є субдискретизація колірних компонентів і квантування DCT-коефіцієнтів. Субдискретизація зменшує кількість даних у кольорових компонентах, а квантування знижує точність частотних коефіцієнтів. Саме квантування є головною причиною появи спотворень у JPEG-зображенні.

Квантування є процесом округлення або наближеного подання значень. У результаті частина точності втрачається. Якщо квантування слабке, різниця між оригінальним і відновленим зображенням може бути майже непомітною. Якщо квантування сильне, з’являються видимі спотворення: втрата дрібних деталей, розмиття, зміни текстур і блокові артефакти.

Для оцінювання якості стиснених зображень часто використовують традиційні метрики, зокрема MSE та PSNR. Середньоквадратична помилка MSE оцінює середню різницю між відповідними пікселями двох зображень. Формула (3.2), [2] описує середньоквадратичну помилку MSE між оригінальним і спотвореним зображеннями

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2, \tag{3.2}$$

де xi - значення і-го пікселя оригінального зображення; yi - значення і-го пікселя стисненого або спотвореного зображення; n - загальна кількість пікселів; MSE - середньоквадратична помилка.

Однак MSE і PSNR не завжди точно відповідають людському сприйняттю якості. Вони порівнюють числові значення пікселів, але не враховують повністю структуру зображення, контури, текстури та характер спотворень. Саме тому в роботі розглядається SSIM. Водночас, відповідно до зауваження викладача, не можна заздалегідь стверджувати, що SSIM завжди кращий за PSNR для JPEG-зображень. Це потрібно перевіряти експериментально.

3.3 Формат JPEG та базовий алгоритм стиснення

JPEG є одним із найпоширеніших стандартів стиснення нерухомих зображень. Назва JPEG походить від Joint Photographic Experts Group. У джерелі [1] зазначено, що JPEG є першим міжнародним стандартом стиснення напівтонових нерухомих зображень. Він може застосовуватися як до зображень у відтінках сірого, так і до кольорових зображень.

Важливо розрізняти стандарт JPEG і конкретний файловий формат. Acharya і Ray [1] підкреслюють, що JPEG визначає синтаксис стисненого потоку даних, але не задає єдиний обов’язковий формат файлу. Для зберігання JPEG-стиснених зображень можуть використовуватися різні формати, наприклад JFIF. JPEG підтримує кілька режимів роботи: послідовний режим без втрат, послідовний DCT-режим із втратами, прогресивний DCT-режим та ієрархічний режим. Найбільш поширеним є базовий послідовний DCT-режим, або baseline JPEG. Саме він використовується у більшості практичних випадків стиснення фотографічних зображень.

Перед стисненням кольорове зображення зазвичай перетворюється з простору RGB у простір YCbCr. Компонента Y відповідає яскравості, а компоненти Cb і Cr - кольоророзницевій інформації. Такий перехід є корисним, оскільки людська зорова система більш чутлива до яскравості, ніж до кольорових змін. Після перетворення кольору компоненти Cb і Cr можуть бути субдискретизовані. Наприклад, у форматі 4:2:0 кожна кольоророзницева

компонента має меншу роздільну здатність, ніж компонента яскравості. Це дозволяє зменшити обсяг даних без різкого погіршення видимої якості.

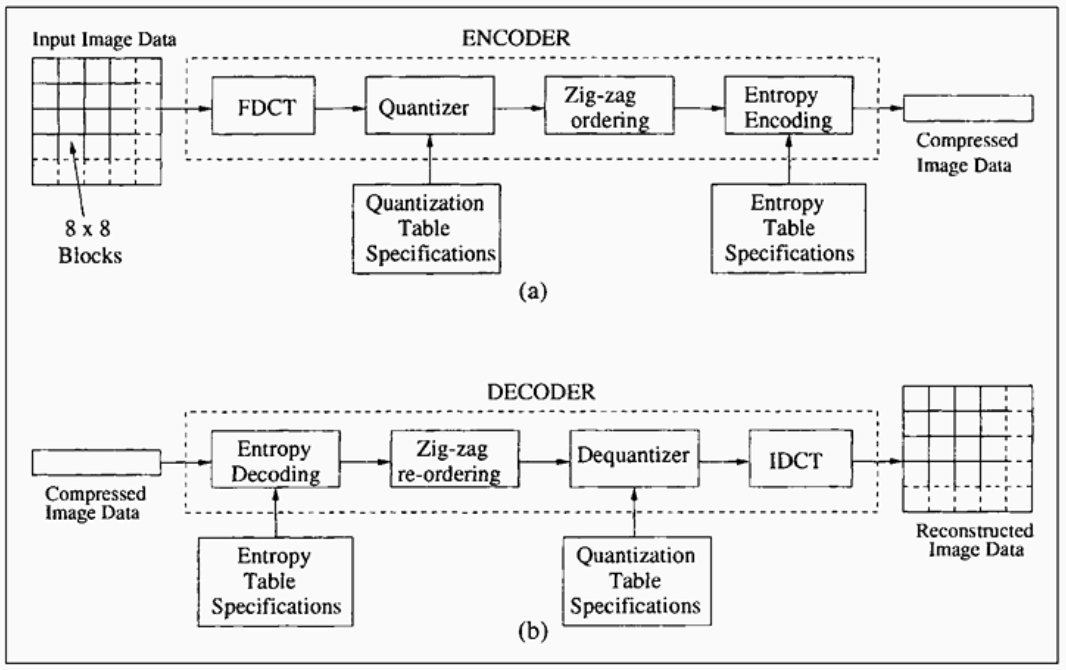


Рисунок 3.2 – Основні етапи baseline JPEG-стиснення: (a) compression, (b) decompression.

На рисунку 3.2 показано загальну схему роботи базового алгоритму JPEG. У верхній частині рисунка наведено процес кодування зображення. Спочатку вхідне зображення поділяється на блоки розміром 8 × 8 пікселів. До кожного блока застосовується пряме дискретне косинусне перетворення FDCT, після чого отримані коефіцієнти квантуються з використанням таблиць квантування. Далі коефіцієнти впорядковуються у zig-zag послідовності та стискаються за допомогою ентропійного кодування. У нижній частині рисунка зображено зворотний процес декодування. Стиснені дані проходять ентропійне декодування, після чого виконується зворотне zig-zag впорядкування, деквантування та обернене дискретне косинусне перетворення IDCT. У результаті формується відновлене зображення. Оскільки під час квантування

частина інформації втрачається, відновлене зображення може відрізнятися від початкового. Саме цей етап є основним джерелом втрат у JPEG-стисненні.

Після підготовки компонент зображення поділяється на блоки розміром 8 × 8 пікселів. Кожен блок обробляється окремо. Перед застосуванням DCT значення пікселів зсуваються за рівнем: для 8-бітних зображень із кожного значення віднімається 128. Це переводить значення пікселів у діапазон, приблизно симетричний відносно нуля. Далі до кожного блока застосовується пряме дискретне косинусне перетворення. Воно переводить блок із просторової області в частотну область. Формула (3.3), [1] описує пряме дискретне косинусне перетворення FDCT для блока 8 × 8 у baseline JPEG

$$F(u,v) = \frac{1}{4}C(u)C(v)\sum_{x=0}^{7}\sum_{y=0}^{7} f(x,y)cos\left[\frac{(2x+1)u\pi}{16}\right]cos\left[\frac{(2y+1)v\pi}{16}\right], \quad (3.3)$$

де f(x, y) – значення пікселя у блоці після зсуву рівня; F(u, v) – DCT-коефіцієнт; x, y – координати пікселя у блоці; u, v – частотні індекси; C(u), C(v) – нормувальні коефіцієнти.

Формула (3.4), [1] задає значення нормувального коефіцієнта C(k), який використовується у FDCT

$$C(k) = \begin{cases} \frac{1}{\sqrt{2}}, & k = 0, \\ 1, & k \neq 0. \end{cases}, \quad (3.4)$$

де C(k) – нормувальний коефіцієнт; k – частотний індекс; значення k = 0 відповідає нульовій частоті.

Після DCT блок пікселів перетворюється на блок із 64 коефіцієнтів. Коефіцієнт F(0,0) називається DC-коефіцієнтом і описує середній рівень яскравості блока. Інші 63 коефіцієнти називаються AC-коефіцієнтами і описують зміни яскравості різної просторової частоти.

Наступним етапом є квантування DCT-коефіцієнтів. Саме цей етап є основним джерелом втрат у baseline JPEG. Під час квантування кожен DCT-коефіцієнт ділиться на відповідний елемент матриці квантування та округлюється. Формула (3.5), [1] описує квантування DCT-коефіцієнтів у JPEG

$$F_q(u, v) = round(\frac{F(u,v)}{Q(u,v)}), \quad (3.5)$$

де F(u, v) – DCT-коефіцієнт до квантування; Q(u, v) – елемент матриці квантування; Fq(u, v) – квантований DCT-коефіцієнт; u, v – частотні індекси.

Матриця квантування визначає, наскільки сильно будуть округлюватися коефіцієнти. Для низькочастотних коефіцієнтів зазвичай використовуються менші значення, а для високочастотних — більші. Це дозволяє точніше зберегти важливу інформацію і грубіше квантувати менш помітні деталі. У джерелі [1] наведені приклади матриць квантування для яскравісної та кольоро різницевих компонент (табл. 3.2 – 3.3).

Таблиця 3.2 – Матриці квантування у baseline JPEG. Матриця квантування для яскравості

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

Таблиця 3.3 – Матриці квантування у baseline JPEG. Матриця квантування для колірності

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

Якість JPEG-зображення може змінюватися залежно від коефіцієнта якості. У книзі Acharya і Ray [1] описується масштабування матриць квантування залежно від параметра Q-JPEG. Формула (3.6), [1] описує масштабувальний коефіцієнт для матриці квантування залежно від параметра Q-JPEG

$$\alpha = \begin{cases} \frac{50}{Q\text{-}JPEG}, & 1 \le Q\text{-}JPEG \le 50, \\ 2 - \frac{Q\text{-}JPEG}{50}, & 50 \le Q\text{-}JPEG \le 100. \end{cases} \quad (3.6)$$

де Q-JPEG — параметр якості JPEG; α — коефіцієнт масштабування матриці квантування; Q(u, v) — елемент початкової матриці квантування.

Після квантування багато високочастотних коефіцієнтів стають нульовими. Для ефективного кодування JPEG використовує zig-zag порядок обходу коефіцієнтів. Такий порядок дозволяє спочатку розмістити низькочастотні коефіцієнти, а потім високочастотні, серед яких часто багато нуліво як зображено на рисунку 3.3.

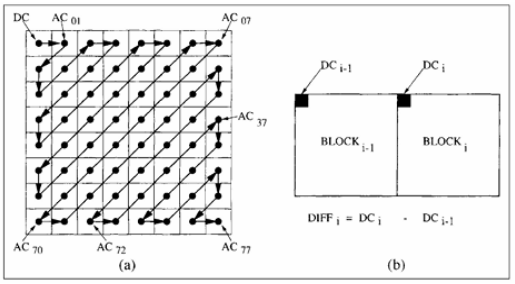


Рисунок 3.3 — Zig-zag порядок обходу DCT-коефіцієнтів у блоці 8 × 8

DC-коефіцієнти сусідніх блоків часто близькі між собою, тому JPEG кодує не саме значення DC, а різницю між DC-коефіцієнтом поточного та попереднього блока. Формула (3.7), [1] описує диференціальне кодування DC-коефіцієнтів у JPEG

$$DIFF_i = DC_i - DC_{i-1}, \tag{3.7}$$

де DCi – DC-коефіцієнт поточного блока; DCi-1 – DC-коефіцієнт попереднього блока; DIFF – різниця між DC-коефіцієнтами.

AC-коефіцієнти кодуються з урахуванням кількості нулів перед ненульовим коефіцієнтом. Для цього використовується run-length coding. Після цього отримані символи кодуються за допомогою кодування Хаффмана. Таким чином, JPEG поєднує перетворення, квантування та ентропійне кодування.

3.4 Артефакти JPEG-стиснення та значення для оцінювання SSIM

JPEG дозволяє значно зменшити розмір файлу, але при сильному стисненні у зображенні з’являються характерні спотворення. Acharya і Ray [1] зазначають, що при низькому значенні параметра якості стають помітними блокові артефакти. Вони виникають через те, що зображення обробляється незалежними блоками 8 × 8. При сильному квантуванні між сусідніми блоками можуть з’являтися помітні межі. Такі спотворення називаються blocking artifacts

(рисунок 3.4). Крім того, через втрату високочастотних коефіцієнтів можуть зникати дрібні деталі, текстури та тонкі контури. Основні типи спотворень після JPEG-стиснення наведено в таблиці 3.4.

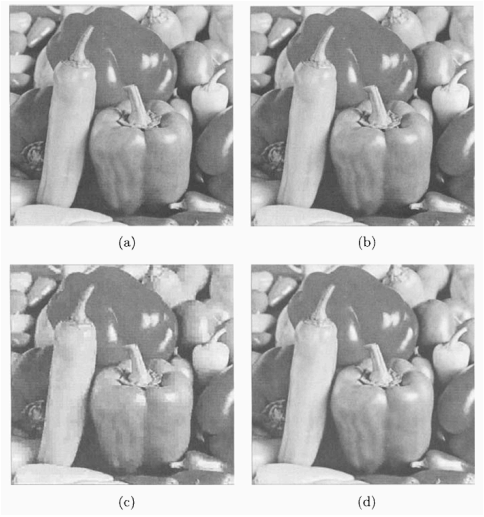


Рисунок 3.4 — Приклад впливу JPEG-стиснення на якість зображення; (a) Оригінальне зображення Перця, (b) стиснене за базовим стандартом JPEG із коефіцієнтом якості 75 (1,57 біта/піксель), (c) стиснене за базовим стандартом JPEG із коефіцієнтом якості 10 (0,24 біта/піксель) та (d) стиснене за новим стандартом JPEG2000 із тією самою швидкістю передачі даних (0,24 біта/піксель).

Таблиця 3.4 – Типові спотворення JPEG-зображень

Тип спотворення	Причина виникнення	Візуальний прояв
Блокові артефакти	Незалежна обробка блоків 8 × 8	Помітні межі між блоками
Розмиття деталей	Грубе квантування високих частот	Втрата дрібних текстур
Дзвін біля контурів	Спотворення високочастотних складових	Світлі або темні ореоли біля меж
Втрата кольорових деталей	Зменшення точності колірних компонентів	Спрощення кольорових переходів
Постеризація	Надмірне зменшення кількості відтінків	Ступінчасті переходи яскравості

Для теми даної роботи важливо, що JPEG-стиснення створює не випадкові, а структурно організовані спотворення. Вони залежать від блокової структури алгоритму, сили квантування та властивостей конкретного зображення. Саме тому оцінювання якості JPEG-зображень не можна зводити лише до формального обчислення однієї метрики. SSIM був запропонований як метрика, що враховує яскравість, контраст і структуру зображення. Через це можна очікувати, що SSIM буде корисним для аналізу JPEG-артефактів, оскільки вони часто змінюють локальну структуру блоків, контурів і текстур. Однак, відповідно до зауваження викладача, у роботі не слід заздалегідь стверджувати, що SSIM завжди кращий за PSNR або MSE.

У статті Dosselmann і Yang [2] показано, що SSIM має тісний зв'язок із MSE. Автори цієї статті звертають увагу, що в багатьох випадках SSIM і MSE можуть давати подібне ранжування спотворених зображень. Це важливо для даної кваліфікаційної роботи, оскільки її мета полягає не тільки в описі SSIM, а й у перевірці ефективності його застосування для оцінювання якості JPEG-зображень, стиснутих із втратами. У практичній частині роботи доцільно порівнювати SSIM з PSNR саме на JPEG-зображеннях з різними рівнями стиснення. Для кожного зображення потрібно змінювати параметр якості JPEG, обчислювати значення PSNR і SSIM, а також візуально аналізувати отримані артефакти. Такий підхід дозволить перевірити, чи справді SSIM краще відображає суб'єктивну якість JPEG-зображень, чи в окремих випадках PSNR дає не гірший або навіть стабільніший результат. Отже, JPEG є класичним прикладом стиснення зображень із втратами. Його ефективність забезпечується поєднанням перетворення кольору, блокового DCT, квантування, zig-zag обходу та ентропійного кодування. Водночас саме ці етапи створюють характерні спотворення, які потрібно враховувати при оцінюванні якості. Тому подальший аналіз SSIM має бути спрямований не на загальне твердження про його перевагу, а на експериментальну перевірку його ефективності для JPEG-стиснених зображень.

4 ОСНОВИ ОЦІНКИ ЯКОСТІ ЗОБРАЖЕНЬ

4.1 Традиційні метрики оцінювання якості

Оцінювання якості зображень є важливою задачею цифрової обробки, особливо після стиснення з втратами. Після JPEG-стиснення відновлене зображення вже не є точною копією оригіналу, тому необхідно визначити, наскільки значними є спотворення та як вони впливають на сприйняття якості. Найпростішим підходом є порівняння відповідних пікселів оригінального та спотвореного зображень. До таких методів належить середньоквадратична помилка MSE. Вона показує середнє значення квадратів різниць між пікселями двох зображень. Формула (4.1), [2] описує середньоквадратичну помилку MSE

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2 \quad , \tag{4.1}$$

де xi – значення і-го пікселя оригінального зображення; yi – значення і-го пікселя спотвореного зображення; n – кількість пікселів; MSE – середньоквадратична помилка.

Чим менше значення MSE, тим меншою є числова різниця між зображеннями. На основі MSE часто обчислюють PSNR — пікове відношення сигнал/шум. Ця метрика подає помилку в логарифмічній формі та є зручною для порівняння результатів стиснення. Формула (4.2) описує пікове відношення сигнал/шум PSNR через MSE

$$PSNR = 10log_{10}(\frac{L^2}{MSE}), \tag{4.2}$$

де L – максимальне можливе значення пікселя; MSE – середньоквадратична помилка; PSNR – пікове відношення сигнал/шум.

Перевагою MSE та PSNR є простота обчислення. Однак ці метрики мають суттєве обмеження: вони оцінюють тільки числову різницю між пікселями. Вони не враховують, як саме розташовані помилки у зображенні та чи руйнують вони контури, текстури або інші важливі структурні елементи.

4.2 Недоліки піксельних метрик

У звіті з практики було розглянуто приклади, які показують недосконаліть традиційних метрик. Два зображення можуть мати однакове або близьке значення MSE, але виглядати для людини по-різному. Це пояснюється тим, що людська зорова система сприймає не окремі пікселі, а структуру сцени. Наприклад, рівномірна зміна яскравості може бути менш помітною, ніж випадковий шум, хоча числова помилка може бути схожою. Так само перестановка пікселів може повністю зруйнувати текстуру, але проста піксельна метрика не завжди коректно відображає таку зміну.

Отже, для оцінювання якості стиснених зображень важливо враховувати не тільки величину помилки, а й її характер. Особливо це актуально для JPEG, оскільки цей формат створює структурні спотворення: блокові артефакти, втрату деталей, розмиття текстур і порушення контурів.

4.3 Індекс структурної подібності SSIM

Для врахування структурних особливостей зображення було запропоновано індекс структурної подібності SSIM. Його основна ідея полягає в тому, що якість зображення потрібно оцінювати не тільки через різницю між пікселями, а через зміну яскравості, контрасту та структури. SSIM порівнює два зображення або дві локальні області зображень за трьома складовими:

яскравість; контраст; структура. Формула (4.3), [2] описує загальний вигляд індексу SSIM як добуток трьох складових

$$SSIM(x, y) = [l(x, y)]^{\alpha}[c(x, y)]^{\beta}[s(x, y)]^{\gamma}, \quad (4.3)$$

де x і y – відповідні області оригінального та спотвореного зображень; l(x, y) – складова порівняння яскравості; c(x, y) – складова порівняння контрасту; s(x, y) – складова порівняння структури; α, β, γ – вагові коефіцієнти.

Складова яскравості порівнює середні значення двох областей. Складова контрасту порівнює середньоквадратичні відхилення. Структурна складова використовує коваріацію, тобто показує, наскільки подібно змінюються відповідні області двох зображень. Значення SSIM зазвичай інтерпретується так: чим ближче воно до 1, тим більш подібними є зображення. Якщо значення зменшується, це означає, що між оригінальним і спотвореним зображенням з’явилися помітні структурні відмінності.

4.4 Локальне обчислення SSIM та карта якості

На практиці SSIM обчислюється локально, тобто не для всього зображення одразу, а для невеликих вікон. Це важливо, тому що різні частини зображення можуть мати різний рівень спотворень. Наприклад, одна область може залишатися майже незмінною, а інша може містити блокові артефакти JPEG або втрату текстур. Після обчислення локальних значень SSIM можна сформувати карту якості. Світлі області такої карти відповідають високій подібності, а темні — сильнішим структурним спотворенням. Карта якості дозволяє не тільки отримати числову оцінку, а й побачити, у яких частинах зображення виникають найбільші зміни. Для отримання одного загального значення якості використовується середнє значення локальних оцінок – MSSIM. Формула (4.4), [1] а описує середнє значення локальних оцінок SSIM

$$MSSIM(X,Y) = \frac{1}{N}\sum_{j=1}^{N_s} SSIM(x_j, y_j), \quad (4.4)$$

де X і Y – повні зображення; x_j і y_j – відповідні локальні області; N_s – кількість локальних областей; SSIM(xj, yj) – локальне значення SSIM.

Таким чином, SSIM дозволяє оцінювати якість зображення з урахуванням локальних структурних змін. Це робить його корисним інструментом для аналізу спотворень після стиснення.

4.5 Обмеження SSIM при оцінюванні стиснених зображень

Незважаючи на переваги SSIM, його не можна розглядати як універсально кращу метрику. У коментарях до звіту з практики було зазначено, що для стиснених зображень SSIM не завжди показує себе краще за PSNR. Це важливе зауваження для даної дипломної роботи.

Dosselmann і Yang [2] показують, що SSIM має тісний зв’язок із MSE. Автори звертають увагу, що в багатьох випадках SSIM і MSE можуть давати подібне ранжування спотворених зображень. Це означає, що структурний підхід не завжди забезпечує принципово іншу або кращу оцінку якості. Для JPEG-зображень це особливо важливо. JPEG створює характерні спотворення: блокові артефакти, втрату високочастотних деталей, розмиття текстур і зміни біля контурів. Теоретично SSIM повинен добре реагувати на такі структурні зміни, але на практиці це потрібно перевіряти експериментально.

Тому в подальшому розділі доцільно не просто описати SSIM, а порівняти його з PSNR та MSE на стиснених зображеннях. Саме це дозволить визначити, чи є SSIM ефективним показником якості для JPEG-стиснення, чи його переваги в цьому випадку обмежені.

Таблиця 4.1 — Порівняння основних метрик оцінювання якості зображень

Метрика	Що оцінює	Переваги	Обмеження
MSE	середню квадратичну різницю між пікселями	проста, швидко обчислюється	не враховує структуру зображення
PSNR	відношення сигнал/шум через MSE	зручна для порівняння стиснення	не завжди відповідає сприйняттю людини
SSIM	яскравість, контраст і структуру	враховує локальні структурні зміни	не завжди має перевагу над MSE/PSNR для стиснених зображень
MSSIM	середнє локальних значень SSIM	дає загальну оцінку всього зображення	результат залежить від способу локального усереднення

Отже, основні метрики якості мають різне призначення. MSE та PSNR є простими і зручними, але не враховують структуру. SSIM враховує структурну подібність, однак його ефективність для JPEG-зображень потребує окремої перевірки. Саме тому наступний розділ має бути присвячений аналізу експериментальних результатів і порівнянню поведінки SSIM з традиційними метриками.

5 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

5.1 Мета експериментального аналізу

У попередніх розділах було розглянуто теоретичні основи JPEG-стиснення, традиційні метрики оцінювання якості зображень та індекс структурної подібності SSIM. Однак для теми даної роботи важливо не тільки описати SSIM, а й критично оцінити його ефективність під час аналізу зображень, стиснених із втратами. У статті Dosselmann і Yang [2] виконується детальне порівняння SSIM із традиційною середньоквадратичною помилкою MSE. Автори не просто розглядають SSIM як сучасну метрику, а перевіряють, наскільки її результати справді відрізняються від результатів MSE. Це важливо для даної дипломної роботи, оскільки JPEG-стиснення створює характерні структурні спотворення, але це не означає, що SSIM завжди буде кращим показником якості. Основна мета експериментального аналізу полягає у визначенні того, чи дає SSIM принципово іншу оцінку якості порівняно з MSE, чи ці метрики часто поводяться подібно. Якщо SSIM і MSE однаково ранжують спотворені зображення, то практична перевага SSIM для оцінювання JPEG-зображень може бути обмеженою. На рисунку 5.1 наведено приклад із роботи Dosselmann і Yang [2], який демонструє проблему оцінювання якості за однаковим або близьким значенням MSE. Різні типи спотворень мають подібний рівень середньоквадратичної помилки, але візуально сприймаються по-різному. Особливо важливим для цієї роботи є приклад JPEG-стиснення, оскільки він показує, що одна числова метрика не завжди достатньо точно описує характер втрати якості. На рисунку 5.1 показан приклад зображень де значеня MSE доволі близькі але з різним візулаьним сприйняттям якості.

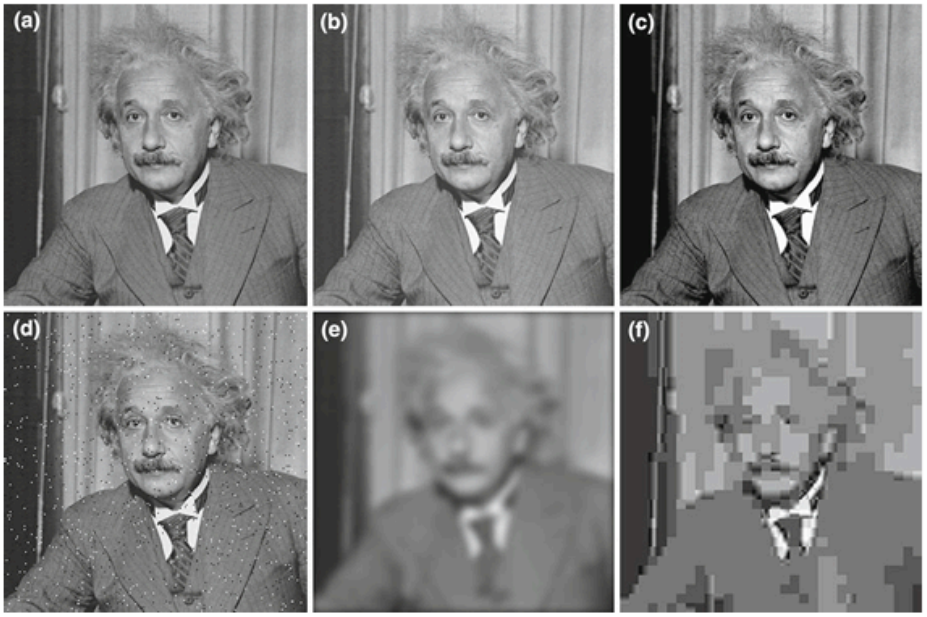


Рисунок 5.1 – Приклад зображень із близьким значенням MSE, але різним візуальним сприйняттям якості; a - оригінальне зображення; b - зміна яскравості; c - зміна контрасту; d - шум; e - гаусове розмиття; f - JPEG-стиснення.

5.2 Методика експериментів Dosselmann і Yang

Dosselmann і Yang [2] використали набір із 15 класичних тестових зображень. Ці зображення відрізнялися за текстурою, деталізацією, освітленням і змістом сцени. Для кожного зображення формувалося 5 наборів спотворених варіантів. У кожному наборі було 10 спотворених зображень. Таким чином, усього було отримано 75 тестових наборів. Для кожного спотвореного зображення автори обчислювали значення MSE та SSIM. Після цього зображення в межах одного набору ранжувалися за якістю. Найгірше зображення отримувало найнижчий ранг, а найкраще — найвищий. Далі порівнювалося, наскільки подібно MSE і SSIM впорядковують одні й ті самі зображення. Набір початкових тестових зображень, використаний авторами, показано на рисунку 5.2. Він містить зображення з різним рівнем деталізації, текстурності та контрасту. Це важливо, оскільки поведінка метрик якості може залежати від типу сцени: для гладких областей, текстур і зображень із великою кількістю контурів результати можуть відрізнятися.



Рисунок 5.2 – Початкові тестові зображення, використані в експериментах Dosselmann і Yang

Для порівняння ранжування використовувався коефіцієнт рангової кореляції Спірмена. У статті на стор. 85 описано коефіцієнт Спірмена, але сама формула там не розписана. Для вставлення можна використати стандартну формулу (5.1), яка описує коефіцієнт рангової кореляції Спірмена, який використовується для порівняння порядку зображень за MSE та SSIM

$$r_s = 1 - \frac{6\sum_{i=1}^{n} d_i^2}{n(n^2-1)}, \tag{5.1}$$

де r – коефіцієнт рангової кореляції; значення r, близьке до 1, означає

сильну подібність ранжування; значення r, близьке до 0, означає слабкий зв’язок між ранжуваннями; від’ємне значення r означає протилежний характер ранжування.

У статті [2] використовувалися різні типи спотворень: шум, розмиття, JPEG-стиснення, JPEG2000-стиснення, різкість та квантування. Це дозволило перевірити поведінку SSIM не тільки для одного типу дефектів, а для різних ситуацій.

Таблиця 5.1 — Методика експериментального дослідження за Dosselmann і Yang [2]

Параметр	Значення
Кількість початкових зображень	15
Кількість наборів для кожного зображення	5
Кількість спотворених зображень у наборі	10
Загальна кількість наборів	75
Метрики для порівняння	MSE та SSIM
Спосіб порівняння	ранжування зображень за якістю
Показник зв’язку	коефіцієнт рангової кореляції Спірмена
Типи спотворень	шум, розмиття, JPEG, JPEG2000, квантування, підвищення різкості

5.3 Результати рангового порівняння SSIM і MSE

Результати експериментів показали високий рівень зв’язку між MSE та SSIM. У статті [2] наведено значення коефіцієнта Спірмена для 75 тестових наборів. Значення кореляції змінювалися від 0,6364 до 1,0000. Середнє значення становило 0,9116, а дисперсія — 0,007. Це означає, що в багатьох випадках MSE та SSIM ранжували спотворені зображення подібним чином. Іншими словами, хоча SSIM має складнішу математичну модель і враховує структуру, його практична поведінка часто наближалася до поведінки MSE.

Таблиця 5.2 — Основні результати рангового порівняння SSIM і MSE

Показник	Значення
Мінімальне значення кореляції	0,6364
Максимальне значення кореляції	1,0000
Середнє значення кореляції	0,9116
Дисперсія	0,007
Загальна кількість коефіцієнтів кореляції	75
Основний висновок	SSIM і MSE часто дають подібне ранжування зображень

Окремо автори розглянули приклад для зображення “mandrill”. У цьому випадку коефіцієнт кореляції дорівнював 0,9152, тобто був близьким до середнього значення по всьому експерименту. Це підтвердило загальну тенденцію: SSIM та MSE часто оцінюють порядок якості спотворених зображень схожим чином.

Водночас між метриками були й певні відмінності. Автори зазначають, що MSE сильніше реагує на шум, тоді як SSIM може бути чутливішим до розмиття. Це пояснюється локальним характером обчислення SSIM та тим, що розмиття безпосередньо змінює локальний контраст і структуру зображення.

5.4 Формальний зв’язок між SSIM та MSE

Крім емпіричного аналізу, Dosselmann і Yang [2] виконали формальний математичний аналіз зв’язку між SSIM та MSE. Для цього вони розглянули спрощений варіант SSIM без стабілізаційних констант. Такий підхід дозволив показати, що між SSIM і локальною MSE існує функціональна залежність. Формула (5.2), [2] описує спрощений варіант SSIM без стабілізаційних констант C1, C2 та C3

$$SSIM^{*}(x,y) = \left(\frac{2\mu_x\mu_y}{\mu_x^2+\mu_y^2}\right)\left(\frac{2\sigma_x\sigma_y}{\sigma_x^2+\sigma_y^2}\right)\left(\frac{\sigma_{xy}}{\sigma_x\sigma_y}\right), \quad (5.2)$$

де x і y – локальні області оригінального та спотвореного зображень; μx, μy – середні значення яскравості; σx, σy – середньоквадратичні відхилення; σxy – коваріація між областями.

Після перетворення цієї формули автори отримали вирази, які дозволяють перейти від локального SSIM до локальної MSE і навпаки. Це є важливим результатом, оскільки він показує, що SSIM не є повністю незалежною від MSE метрикою. Формули (5.3) та (5.4), [2] описують функціональний зв’язок між локальним SSIM та локальною MSE

$$MSE^{*} = \frac{k_1(l_1+l_2-k_1)SSIM^{*}(x,y)}{k_2} + l_1 + l_2 - k_2 \quad , \quad (5.3)$$

$$SSIM\ *\ (x,y)\ =-\ \frac{k_2(MSE^*-l_1-l_2+k_2)}{k_1(l_1\ +l_2-k_1)}, \tag{5.4}$$

де MSE* — локальне значення середньоквадратичної помилки; SSIM* — спрощене локальне значення SSIM; l_1 , l_2 , k_1, k_2— допоміжні величини, пов’язані з яскравістю та енергією локальних областей.

Автори підкреслюють, що найбільший вплив на відмінність між SSIM та MSE має величина l2, яка пов’язана із сумою квадратів значень пікселів спотвореного зображення. Однак ця величина також не показує, де саме розташоване спотворення і наскільки воно помітне для людини. Зв’язок між MSE та SSIM автори також показують графічно. На рисунку 5.3 наведено залежність між значеннями MSE та SSIM для зображень “Goldhill” і “man”. Точки на графіках утворюють окремі послідовності, що відповідають різним типам спотворень. Це підтверджує, що між метриками існує помітний зв’язок, але характер цього зв’язку залежить від виду дефекту.

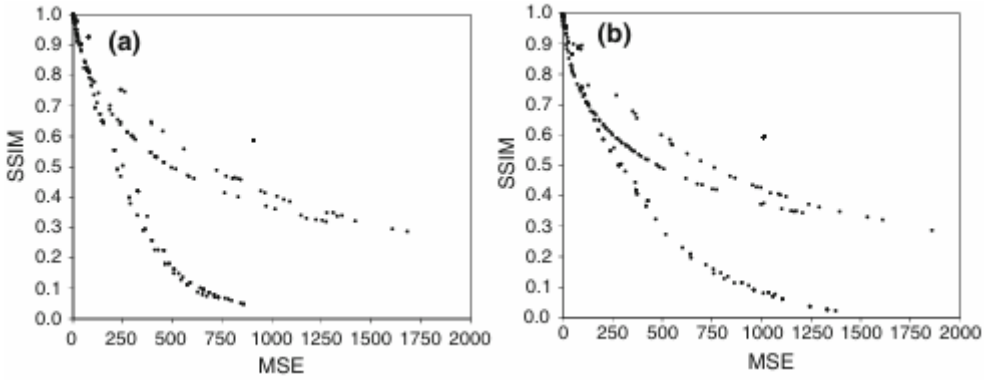


Рисунок 5.3 – Графіки залежності між MSE та SSIM для тестових зображень “Goldhill” і “man”

Отже, навіть формальний аналіз підтверджує, що SSIM не повністю вирішує проблему відповідності автоматичної метрики людському сприйняттю.

5.5 Значення результатів для оцінювання JPEG-зображень

Для даної дипломної роботи найважливішим є висновок про поведінку SSIM під час оцінювання зображень, стиснених із втратами. JPEG-стиснення створює характерні артефакти: блокові межі, втрату дрібних деталей, розмиття текстур і спотворення біля контурів. Теоретично SSIM повинен краще реагувати на такі структурні зміни, ніж MSE або PSNR. Однак результати Dosselmann і Yang [2] показують, що практична перевага SSIM не завжди є очевидною. У багатьох випадках SSIM і MSE дають схоже ранжування зображень. Крім того, у статті зазначено, що для випадків стиснення та квантування між ранжуваннями MSE і SSIM не спостерігалося суттєвих розбіжностей.

Це має пряме значення для оцінювання JPEG-зображень. Якщо SSIM та MSE подібно реагують на JPEG-стиснення, то використання SSIM не завжди дає нову інформацію порівняно з традиційними метриками. Тому SSIM доцільно застосовувати не як єдиний показник якості, а разом із PSNR, MSE та візуальним аналізом артефактів. Результати дослідження були занесині в таблицю 5.3.

Таблиця 5.3 – Практичне значення результатів Dosselmann і Yang для JPEG-стиснення, частина 1

1. Спостереження	2. Значення для дипломної роботи
SSIM і MSE мають високий середній рівень кореляції	SSIM не можна автоматично вважати принципово кращою метрикою
SSIM чутливіший до розмиття	Метрика може бути корисною для аналізу втрати чіткості
MSE сильніше реагує на шум	Для шумових спотворень традиційна метрика може бути не гіршою

Продовження таблиці 5.3

1. Спостереження	2. Значення для дипломної роботи
Для стиснення розбіжності між SSIM і MSE не є суттєвими	Для JPEG потрібне окреме порівняння SSIM з PSNR/MSE
SSIM має математичний зв’язок із MSE	Метрика не є повністю незалежною альтернативою MSE

Отже, результати експериментальних досліджень підтверджують зауваження, зроблене у попередніх розділах: SSIM є корисною метрикою, але її не можна розглядати як універсально найкращий спосіб оцінювання якості. Для JPEG-зображень доцільно застосовувати комплексний підхід, у якому SSIM порівнюється з PSNR, MSE та візуальною оцінкою характерних артефактів. Загальний висновок за результатами статті Dosselmann і Yang полягає в тому, що між SSIM та MSE існує як статистичний, так і формальний зв’язок. Це означає, що SSIM може давати корисну інформацію про структурні зміни, але його результати часто залишаються близькими до результатів традиційних метрик. Саме тому при аналізі ефективності SSIM для стиснених із втратами зображень необхідно перевіряти його поведінку експериментально, а не виходити лише з теоретичних переваг структурного підходу.

ВИСНОВКИ

У кваліфікаційній роботі було досліджено ефективність застосування індексу структурної подібності SSIM для оцінювання якості зображень, стиснених із втратами. Основну увагу було приділено JPEG-стисненню, оскільки цей формат широко використовується для зберігання та передавання фотографічних зображень і створює характерні візуальні спотворення.

У першому розділі було розглянуто цифрові зображення як об’єкти обробки та оцінювання якості. Було показано, що зображення можна подати як двовимірний дискретний сигнал або як матрицю пікселів. Також було проаналізовано традиційні метрики якості, зокрема MSE та PSNR. Їхньою перевагою є простота обчислення, однак вони оцінюють переважно числову різницю між відповідними пікселями і не завжди враховують особливості людського зорового сприйняття.

У другому розділі було досліджено ортогональні перетворення, які застосовуються при стисненні зображень. Особливу увагу було приділено дискретному косинусному перетворенню DCT. Було встановлено, що DCT дозволяє перетворити блок пікселів у набір частотних коефіцієнтів, у якому основна енергія зображення зосереджується переважно в низькочастотній області. Саме ця властивість робить DCT ефективною основою для JPEG-стиснення.

У третьому розділі було розглянуто стиснення зображень із втратами та формат JPEG. Було показано, що baseline JPEG використовує перетворення кольору, поділ зображення на блоки 8×8, DCT, квантування, zig-zag порядок обходу коефіцієнтів та ентропійне кодування. Основним джерелом втрат є квантування DCT-коефіцієнтів. Саме воно дозволяє суттєво зменшити розмір файлу, але водночас спричиняє втрату деталей, розмиття та блокові артефакти.

У четвертому розділі було розглянуто основи оцінювання якості зображень та індекс структурної подібності SSIM. Було встановлено, що SSIM

оцінює подібність двох зображень за трьома складовими: яскравістю, контрастом і структурою. Це робить його більш наближеним до структурного аналізу зображення, ніж прості піксельні метрики. Водночас було підкреслено, що SSIM не можна автоматично вважати універсально кращим за MSE або PSNR, особливо для зображень, стиснених із втратами.

У п’ятому розділі було проаналізовано результати експериментальних досліджень Dosselmann і Yang. Було показано, що між SSIM та MSE існує високий рівень статистичного зв’язку. У дослідженні середнє значення рангової кореляції між цими метриками становило 0,9116, що свідчить про подібне ранжування спотворених зображень у багатьох випадках. Також було розглянуто формальний зв’язок між спрощеним SSIM і локальною MSE. Отримані результати дозволяють зробити висновок, що SSIM є корисною метрикою для аналізу структурних змін у зображеннях, але його переваги над традиційними метриками не є абсолютними. Для JPEG-зображень SSIM може бути чутливим до втрати структури, розмиття та блокових артефактів, однак у багатьох випадках його результати можуть бути близькими до MSE або PSNR.Таким чином, для оцінювання якості зображень, стиснених із втратами, доцільно застосовувати комплексний підхід. SSIM варто використовувати разом із MSE, PSNR та візуальним аналізом характерних артефактів JPEG-стиснення. Такий підхід дозволяє отримати більш повне уявлення про якість відновленого зображення і не покладатися лише на одну метрику.

Отже, мету роботи було досягнуто: проаналізовано принципи JPEG-стиснення, розглянуто математичні та практичні основи SSIM, досліджено його зв’язок із традиційними метриками та визначено обмеження його застосування для оцінювання якості зображень, стиснених із втратами.

ПЕРЕЛІК ПОСИЛАНЬ

1. Acharya T., Ray A. IMAGE PROCESSING Princiles and Applications : підручник. New Jersey : John Wiley & Sons, 2005. 428 с.

2. Dosselman R., Yang X. D. Acomprehensive assessment of the structural similarity index. *Springer-Verlag London Limited*. 2009. Т. 5. С. 81–91.

3. T. Acharya and A. Mukherjee, “High-speed Parallel VLSI Architectures for Image Decorrelation,” International Journal of Pattern Recognition and Artificial Intelligence, 9(2), 1995, 343-365.

4. T. Acharya and P. Tsai, JPEG2000 Standard for Image Compression: Concepts, Algorithms and VLSI Architectures, Wiley, Hoboken, NJ, 2004.

5. W. B. Pennebaker and J. L. Mitchell, JPEG Still Image Data Compression Staadard. Chapman & Hall, New York, 1993.

6. H. Lohscheller, “A Subjectively Adapted Image Communication System,” IEEE Transactions on Communications, COM-32, Vol. 12, 1984. 1316- 1322.

7. C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, Urbana, IL, 1949.

8. B. McMillan, “The Basic Theorems of Information Theory,” Ann. Math. Statist., Vol. 24, 1953, 196-219.

9. R. M. Fano, Transmission of Information, MIT Press, Cambridge, MA, 1949.

10. I. H. Witten, R. M. Neal, and J. G. Cleary, “Arithmetic Coding for Data Compression,” Communications of the ACM, 30(6), June 1987.