

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
(повна назва)

Кафедра Інформаційно-мережної інженерії
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА Пояснювальна записка

рівень вищої освіти перший (бакалаврський)

Оцінювання якості зображення на основі індексу структурної подібності SSIM

(тема)

Виконав:

здобувач 3 року навчання,
групи ТРИМІу-22-1

Артеменко Микита

(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

Завідувач кафедри ІМІ

(підпис)

Валерій Безрук

(власне ім'я, прізвище)

2025 р.

Не містить відомостей, заборонених до відкритого публікування

Студент _____ / Артеменко М.О. /
(підпис) (прізвище та ініціали)

Керівник _____ / Федоров О.В /
(підпис) (прізвище та ініціали)

Харківський національний університет радіоелектроніки

Факультет Інфокомунікацій
Кафедра Інформаційно мережної інженерії
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 Артеменко Микиті Олександровичу
(прізвище, ім'я, по батькові)

- Тема роботи Оцінювання якості зображення на основі індексу структурної подібності SSIM
затверджена наказом університету від 23 травня 2025 р. № 411Ст
- Термін подання здобувачем роботи до екзаменаційної комісії 16 липня 2025 р.
- Вихідні дані до роботи Дослідити доцільність використання індексу структурної подібності SSIM в задачах оцінювання якості стиснених із втратами зображень, зокрема JPEG зображень. З'ясувати чутливість індексу SSIM до різних типів шуму, а саме шуму компресії та адитивного білого гаусового шуму. Виконати порівняльний аналіз двох метрик якості зображень: індексу SSIM та PSNR з використанням тестових зображень. Зробити висновки.
- Перелік питань, що потрібно опрацювати в роботі _____
Вступ
1 Особливості системи зору людини
2 Стандарт JPEG кодування сталих зображень
3 Структурні підходи до оцінки якості зображення
4 Індекс структурної подібності
5 Оцінювання якості зображень з використанням індексу структурної подібності SSIM
6 Результати експериментальних досліджень
Висновки

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.	26.05.2025	вик.
2.	Підбір літератури за темою роботи	26.05 – 13.06.2025	вик.
3.	Виконання розділу 1	28.05 – 15.06.2025	вик.
4.	Виконання розділу 2	05.06 – 20.06.2025	вик.
5.	Виконання розділу 3	16.06 – 05.07.2025	вик.
6.	Виконання розділу 4	20.06 – 06.07.2025	вик.
7.	Виконання розділу 5	30.06 – 08.07.2025	вик.
8.	Виконання розділу 6	20.06 – 10.07.2025	вик.
9.	Оформлення презентаційного матеріалу, підготовка до захисту в ЕК	05.07 – 15.07.2025	вик.

Дата видачі завдання 26 травня 2025р.

Здобувач _____
(підпис)

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

РЕФЕРАТ

Пояснювальна записка: 75 с., 29 рис., 2 таб., 26 джерел, 1 додаток.

Об'єктом дослідження є методи оцінювання якості спотворених зображень.

Мета кваліфікаційної роботи – дослідження методів оцінювання якості спотворених зображень. Порівняльний аналіз таких метрик якості зображень як PSNR та індекс SSIM.

В роботі було виконано порівняльний аналіз двох метрик якості зображень: індексу SSIM та PSNR. Розглянуто два типи спотворень: шум JPEG компресії і адитивний білий гаусів шум. Індекс SSIM виявився більш чутливим до спотворень у вигляді адитивного білого гаусового шуму, хоча шуми компресії є візуально більш помітними. Останнє свідчить на користь традиційної метрики PSNR, яка крім того є значно простішою в обчислювальному плані.

ЯКІСТЬ, СПОТВОРЕННЯ, КОМПРЕСІЯ, ІНДЕКС, SSIM, JPEG, PSNR, ШУМ, ОБЧИСЛЕННЯ

ABSTRACT

Explanatory note: 75 p., 29 fig., 2 tabs, 26 sources, 1 app.

Object of study is methods for assessing the quality of distorted images.

Purpose of the qualification work is to study methods for assessing quality of distorted images. Comparative analysis of such image quality metrics as PSNR and SSIM index.

The paper presents a comparative analysis of two image quality metrics: SSIM and PSNR. Two types of distortion are considered: JPEG compression noise and additive white Gaussian noise. The SSIM index proved to be more sensitive to distortions in the form of additive white Gaussian noise, although compression noise is more visually noticeable. The latter is in favour of the traditional PSNR metric, which is also much simpler to compute.

QUALITY, DISTORTION, COMPRESSION, INDEX, SSIM, JPEG, PSNR, NOISE, COMPUTATION

ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ	8
ВСТУП.....	9
1 ОСОБЛИВОСТІ СИСТЕМИ ЗОРУ ЛЮДИНИ	10
1.1 Будова зорової системи людини.....	10
1.2 Коефіцієнт контрастності.....	11
1.3 Просторова частота	13
1.4 Ефект маскування.....	16
1.5 Смуги Маха.....	16
2 СТАНДАРТ JPEG КОДУВАННЯ СТАЛИХ ЗОБРАЖЕНЬ	18
2.1 Стандарт JPEG	18
2.2 Стандарт JPEG 2000	22
3 СТРУКТУРНІ ПІДХОДИ ДО ОЦІНКИ ЯКОСТІ ЗОБРАЖЕННЯ.....	27
4 ІНДЕКС СТРУКТУРНОЇ ПОДІБНОСТІ	32
5 ОЦІНЮВАННЯ ЯКОСТІ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ ІНДЕКСУ СТРУКТУРНОЇ ПОДІБНОСТІ SSIM	41
6 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ	47
ВИСНОВКИ	60
ПЕРЕЛІК ПОСИЛАНЬ	61
ДОДАТОК А СЛАЙДИ ПРЕЗЕНТАЦІЇ	64

ПЕРЕЛІК СКОРОЧЕНЬ

- JND – just noticeable difference
HVS – human visual system
MTF – modulation transfer function
DCT – discrete cosine transform
DC – direct current
AC – alternating current
RLE - run-length encoding
QC - quantization coefficient
QF – quality factor
ЛПР – ледь помітна різниця
СЗР – система зору людини
МПФ – модуляційно-передатна функція
ДКП – дискретне косинусне перетворення

ВСТУП

Сигнали, що відповідають цифровим зображенням зазвичай подаються як двовимірні (2D) масиви дискретних відліків сигналу. Якщо ми переставимо відліки сигналу в одновимірний (1D) вектор, то кожне зображення стане єдиною точкою у багатовимірному *просторі зображення*, розмірність якого дорівнює кількості дискретів у сигналі зображення. Було зазначено, що кластер *природних* сигналів зображення займає надзвичайно малу частину такого простору зображення [15, 16]. Протягом своєї тривалої еволюції та розвитку система зору людини (СЗЛ) була значною мірою піддана впливу природного *візуального* середовища, і різноманітні дані показали, що СЗЛ добре адаптована для вилучення корисної інформації з природних сцен [17]. Метрика якості зображення, яка має на меті передбачити поведінку СЗЛ при оцінці якості, також потребує «адаптації» до властивостей природних сигналів зображення.

Дана кваліфікаційна робота присвячена вивченню оцінювання якості зображення на основі індексу структурної подібності SSIM [19].

1 ОСОБЛИВОСТІ СИСТЕМИ ЗОРУ ЛЮДИНИ

1.1 Будова зорової системи людини

Критичною проектною метою системи кодування цифрового відео/зображень є те, щоб відео/зображення, створені системою, були прийнятними та приємними для глядача, тобто кінцевим спостерігачем є людське око [1]. Для досягнення цієї мети необхідно враховувати реакцію системи зору людини (СЗЛ). СЗЛ — це система, за допомогою якої людина спостерігає, інтерпретує та реагує на візуальні стимули. Спрощений поперечний розріз людського ока показано на рисунку 1.1 [1]. Зображення фокусується на поверхні сітківки кришталиком, який змінює форму під контролем м'язів, щоб забезпечити належне фокусування на близьких і віддалених об'єктах. Райдужка контролює діафрагму кришталика, а отже, кількість світла, що потрапляє в око. Сітківка складається з масиву колбочок (фоторецепторів, чутливих до кольору при високих рівнях освітлення) і паличок (фоторецепторів, чутливих до яскравості при низьких рівнях освітлення). Більш чутливі колбочки сконцентровані в центральній області (фовеа), що означає, що кольоровий зір високої роздільної здатності досягається лише на невеликій ділянці в центрі поля зору [1]. Нерви, що з'єднуються з сітківкою, виходять з очного яблука через зоровий нерв. Людський мозок обробляє та інтерпретує візуальну інформацію, частково на основі отриманої інформації (зображення, виявлене сітківкою) і частково на основі попередньо засвоєних відповідей (таких як відомі форми об'єктів) [1].

Кінцевим оцінювачем якості зображення є людина-спостерігач. Отже, вигідно та важливо вбудовувати СЗЛ в програми обробки зображень. Таким чином, оцінку якості зображення можна суб'єктивно покращити, включивши модель СЗЛ з такими параметрами, як орієнтація

та кут поля, що впливають на сприйняття людського ока. Робота СЗЛ - це велика та складна галузь дослідження. Деякі з важливих особливостей СЗЛ наведені нижче [1].

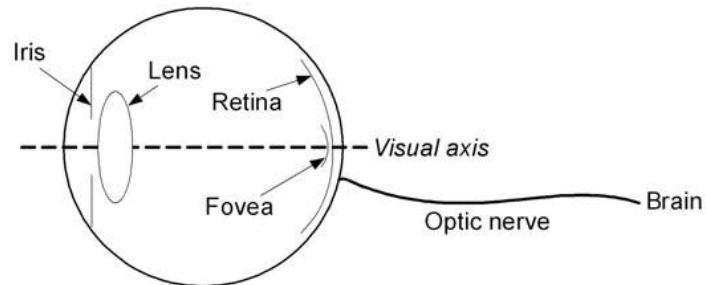


Рисунок 1.1 - Спрощений поперечний розріз людського ока

1.2 Коефіцієнт контрастності

Відомо, що реакція ока на зміни інтенсивності освітлення є нелінійною. Розглянемо ділянку світлової інтенсивності, оточену фоном інтенсивністю I . Ледь помітна різниця (just noticeable difference, JND), ΔI , інтенсивності цілі (об'єкта) має бути функцією I . Зокрема в [2] мова йде про те, що чутливість людських очей до розрізнення відмінностей в інтенсивності залежить не тільки від самої різниці, але й від рівня інтенсивності. В широкому діапазоні інтенсивностей виявлено, що співвідношення $\Delta I / I$, яке носить назву коефіцієнта Вебера (Weber fraction), є майже постійним зі значенням близьким до 0,02 [3, 4]. Цей результат не діє при дуже низьких або дуже високих інтенсивностях, де JND помітно зростає, як показано на рис. 1.2 [1]. Загальний вираз $(I + \Delta I) / I$ представляє коефіцієнт контрастності, який є постійним у деякому діапазоні інтенсивностей.

В роботі [5] для цифрових зображень з роздільною здатністю 8 bpr (тобто, 8 біт на піксель) було промодельовано властивість незмінності співвідношення Вебера та отримано такі результати: СЗЛ є більш

чутливою на ближчих відстанях до монітора для фонові інтенсивності I рівнів сірого з діапазону від 0 до 50 (зона низьких рівнів сірого). Однак відстань перегляду не впливає на значення JND для фонові інтенсивності вищої за 50. Щодо впливу умов навколишнього освітлення, СЗЛ є більш чутливою, коли в кімнаті темно, для зони низьких рівнів сірого, і вона не залежить від умов освітлення для зони високих рівнів сірого. Щодо впливу яскравості монітора, СЗЛ є більш чутливою до змін при високій яскравості монітора і менш чутливою при низькій яскравості [1].

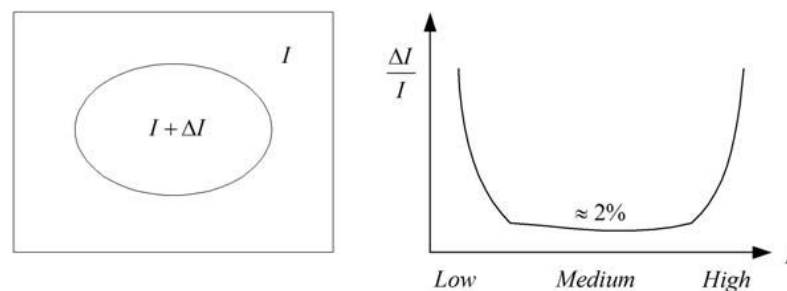


Рисунок 1.2 - Вимірювання чутливості до контрасту

Порівняно зі співвідношенням Вебера, яке є постійним у середньому діапазоні інтенсивності та зростає в нижньому та верхньому діапазонах, ці експериментальні результати показують, що значення JND залишаються майже незмінними в досить широкому діапазоні рівнів сірого від 50 до 255 (лише невелике збільшення спостерігається для значень від 235 до 255). Це означає, що динамічний діапазон яскравості монітора суттєво не розширюється, навіть якщо рівень сірого є максимальним (255). Експерименти з коефіцієнтом контрастності виявили, що СЗЛ стає нечутливою до похибки квантування в нижньому діапазоні яскравості, а в середньому діапазоні - є чутливою [1].

1.3 Просторова частота

Загалом вимірювання порогу яскравості як функції просторової частоти були проведені багатьма дослідниками, і деякі плідні результати зведено в [6]. Зокрема встановлено, що пік *модуляційно-передатної функції* (МПФ, *modulation transfer function*, MTF) становить близько 6,2 циклів/град. МПФ визначається формулою

$$F_{MTF}(f_r) = 2,6(0,0192 + 0,144 f_r) e^{-(0,114 f_r)^{1,1}} \quad (1.1)$$

де f_r позначає *просторову частоту*.

МПФ забезпечує легкий спосіб інтеграції чутливості СЗЛ в додатки, що здійснюють кодування зображень. Приклади використання МПФ включають перетворення зображення в область сприйняття перед кодуванням, оцінку якості зображення та визначення міри спотворення, що використовується при обчисленні бітрейту (бітової швидкості) [1].

Просторова частота позначає частоту, нормалізовану за кутом огляду, як на рис. 1.3, що показує горизонтальне та вертикальне спостереження з постійної відстані, скажімо, у шість разів більшої за вертикальну висоту. Тоді кути огляду обчислюються як [1]

$$\theta_h = 2 \arctan\left(\frac{H/2}{6V}\right), \quad (1.2)$$

$$\theta_v = 2 \arctan\left(\frac{V/2}{6V}\right). \quad (1.3)$$

Максимальні просторові частоти в горизонтальному та вертикальному напрямках визначаються як [1]

$$f_{h,\max} = \frac{H/2}{\theta_h}, \quad (1.4)$$

$$f_{v,\max} = \frac{V/2}{\theta_v}. \quad (1.5)$$

та вимірюються в одиницях циклів/градус.

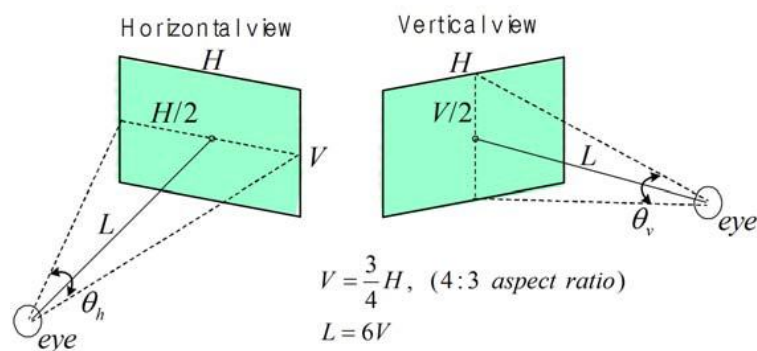


Рисунок 1.3 - Умови спостереження та концепція просторової частоти

Далі, просторова частота об'єкта визначається в горизонтальному напрямку $f_{h,VO}$ та у вертикальному напрямку $f_{v,VO}$ як [1]

$$f_{h,VO} = \frac{f_{h,\max}}{H_{VO}/2}, \quad (1.6)$$

$$f_{v,VO} = \frac{f_{v,\max}}{V_{VO}/2}, \quad (1.7)$$

де H_{VO} та V_{VO} позначають горизонтальний та вертикальний розміри об'єкта відповідно.

Радіальна частота обраховується за допомогою просторових частот в обох напрямках як [1]

$$f_{r,VO} = \sqrt{f_{h,VO}^2 + f_{v,VO}^2} \quad (1.8)$$

Відповідна чутливість об'єкта може бути розрахована за допомогою емпірично виведеної МПФ, як в (1.1).

Матриці квантування використовуються для квантування коефіцієнтів перетворення в JPEG (див. розділ 2) та MPEG. Матриці мають невід'ємні елементи, які збільшуються в горизонтальному та вертикальному напрямках. Коефіцієнти перетворення діляться на значення елементів матриці квантування. Більші значення елементів матриці квантування призводять до грубого квантування високочастотних коефіцієнтів перетворення, тоді як низькочастотні коефіцієнти квантуються з меншими кроками, тобто діляться на менше число. Отже, міжнародні стандарти більш-менш використовують концепцію СЗЛ у кодері [1].

Той факт, що людське око не однаково чутливе до спотворених в наслідок кодування зображень або відео сигналів, можна пояснити з точки зору концепції просторової частоти СЗЛ. Наприклад, припускаючи, що щільність дискретизації становить 64 пікселі/град і зображення має розмірність 512×512 пікселів, найнижча просторова частота в блоці 8×8 становить близько 4 циклів/град, що знаходиться в найчутливішій смузі МПФ. Враховуючи, що блочний ефект кодера на основі ДКП (коли відбувається розбиття зображення на блоки 8×8 пікселів) є домінуючим з великими спотвореннями при малих значеннях показника якості (тобто при високих степенях компресії), концепція просторової частоти відіграє

роль в поясненні того, чому людське око виявляється таким чутливим до цього ефекту [1].

1.4 Ефект маскування

Коли до складу зображення входять області з великою кількістю деталей, тоді спостерігається втрата чутливості ока людини до наявності помилок в цих областях. Це отримало назву ефекту маскування, що з'являється внаслідок присутності ділянок на зображенні із суттєвою кількістю деталей. Дослідники намагалися кількісно оцінити величину маскування, використовуючи різні функції помітності для вимірювання чутливості ока до спотворень [7, 8]. Через зміни відстані перегляду, частоти дискретизації та величезні відмінності у вмісті зображення, суб'єктивно виміряні пороги не застосовуються глобально у всіх випадках [1]. Адаптація до локального вмісту зображення, використовуючи більш складні функції активності для вимірювання функції маскування, була введена в [9]. Ці функції маскування використовуються при розробці квантувачів для досягнення малих значень бітрейту.

1.5 Смуги Маха

Зображення, що містять різкі краї, створюють великі помилки передбачення в предиктивному кодуванні або великі коефіцієнти перетворення в трансформаційному кодуванні, що є основними перешкодами в кодері, які не дозволяють досягати високого стиснення [1]. Смуги Маха (названі на честь австрійського фізика Е. Маха) часто використовуються для пояснення проблеми, що визначається як ілюзія яка підкреслює наявність на зображенні границь в той час як насправді інтенсивність змінюється плавно [10]. Рисунок 1.4 показує набір

вертикальних сірих смуг з вісьмома краями. У кожній смузі рівні сірого є постійними.

На межі між двома краями, права сторона кожної смуги виглядає трохи темнішою, ніж середина смуги, а ліва сторона виглядає трохи світлішою. Перехід від однієї смуги до іншої підкреслюється чутливістю зорової системи людини до інтенсивності.

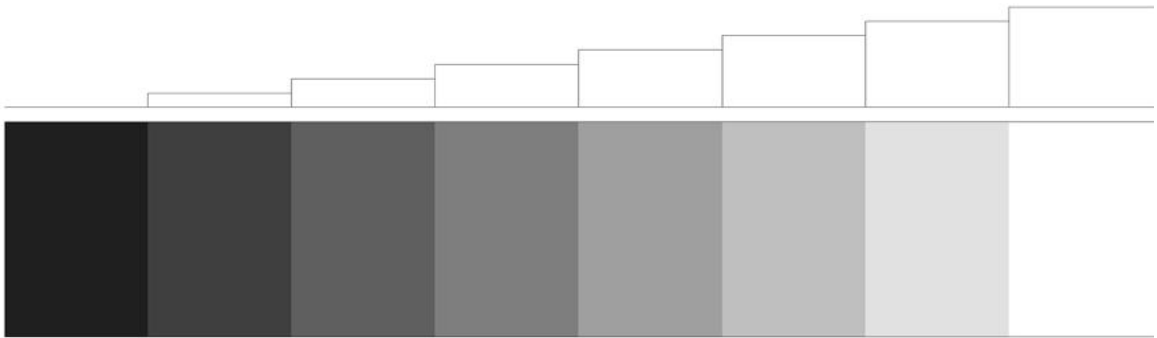


Рисунок 1.4 - Смути сірої шкали з рівними приростами інтенсивності

2 СТАНДАРТ JPEG КОДУВАННЯ СТАЛИХ ЗОБРАЖЕНЬ

2.1 Стандарт JPEG

Стандарт JPEG призначений для стиснення нерухомих монохромних та кольорових зображень з неперервними тонами. Це може бути досягнуто за допомогою одного з чотирьох основних режимів обробки: послідовного, прогресивного, безвтратного та ієрархічного. Стиснення та зберігання є основними цілями базової системи JPEG, яка забезпечує бітрейт 0,25 – 0,5 біта на піксель при помірній або гарній якості зображення [1, 11].

Рисунок 2.1 показує базову систему JPEG. Вхідні дані зображення, що підлягає компресії, попередньо трансформуються шляхом віднімання 2^{p-1} , де p - число біт, яке використовується для зберігання значень пікселів зображення. При $p = 8$ в базовому режимі відбувається зсув на 128. Зворотний зсув рівня на декодері відновлює відліки до 8-бітного подання без знаку [1].

Зсув рівня не впливає на коефіцієнти змінного струму (alternating current, AC) або дисперсії. Він впливає лише на коефіцієнт постійного струму (direct current, DC), зсуваючи до нуля середнє значення інтенсивності напівтонів. Різницеві значення між коефіцієнтами DC також залишаються незмінними. Начальне значення на початку зображення та в інтервалі перезапуску можна встановити рівним нулю [1].

Зображення після зміщення рівня розділяються на блоки 8x8 пікселів в режимі растрового (по строках) сканування для подальшого обчислення двомірного ДКП (дискретного косинусного перетворення). Коефіцієнти перетворення квантуються за допомогою заздалегідь визначених таблиць квантування для яскравості (luminance) та кольоровості (chrominance). Елементи матриці квантування мають обиратись виходячи з таких критеріїв: 1) вони мають бути додатними; 2) узгоджуватись із значеннями

ентропії та енергії відповідних ДКП коефіцієнтів; а також 3) забезпечувати рівень помилки квантування меншим за поріг візуальної чутливості [1]. Під час розробки алгоритму компресії JPEG [12] спеціалісти інтенсивно експериментували з простими і легкими методами квантування і запропонували таблиці для квантування коефіцієнтів DC та AC, які рекомендується використовувати за замовченням. Оскільки елементи матриці квантування зазвичай є різними, кожний з 64 коефіцієнтів ДКП квантується із своїм кроком [1].

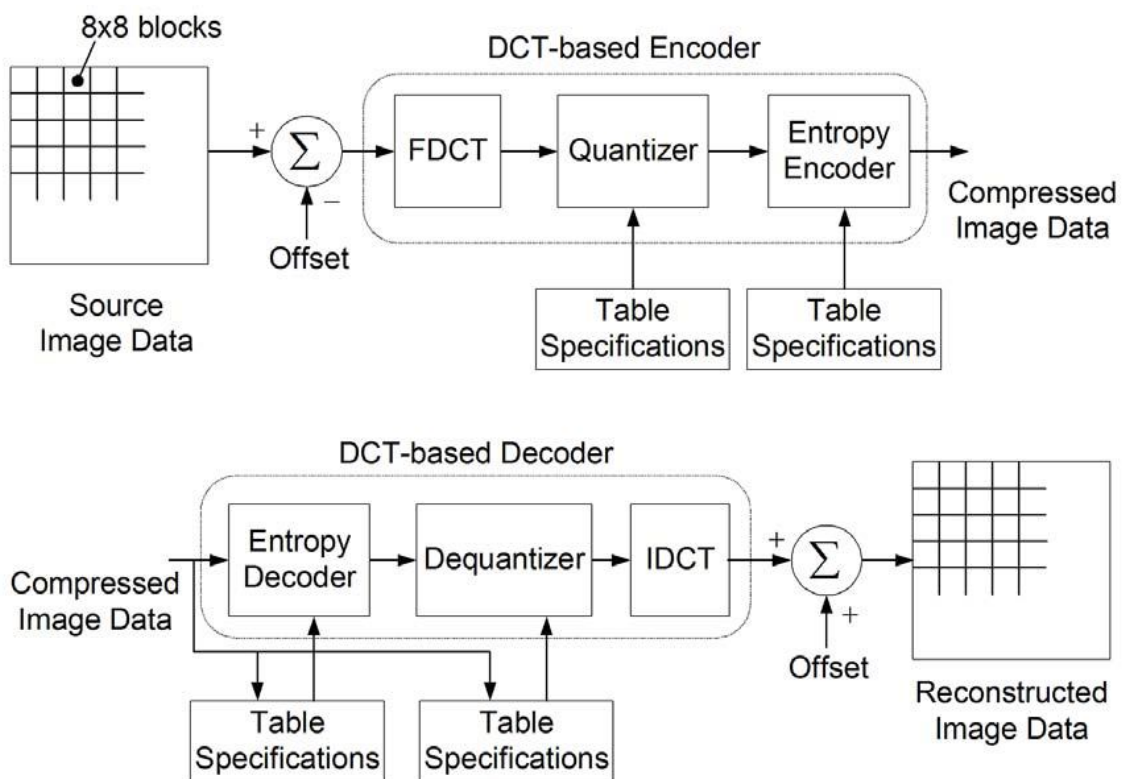


Рисунок 2.1 - Базовий потік кодування та декодування JPEG

Квантовані коефіцієнти ДКП, S_{quv} , визначаються наступним рівнянням [1]:

$$S_{quv} = \text{Round}(S_{uv} / Q_{uv}), \quad (2.1)$$

де S_{uv} - коефіцієнт ДКП, а Q_{uv} - відповідний елемент матриці квантування. Співвідношення (2.1) дозволяє легко та швидко здійснювати рівномірне квантування ДКП коефіцієнтів, які утворюють блоки розміром 8×8 коефіцієнтів. На рис. 2.2 показано матриці квантування, отримані емпірично. Розмір кроку Q_{uv} є функцією частоти коефіцієнта перетворення. Коефіцієнти, які відповідають нижнім частотам, квантуються з меншими кроками, а ті, що відносяться до вищих частот – з більшими кроками, це призводить до більших незворотних помилок квантування.

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

а)

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

б)

Рисунок 2.2 Приклад матриці квантування яскравості а) та кольоровості б), що використовуються в JPEG за замовченням [12]

Оскільки кодер і декодер повинні бути симетричними для ідеальної реконструкції на декодері, елементи матриці квантування повинні бути передані декодеру. Якщо використовуються матриці квантування за замовчуванням, то їх передавати не потрібно. Зворотна операція виконується відповідно до (2.2). Для DC коефіцієнтів попередньо відбувається відновлення фактичних значень за значеннями різниць коефіцієнтів, оскільки квантуванню піддавались саме різниці DC коефіцієнтів, які відносяться до двох сусідніх блоків [1]. Нагадаємо, що AC

коефіцієнти з блоків розміру 8x8 піддаються квантуванню незалежно один від одного.

Реконструкція коефіцієнтів виконується за допомогою тієї ж матриці квантування, а саме [1]:

$$R_{uv} = S_{quv} \cdot Q_{uv}. \quad (2.2)$$

Оскільки блокові артефакти, на появу яких в основному впливають значення DC коефіцієнта, є чутливими до відгуку просторово-частотної характеристики системи зору людини, коефіцієнт DC обробляється окремо від 63 коефіцієнтів AC. Він кодується диференціально за допомогою наступного передбачення першого порядку [1]:

$$DIFF = DC_i - DC_{i-1}, \quad (2.3)$$

де DC_i і DC_{i-1} - коефіцієнти DC поточного блоку 8x8 пікселів та попереднього блоку 8x8 пікселів відповідно.

На наступному етапі, квантовані 63 AC коефіцієнти, що знаходяться в блоках 8x8, розгортаються в одномірні масиви з використанням зиг-заг сканування як показано на рис. 2.3. Зазначена операція здійснюється в рамках підготовки до ентропійного кодування. Вздовж маршруту, який задає зиг-заг сканування, розташовуються ДКП коефіцієнти, що відповідають зростаючим просторовим частотам. Крім того, дисперсії відповідних ДКП коефіцієнтів як правило зменшуються по мірі просування зиг-заг маршрутом. Також квантування високочастотних коефіцієнтів із великими кроками відповідає властивостям СЗЛ і призводить до значної кількості нулів після квантування. Останнє дає змогу розробляти таблиці кодів із змінною довжиною блоку, які забезпечуватимуть ефективне кодування відомостей про серії (послідовності) нульових коефіцієнтів, які з'являються внаслідок зиг-заг

На рис. 2.4 показана спрощена система кодування JPEG2000. Як і в стандарті JPEG, спочатку застосовується зсув рівня шляхом віднімання 2^{m-1} , де 2^m кількість рівнів сірого в зображенні. Потім обчислюється дискретне вейвлет-перетворення. Коефіцієнти перетворення квантуються та ентропійно кодуються перед формуванням вихідного бітового потоку. Декодер є зворотним до кодера. Стиснений бітовий потік спочатку ентропійно декодується, деквантується та обчислюється зворотне вейвлет-перетворення, в результаті чого отримуються реконструйовані дані зображення [1].

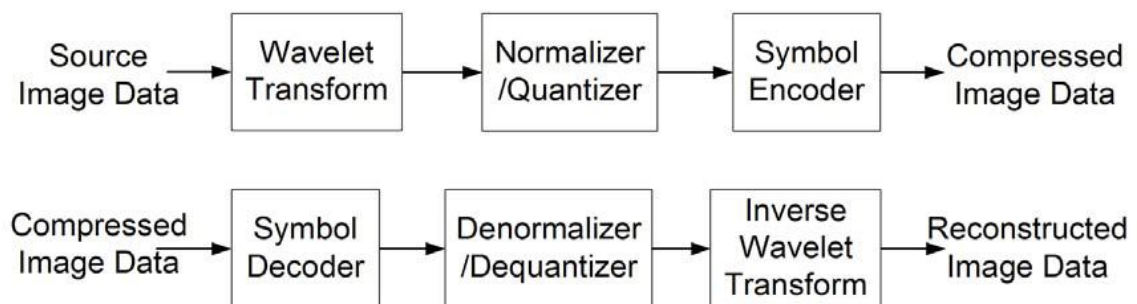


Рисунок 2.4 - Спрощена структурна схема кодера/декодера JPEG2000

Пряме DWT виконується серією одновимірних субсмугових декомпозицій на низькочастотну та високочастотну смуги. Відліки низьких частот представляють собою зменшену версію оригінального набору даних з низькою роздільною здатністю, тоді як відліки високих частот представляють собою зменшену версію з високою роздільною здатністю. Перетворення може використовуватися для зворотного (безвратного, безпомилкового) або незворотного (з втратами) стиснення. Зворотне перетворення за замовчуванням виконується за допомогою фільтра (5, 3), тоді як незворотне перетворення за замовчуванням виконується за допомогою фільтра Добеші (9, 7). Зауважте, що перше число в специфікації фільтра представляє кількість відводів фільтра, що використовуються для низькочастотної фільтрації на етапі аналізу або для

високочастотної фільтрації на етапі синтезу [1]. Детальніше про коефіцієнти фільтра можна знайти в [13, 14]. Після перетворення всі коефіцієнти квантуються.

Квантування - це процес, за допомогою якого точність подання коефіцієнтів зменшується, що призводить до стиснення даних зображення. Це операція з втратами, яка не може бути обернена, якщо розмір кроку квантування не дорівнює 1. Кожен з коефіцієнтів перетворення S_{uv} в субсмузі квантується до значення S_{quv} відповідно до формули [1]:

$$S_{quv} = S_{uv} \left\lfloor \frac{S_{uv}}{\Delta_b} \right\rfloor, \quad (2.4)$$

де Δ_b позначає крок квантування в субсмузі, який визначається [1]:

$$\Delta_b = 2^{R_b - e_b} \cdot \left(1 + \frac{m_b}{2^{11}} \right), \quad (2.5)$$

де R_b - номінальний динамічний діапазон субсмуги, e_b та m_b кількість бітів, виділених для *експоненти* та *мантиси* коефіцієнтів перетворення. Номінальний динамічний діапазон - це сума кількості бітів, що використовуються для представлення оригінального зображення, та бітів посилення аналізу для субсмуги b [1]. Пари *експонента/мантиса* (e_b, m_b) або явно сигналізуються в синтаксисі бітового потоку для кожної субсмуги, що відомо як явне квантування, або сигналізуються лише в бітовому потоці для смуги LL, що відомо як неявне квантування. В останньому випадку решта субсмуг квантуються за допомогою екстрапольованих параметрів субсмуги LL. Нехай e_0 і m_0 кількість бітів, виділених для субсмуги LL, екстрапольовані параметри для субсмуги b будуть [1]:

$$m_b = m_0, \quad (2.6)$$

$$e_b = e_0 \cdot nsd_b - nsd_0, \quad (2.7)$$

де nsd_b позначає кількість рівнів декомпозиції субсмути від оригінального зображення до субсмути b .

Основна відмінність між JPEG на основі ДКП та JPEG2000 на основі вейвлетів полягає у відсутності процесу розбиття на блоки 8x8 пікселів в останньому (тобто в JPEG2000). Поділ на блоки непотрібний, оскільки вейвлет-перетворення є за своєю суттю локальним (тобто базисні функції перетворення змінюються за тривалістю), так і обчислювально ефективним. На рис. 2.5 показано два стовпці реконструйованих зображень, стиснених за допомогою JPEG (лівий стовпець) та JPEG2000 (правий стовпець). Зверху вниз показані реконструйовані зображення, збільшені фрагменти зображення та зображення помилок. Відмова від процесу поділу на блоки 8x8 пікселів в JPEG2000 усуває артефакти блокування, що є характерним для кодеку на основі ДКП при високих коефіцієнтах стиснення [1].

Крім зменшення помилки відновлення, що вимірюється в термінах середньоквадратичного відхилення (rms), як показано в табл. 2.1 [1], JPEG2000 підвищує якість зображення з боку суб'єктивного сприйняття.

Таблиця 2.1 – Результати порівняльного аналізу кодеків JPEG і JPEG2000 для зображення *Girl*

Кодек	Коефіцієнт стиснення	Похибка (rms)	Параметри стиснення
JPEG	21,85	7,63	Quality factor=4
JPEG2000	22,53	6,26	$m = 8, e = 8,6$

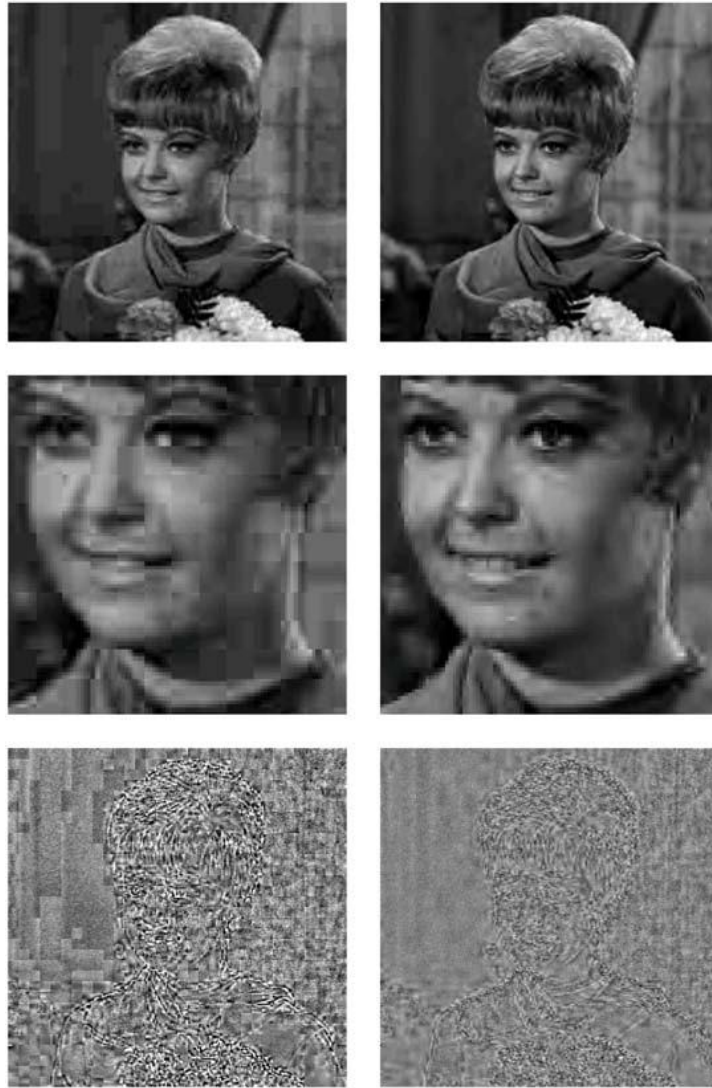


Рисунок 2.5 - Порівняння між кодеками JPEG (ліва колонка) та JPEG2000 (права колонка). Зверху вниз показано відновлене зображення, збільшені фрагменти зображення для розрізнення деталей та зображення з помилками

3 СТРУКТУРНІ ПІДХОДИ ДО ОЦІНКИ ЯКОСТІ ЗОБРАЖЕННЯ

Одна відмінна риса, яка відрізняє сигнали природних зображень від «типового» зображення, випадково вибраного з простору зображень, полягає в тому, що вони високо структуровані — зразки сигналів демонструють сильні залежності між собою. Ці залежності несуть важливу інформацію про структури об'єктів у візуальній сцені. Метрика якості зображення, яка ігнорує такі залежності, може не забезпечити ефективних прогнозів якості зображення. Для прикладу скористаємось метрикою помилки Мінковського. У просторовій області метрика Мінковського між еталонним зображенням \vec{x} (яке передбачає ідеальну якість) та спотвореним зображенням \vec{y} визначається як [20]

$$E_p = \left(\sum_{i=1}^N |x_i - y_i|^p \right)^{1/p}, \quad (3.1)$$

де x_i та y_i — це i -ті відліки зображень \vec{x} , а \vec{y} , відповідно, N — кількість відліків зображення, а p відноситься до показника ступеня та змінюється в діапазоні $p \in [1, \infty)$. На рис. 3.1 показано два спотворені зображення, отримані з одного й того ж оригінального зображення [20]. Перше спотворене зображення було отримано шляхом додавання константи до всіх відліків сигналу, а друге було створено тим самим методом, за винятком того, що знаки константи обирались випадковим чином як додатні або від'ємні. Легко показати, що відстані Мінковського між оригінальним зображенням та обома спотвореними зображеннями абсолютно однакові, незалежно від значення показника ступеня p , що використовувався. Однак візуальна якість двох спотворених зображень різночче відрізняється [20]. Інший приклад наведено на рис. 3.2, де

зображення В було згенеровано шляхом додавання незалежного білого гаусового шуму до оригінального текстурного зображення А. На зображенні С значення вибірок сигналу залишилися такими ж, як і на зображенні А, але просторове впорядкування відліків (пікселів) було змінено (за допомогою процедури сортування). Зображення D було отримано з зображення В, дотримуючись тієї ж процедури перепорядкування, що й для створення зображення С. Знову ж таки, відстані Мінковського між зображеннями А та В, а також зображеннями С та D абсолютно однакові, незалежно від того який степінь p використовується. Однак зображення D видається зашумленим в значно більшій мірі в порівнянні з зображенням В [20].

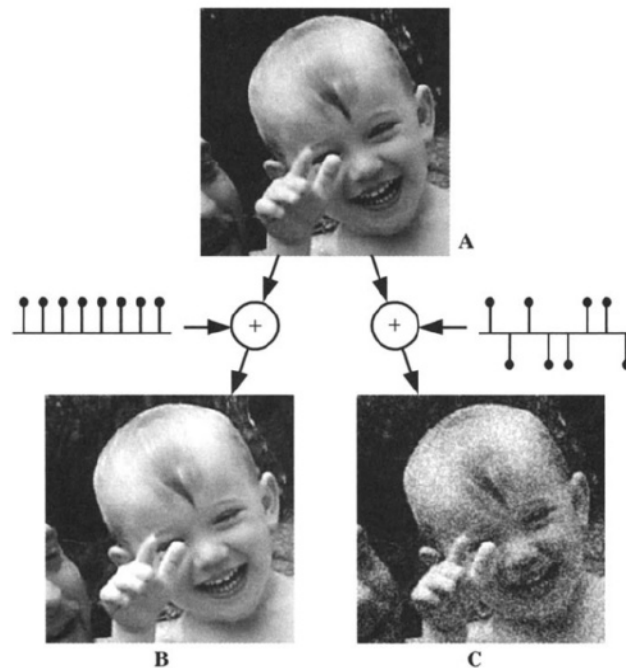


Рисунок 3.1 – Ілюстрація недоліків відстані Мінковського для прогнозування якості зображення. А: початкове зображення; В: зображення, спотворене додаванням додатної константи; С: зображення спотворене додаванням тієї ж константи, але з випадковим знаком. Зображення В та С мають однакову відстань Мінковського відносно зображення А, але різоче різну візуальну якість

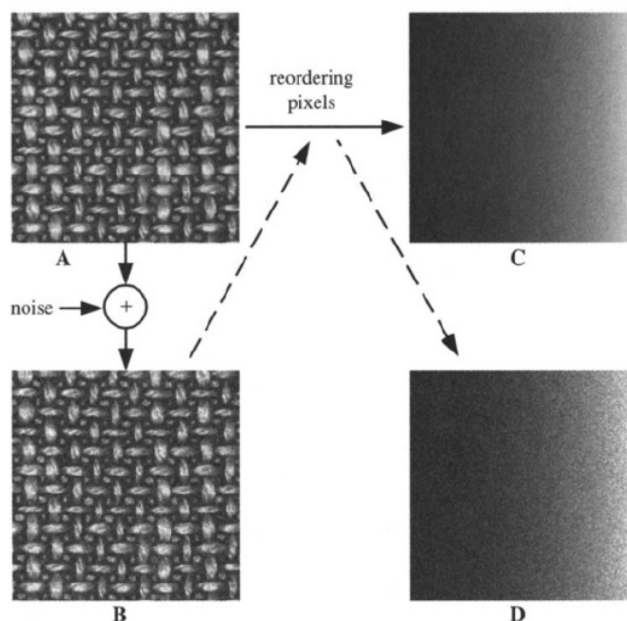


Рисунок 3.2 – Помилка відстані Мінковського для прогнозування якості зображення. А: Оригінальне текстурне зображення; В: спотворене зображення шляхом додавання незалежного білого гаусового шуму; С: перевпорядкування пікселів на зображенні А (шляхом сортування значень інтенсивності пікселів); D: перевпорядкування пікселів на зображенні В шляхом дотримання того ж перевпорядкування, яке використовувалося для створення зображення С. Відстані Мінковського між зображеннями А та В, а також зображеннями С та D однакові, але зображення D виглядає набагато шумнішим, ніж зображення В

Існують різні способи пояснити очевидну неефективність відстані Мінковського в цих прикладах. Один із способів — звернутись до психофізичних особливостей людського зору. Тут ми наводимо більш пряме пояснення, засноване на математичних властивостях відстані Мінковського. Зауважте, що неявним припущенням відстані Мінковського є те, що всі відліки (пікселі) сигналу є незалежними [20]. Як результат, порядок розташування відліків сигналу не впливає на значення рівня вимірних спотворень. Це різко контрастує з тим фактом, що природні сигнали, які відповідають зображенням є високоструктурованими; тобто,

впорядкованість і шаблон притаманні відлікам сигналу несуть більшу частину візуальної інформації на зображенні. Отже, «правильний» показник якості зображення повинен бути здатним відображати структурну інформацію або відчувати структурні зміни в сигналах, що відповідають зображенням [20].

На рисунку 3.3 показано одне з потенційних рішень для подолання цієї проблеми. Ідея полягає в тому, щоб застосувати перетворення зображення перед визначенням відстані Мінковського, щоб відліки сигналу в області перетворення стали незалежними (або принаймні декорельованими) [20]. Додатковою вимогою до перетворення T є те, що воно не має призводити до втрат або «візуально» виглядати без втрат, в тому сенсі, що вся важлива візуальна інформація зберігається після перетворення (резонно, має існувати зворотне перетворення, яке може реконструювати сигнали зображення в просторовій області). Оскільки таке перетворення може усунути залежності між відліками сигналу, що відповідає зображенню, без втрати важливої візуальної інформації, можна сказати, що «структура» сигналу, що відповідає зображенню, добре відображається поданням сигналу в області перетворення [20].

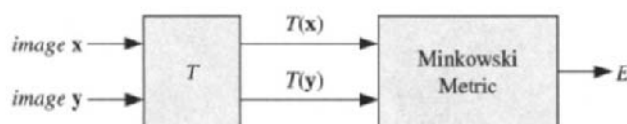


Рисунок 3.3 – Застосування перетворення зображення перед визначенням відстані Мінковського може потенційно зменшити залежності між відліками сигналу, тим самим покращуючи метрику якості зображення

Структура, представлена на рис. 3.3, ґрунтується на мотивації відтворення обчислювальних аспектів ранніх стадій розвитку СЗЛ. Така аналогія є доцільною з точки зору обчислювальної нейронауки [20]. У цьому контексті десятиліття тому було висунуто припущення, що роль

ранніх біологічних сенсорних систем полягає в усуненні надлишків у сенсорних сигналах, що призводить до набору нейронних відповідей, які є статистично незалежними, відомих як принцип «ефективного кодування» [17, 18, 20].

В наступному розділі зосереджено увагу на іншому підході до оцінки якості зображення: методах *на основі структурної подібності* [19]. Замість того, щоб намагатися розробити ідеальне перетворення, яке може повністю усунути залежності в сигналі, як це запропоновано на рис. 3.3, ці методи замінюють метрику похибки Мінковського різними вимірюваннями, які адаптовані до структури опорного вхідного сигналу. Нижче сформульовано алгоритми індексу структурної подібності та описано інтуїтивні принципи їх розробки, а також демонструється як ці алгоритми використовуються для оцінки якості зображення [20].

4 ІНДЕКС СТРУКТУРНОЇ ПОДІБНОСТІ

Найфундаментальніший принцип, що лежить в основі структурних підходів до оцінки якості зображення, полягає в тому, що СЗЛ високо адаптована для вилучення структурної інформації з візуальної сцени, і тому вимірювання структурної подібності (або спотворення) повинно забезпечити гарне наближення до сприймаємої якості зображення. Залежно від того, як визначено *структурну інформацію* та *структурне спотворення*, можуть існувати різні способи розробки алгоритмів оцінки якості зображення. Індекс структурної подібності (SSIM) є конкретною реалізацією з точки зору формування зображення [20].

Щоб зрозуміти інтуїцію методу індексу SSIM, давайте ще раз розглянемо простір зображень, описаний в попередньому розділі. На рис. 4.1, еталонне зображення (оригінальне зображення "Ейнштейн") представлено як вектор у просторі зображень [20]. Будь-яке спотворення зображення можна інтерпретувати як додавання вектора спотворення до центрального вектора еталонного зображення. Зокрема, вектори спотворення з однаковою нормою визначають гіперсферу середньоквадратичної помилки (MSE) у просторі зображень. Однак, як показано на рис. 4.1, зображення, що знаходяться на одній гіперсфері, можуть мати різне візуальне якість [20]. Це означає, що норми вектора спотворення недостатньо як корисного показника якості зображення, і що напрямки цих векторів мають важливіші перцептивні (тобто такі, що сприймаються) значення. Фактично, ми хочемо відокремити вплив освітлення від решти інформації, яка представляє структури об'єкта. Інтуїтивно зрозуміло, що основним впливом зміни освітлення на зображення є зміна середньої локальної яскравості та контрастності, і така зміна не повинна мати сильного впливу на якість зображення, що сприймається [20]. Це підтверджується рис. 4.1, де

зображення лише зі змінами яскравості або контрастності мають набагато кращу якість, ніж інші зображення з серйозними «структурними» спотвореннями [20].

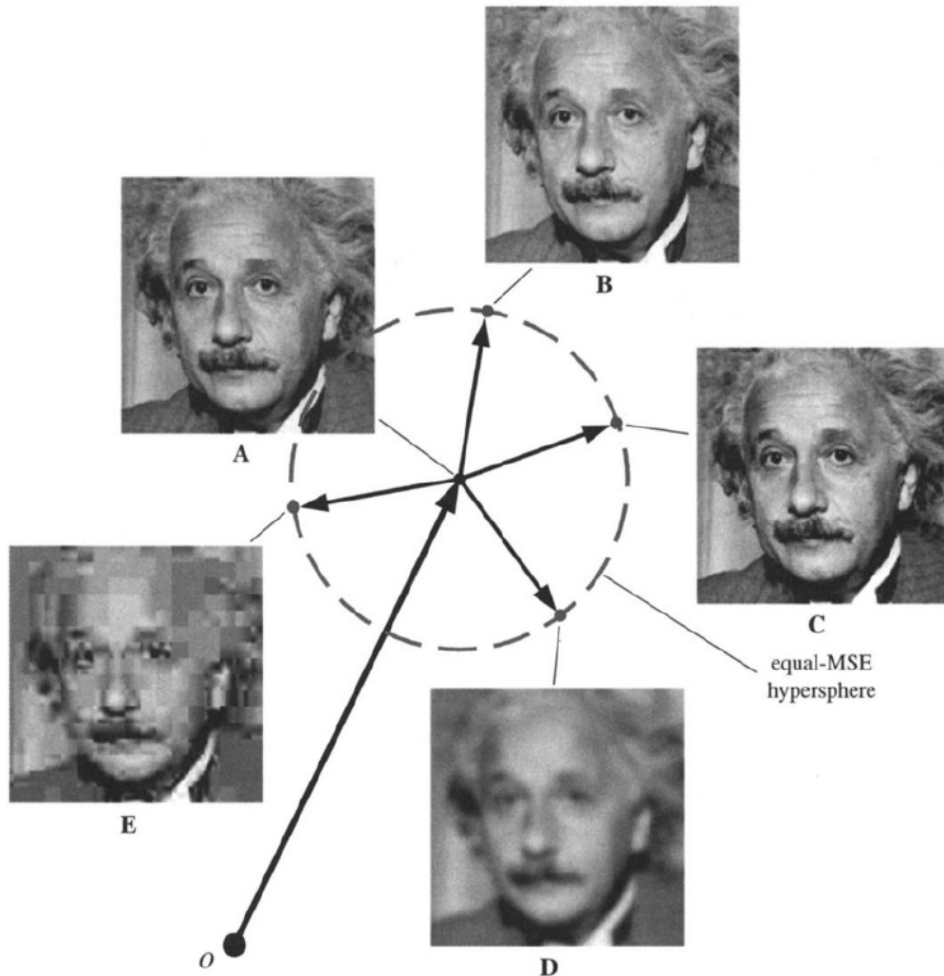


Рисунок 4.1 – Зображення можна представити як вектор у просторі зображень, розмірність якого дорівнює кількості пікселів у зображенні. Зображення з однаковою середньоквадратичною похибкою (MSE) відносно вихідного зображення утворюють гіперсферу в просторі зображень, але зображення, що знаходяться на одній гіперсфері, мають різне візуальне якість. А: початкове (еталонне) зображення; В: зображення зміщене за середнім значенням $MSE = 144$; С: контрастно розтягнуте зображення, $MSE = 144$; D: розмите зображення, $MSE = 144$; Е: стиснене зображення JPEG, $MSE = 142$

На рисунку 4.2 показано, як зміни яскравості та контрастності можна відокремити від структурних спотворень у просторі зображення. Зміни яскравості можна охарактеризувати рухом вздовж напрямку \vec{x} , визначеного $x_1 = x_2 = \dots = x_N$, який є перпендикулярним до гіперплощини $\sum_{i=1}^N x_i = 0$ [20]. Зміни контрастності визначаються напрямом $\vec{x} - \bar{x}$. У просторі зображення два вектори, що визначають зміни яскравості та контрастності, утворюють двовимірний підпростір (площину), який є узгодженим з вектором еталонного зображення \bar{x} . Наприклад, на рис. 4.1 площина, визначена опорним зображенням А, містить не тільки саме еталонне зображення, але й зображення В та С. Спотворення зображення, що залишилося, відповідає повороту такої площини на певний кут, що ми інтерпретуємо як структурну зміну на рис. 4.2 [20].

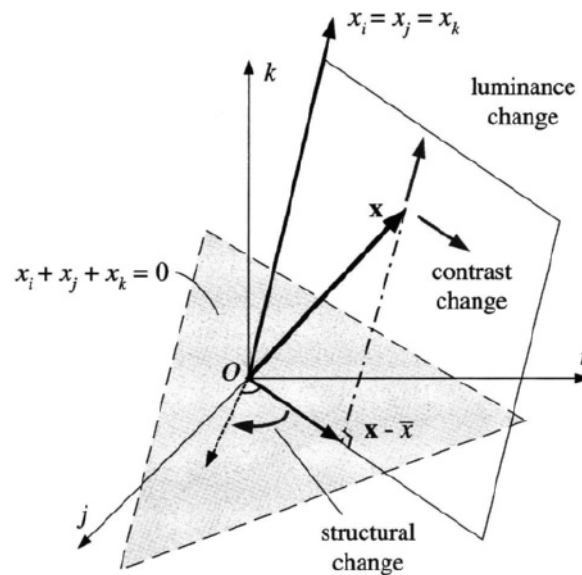


Рисунок 4.2 – Відокремлення яскравості, контрасту та структурних змін від еталонного зображення \bar{x} в просторі зображень. Це ілюстрація в тривимірному просторі. На практиці кількість вимірів дорівнює кількості пікселів зображення

Структурна схема алгоритму розрахунку індексу SSIM показана на рис. 4.3. Спочатку яскравість кожного сигналу оцінюється як середня інтенсивність [20]:

$$\mu_x = \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (4.1)$$

Функція порівняння яскравості $l(\bar{x}, \bar{y})$ тоді є функцією від μ_x та μ_y [20]:

$$l(\bar{x}, \bar{y}) = l(\mu_x, \mu_y). \quad (4.2)$$

По-друге, ми видаляємо середню інтенсивність із сигналу. Отриманий сигнал $\bar{x} - \mu_x$ відповідає проекції вектора \bar{x} на гіперплощину $\sum_{i=1}^N x_i = 0$, як показано на рис. 4.2. Ми використовуємо стандартне відхилення як оцінку контрасту сигналу. Незміщена оцінка в дискретній формі задається як [20]

$$\sigma_x = \left(\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2}. \quad (4.3)$$

Таким чином порівняння контрасту $c(\bar{x}, \bar{y})$ є порівнянням σ_x та σ_y [20]:

$$c(\bar{x}, \bar{y}) = c(\sigma_x, \sigma_y). \quad (4.4)$$

По-третє, сигнал нормалізується (ділиться) на власне стандартне відхилення, так що два сигнали, що порівнюються, мають одиничне стандартне відхилення. Порівняння структур $s(\bar{x}, \bar{y})$ проводиться для цих нормалізованих сигналів [20]:

$$s(\vec{x}, \vec{y}) = s\left(\frac{\vec{x} - \mu_x}{\sigma_x}, \frac{\vec{y} - \mu_y}{\sigma_y}\right). \quad (4.5)$$

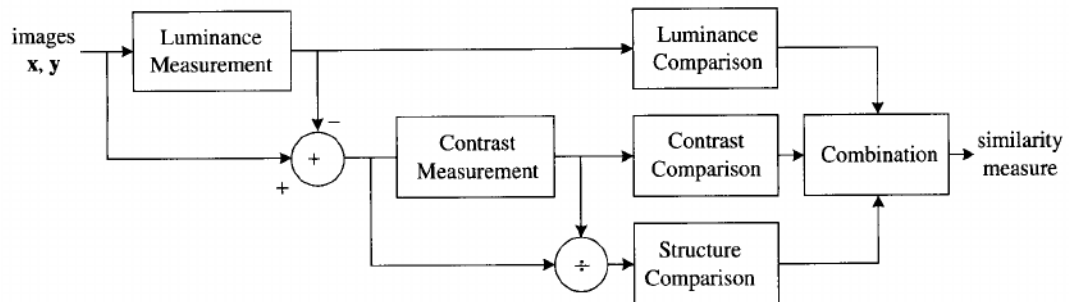


Рисунок 4.3 – Схема системи вимірювання подібності зображень

Зрештою, три компоненти об'єднуються для отримання загальної міри подібності [20]:

$$S(\vec{x}, \vec{y}) = f(l(\vec{x}, \vec{y}), c(\vec{x}, \vec{y}), s(\vec{x}, \vec{y})). \quad (4.6)$$

Важливим моментом є те, що ці три компоненти є відносно незалежними, що фізично доцільно, оскільки зміна яскравості та/або контрастності мало впливає на структури об'єктів у сцені.

Щоб завершити визначення міри подібності в рівнянні (4.6), нам потрібно визначити три функції $l(\vec{x}, \vec{y})$, $c(\vec{x}, \vec{y})$ та $s(\vec{x}, \vec{y})$, а також комбінаційну функцію $f(\cdot)$. Крім того, ми також хотіли б, щоб міра подібності задовольняла такі умови [20]:

1. Симетрія: $S(\vec{x}, \vec{y}) = S(\vec{y}, \vec{x})$. Під час кількісної оцінки подібності між двома сигналами зміна порядку вхідних сигналів не повинна впливати на результуючий вимір.

2. Обмеженість: $S(\vec{x}, \vec{y}) \leq 1$. Верхня межа може служити показником того, наскільки близькі два сигнали до ідеального співпадіння. Зверніть увагу, що метрики типу відношення сигнал/шум зазвичай необмежені.

3. Унікальний максимум: $S(\bar{x}, \bar{y}) = 1$ тоді і тільки тоді, коли $\bar{x} = \bar{y}$.

Максимальне значення досягається лише тоді, коли сигнали, що порівнюються, ідентичні. Іншими словами, міра подібності повинна кількісно визначати будь-які відмінності, які можуть існувати між вхідними сигналами.

Порівняння яскравості визначається як [20]

$$I(\bar{x}, \bar{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (4.7)$$

де константа C_1 включена, щоб уникнути нестабільності, коли сума $\mu_x^2 + \mu_y^2$ дуже близька до нуля. Гарним вибором є [20]

$$C_1 = (K_1 \cdot L)^2, \quad (4.8)$$

де L — динамічний діапазон значень пікселів (255 для 8-бітних зображень за шкалою відтінків сірого), а $K_1 \ll 1$ — невелика константа. Подібні міркування також застосовуються до порівняння контрасту та порівняння структури, як описано нижче. Легко побачити, що рівняння (4.7) відповідає трьом властивостям, переліченим вище [20].

Рівняння (4.7) також пов'язане із законом Вебера, який широко використовується для моделювання адаптації до освітлення (також відомої як маскування яскравості) в СЗЛ. Згідно із законом Вебера, величина ледь помітної зміни яскравості ΔI приблизно пропорційна фоновій яскравості I для широкого діапазону значень яскравості. Іншими словами, СЗЛ чутлива до відносної, а не абсолютної зміни яскравості. Якщо через R позначити відношення яскравості спотвореного сигналу відносно опорного сигналу, тоді можна записати $\mu_y = R \cdot \mu_x$. Підставляючи це в рівняння (4.7), отримуємо [20]

$$l(\vec{x}, \vec{y}) = \frac{2R + C_1 / \mu_x^2}{1 + R^2 + C_1 / \mu_x^2}. \quad (4.9)$$

Якщо ми припустимо, що значення C_1 є достатньо малим (відносно μ_x^2), щоб його можна було ігнорувати, тоді $l(\vec{x}, \vec{y})$ є функцією лише від R , а не від $\Delta I = \mu_y - \mu_x$. У цьому сенсі це якісно узгоджується із законом Вебера.

Функція порівняння контрасту має подібний вигляд [20]:

$$c(\vec{x}, \vec{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (4.10)$$

де C_2 — невід'ємна константа

$$C_2 = (K_2 \cdot L)^2 \quad (4.11)$$

і K_2 задовольняє умові $K_2 \ll 1$. Це визначення знову задовольняє три перелічені вище властивості. Важливою особливістю цієї функції є те, що при однаковій величині $\Delta\sigma = \sigma_y - \sigma_x$ цього показника вона менш чутлива до випадку високого основного контрасту, σ_x ніж до випадку низького основного контрасту. Це пов'язано з функцією маскуванню контрасту СЗЛ [20].

Порівняння структур проводиться після віднімання яскравості та нормалізації контрасту. Геометрично ми можемо пов'язати структури двох зображень з напрямком двох одиничних векторів $(\vec{x} - \mu_x) / \sigma_x$ та $(\vec{y} - \mu_y) / \sigma_y$, кожен з яких лежить у гіперплощині, $\sum_{i=1}^N x_i = 0$ як показано на рис. 4.2. Кут між двома векторами забезпечує просту та ефективну міру для кількісної оцінки структурної подібності [20].

Зокрема, коефіцієнт кореляції між \vec{x} та \vec{y} відповідає косинусу кута. Таким чином, ми визначаємо функцію порівняння структур як [20]:

$$s(\vec{x}, \vec{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}. \quad (4.12)$$

Як і у випадку з мірами яскравості та контрастності, ми вводимо невелику константу як у знаменнику, так і в чисельнику. Значення σ_{xy} можна оцінити як [20]:

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y). \quad (4.13)$$

Зверніть увагу, що $s(\vec{x}, \vec{y})$ може набувати від'ємних значень. Від'ємні значення структурної подібності відповідають випадкам, коли локальні структури зображення інвертовані.

Нарешті, ми об'єднуємо три порівняння (4.7), (4.10) та (4.12). Результатом є клас мір подібності зображень, які ми разом називаємо індексами структурної подібності (SSIM), між сигналами \vec{x} та \vec{y} [20]:

$$SSIM(\vec{x}, \vec{y}) = [l(\vec{x}, \vec{y})]^\alpha \cdot [c(\vec{x}, \vec{y})]^\beta \cdot [s(\vec{x}, \vec{y})]^\gamma, \quad (4.14)$$

де $\alpha > 0$, $\beta > 0$ та $\gamma > 0$ - параметри, що використовуються для коригування відносної важливості трьох компонентів. Легко перевірити, що це визначення задовольняє три умови, наведені вище. Далі ми будемо вважати $\alpha = \beta = \gamma = 1$ та $C_3 = C_2 / 2$ [20].

Це призводить до отримання зазначеного індексу SSIM [19]:

$$SSIM(\vec{x}, \vec{y}) = l(\vec{x}, \vec{y}) \cdot c(\vec{x}, \vec{y}) \cdot s(\vec{x}, \vec{y}) =$$

$$\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \cdot \frac{\sigma_{xy} + C_2/2}{\sigma_x\sigma_y + C_2/2} =$$

$$\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \cdot \frac{2\sigma_{xy} + C_2}{2\sigma_x\sigma_y + C_2} =$$

$$\frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \cdot \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} =$$

$$\frac{(2\mu_x\mu_y + C_1) \cdot (2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1) \cdot (\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4.15).$$

5 ОЦІНЮВАННЯ ЯКОСТІ ЗОБРАЖЕНЬ З ВИКОРИСТАННЯМ ІНДЕКСУ СТРУКТУРНОЇ ПОДІБНОСТІ SSIM

Індекс SSIM (Structural Similarity Index) вимірює структурну подібність між двома сигналами зображення. Якщо один із сигналів зображення вважається ідеальним за якістю, то індекс SSIM можна розглядати як показник якості іншого порівнюваного сигналу зображення [20]. При застосуванні підходу з індексом SSIM до зображень великого розміру доцільно обчислювати його локально, а не глобально. Причин цьому багато. По-перше, статистичні характеристики зображень зазвичай просторово нестационарні. По-друге, спотворення зображення, які можуть залежати або не залежати від локальних статистик зображення, також можуть змінюватися в просторі. По-третє, через нерівномірну функцію вибірки сітківки системи зору людини (СЗЛ) при типових відстанях перегляду людський спостерігач може сприймати з високою роздільною здатністю лише локальну область зображення за один момент часу [20]. Нарешті, локалізоване вимірювання якості може надати просторово змінну карту якості зображення, яка дає більше інформації про деградацію якості зображення. Така карта якості може використовуватися по-різному. Вона може бути використана як показник варіації якості на зображенні. Вона також може бути використана для контролю якості зображення в системах обробки зображень з просторовою змінною, наприклад, при кодуванні зображень з областю інтересу та фовеальною обробкою зображень [22].

В одній з реалізацій підходу з індексом SSIM [23, 24] локальні статистики μ_x , σ_x та σ_{xy} (рівняння (4.1), (4.3) та (4.13)) обчислюються в межах локального квадратного вікна 8×8 . Вікно переміщується піксель за пікселем з верхнього лівого кута в нижній правий кут зображення. На кожному кроці локальні статистики та індекс SSIM обчислюються в межах локального вікна [20]. Однією з проблем цього методу є те, що

результуюча карта індексу SSIM часто демонструє небажані артефакти "блокування", як показано на рис. 5.1 С. Такі "артефакти" небажані, тому що вони створюються вибором методу вимірювання якості (локальне квадратне вікно), а не спотвореннями зображення. В [19] запропоновано використовувати гаусову вагову функцію з круговою симетрією $\vec{w} = \{w_i | i = 1, 2, \dots, N\}$, сума ваг має дорівнювати одиниці, $\sum_{i=1}^N w_i = 1$.

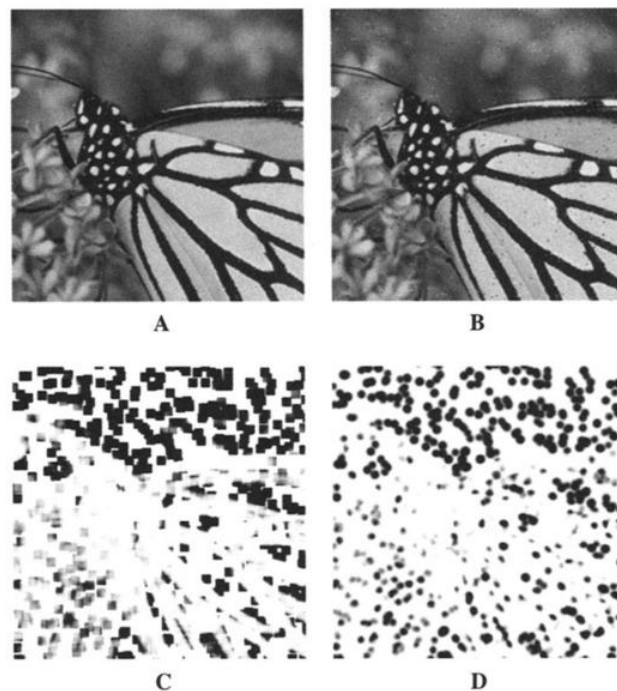


Рисунок 5.1 - Вплив форми локального вікна на карту індексу структурної подібності (SSIM): А - оригінальне зображення; В - зображення, спотворене імпульсним шумом; С - карта індексу SSIM з використанням підходу з залученням квадратного вікна; D - карта індексу SSIM з використанням підходу з використанням згладжуючого вікна. На обох картах індексу SSIM світліший колір вказує на кращу якість

Таким чином, оцінки локальних статистик, μ_x , σ_x та σ_{xy} , набувають форму:

$$\mu_x = \sum_{i=1}^N w_i \cdot x_i, \quad (5.1)$$

$$\sigma_x = \sqrt{\sum_{i=1}^N w_i \cdot (x_i - \mu_x)^2}, \quad (5.2)$$

$$\sigma_{xy} = \sum_{i=1}^N w_i \cdot (x_i - \mu_x) \cdot (y_i - \mu_y). \quad (5.3)$$

При такому підході зі згладженим вікном карти якості демонструють локально ізотропну властивість, як показано на рис. 5.1 D [20].

На рис. 5.2 показано карти індексу SSIM для набору зразків зображень з різними типами спотворень [20]. Для порівняння також включена карта абсолютної помилки для кожного спотвореного зображення. Індекс SSIM та карти абсолютної помилки були скориговані таким чином, що більш світлий колір завжди вказує на кращу якість з точки зору даної міри якості/спотворення [20]. Можна бачити, що спотворені зображення демонструють змінну якість в просторі. Наприклад, на зображенні В шум в області обличчя здається значно більш вираженим, ніж у текстурних областях. Однак карта абсолютної помилки (D) повністю незалежна від базових структур зображення. Навпаки, карта індексу SSIM (C) дає передбачення, яке узгоджується із сприйняттям. На зображенні F схема розподілу бітів низькобітного стиснення JPEG2000 призводить до згладженого представлення деталізованих структур зображення [20]. Наприклад, інформація про текстуру даху будівлі, а також дерев втрачена. Це дуже добре показано картою індексу SSIM (G), але не може бути передбачено за картою абсолютної помилки (H). Деякі інші типи спотворень спричинені низькобітним стисненням JPEG. На зображенні J основні спотворення, які ми спостерігаємо, це ефект блокування на небі та артефакти навколо зовнішніх меж будівлі. Знову ж таки, карта абсолютної

помилки (L) не дає корисного передбачення, а карта індексу SSIM (K) успішно передбачає зміни якості зображення в просторі [20]. З цих зразків зображень ми бачимо, що така проста міра якості зображення, як індекс SSIM, може адаптуватися до різних видів спотворень зображення та забезпечувати узгоджені із сприйняттям передбачення якості [20].

Останнім кроком в системі вимірювання якості зображення є об'єднання карти якості в один єдиний показник якості для всього зображення. Зручним способом є використання зваженого підсумовування [20]. Нехай \vec{X} та \vec{Y} - два зображення, що підлягають порівнянню, а $SSIM(\vec{x}_j, \vec{y}_j)$ - локальний індекс SSIM, обчислений для j -го локального зразка (тобто, $SSIM(\vec{x}_j, \vec{y}_j)$ для всіх значень j утворює карту якості за індексом SSIM, як це було зображено на рис. 5.2), тоді SSIM індекс між \vec{X} та \vec{Y} визначається як [20]:

$$SSIM(\vec{X}, \vec{Y}) = \frac{\sum_{j=1}^{N_s} W_j(\vec{x}_j, \vec{y}_j) \cdot SSIM(\vec{x}_j, \vec{y}_j)}{\sum_{j=1}^{N_s} W_j(\vec{x}_j, \vec{y}_j)}, \quad (5.4)$$

де N_s кількість відліків в карті якості, а $W_j(\vec{x}_j, \vec{y}_j)$ вага, присвоєна j -му відліку. Якщо всі відліки на карті якості мають однакову вагу, то $W_j(\vec{x}_j, \vec{y}_j) = 1$. Це дає середнє значення SSIM (MSSIM), яке використовується в [19].

Існує два випадки, коли бажаним є нерівномірне зважування. По-перше, залежно від застосування, є деякі попередні знання про важливість різних областей зображення, і така карта важливості може бути перетворена на вагову функцію. Наприклад, системи обробки зображень на основі об'єктних областей інтересу часто сегментують об'єкти в сцені і надають різним об'єктам різну важливість. У фовеальній системі обробки

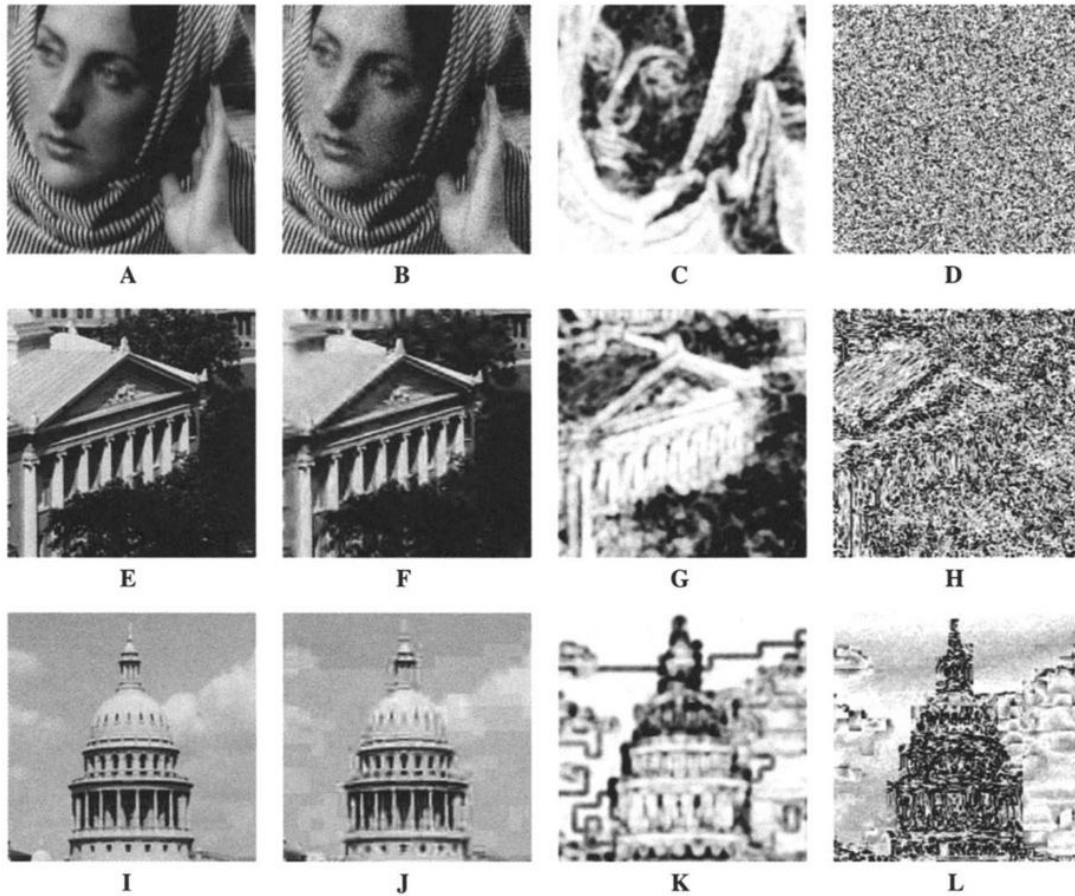


Рисунок 5.2 - Зразки спотворених зображень та їхні карти якості/спотворень (зображення обрізані до розміру 160x160 пікселів для наочності); (А, Е, І) - оригінальні зображення; (В) - зображення, що було спотворено гаусовим шумом; (F) - зображення, стиснене кодером JPEG2000; (J) - зображення, стиснене кодером JPEG; (C, G, K) - карти індексу структурної подібності (SSIM) спотворених зображень, де яскравість вказує на величину локального індексу SSIM (обчисленого за допомогою ковзного квадратного вікна для наочності); (D, H, L) - карти абсолютної помилки спотворених зображень, де затемнення вказує на абсолютне значення локальної різниці пікселів. Зверніть увагу, що на всіх картах якості/спотворень (C, D, G, H, K та L) яскравіший колір вказує на кращу якість з точки зору базової міри якості/спотворення

зображень [22] вагова функція може бути визначена відповідно до характеристики фовеації СЗЛ (тобто зорова роздільна здатність поступово зменшується як функція відстані від точки фіксації). Зверніть увагу, що вагова функція тут визначається лише просторовим положенням j , але не залежить від локального вмісту зображення \vec{x}_j і \vec{y}_j .

У другому випадку вміст зображення також відіграє роль. Було помічено, що різні текстури зображення по-різному привертають фіксацію погляду людини, і тому цим областям може бути призначені різні ваги. В [24] використовується вагова функція, зважена за дисперсією, тобто [20]:

$$W(\vec{x}, \vec{y}) = \sigma_x^2 + \sigma_y^2 + C_2. \quad (5.5)$$

Було помічено, що ця вагова функція корисна для врівноваження екстремального випадку, коли сильні спотворення з високою дисперсією концентруються на деяких невеликих областях зображення.

6 РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНИХ ДОСЛІДЖЕНЬ

Зазвичай природні зображення дуже структуровані, їхні пікселі демонструють сильну кореляцію, особливо коли вони просторово близькі, і ці залежності несуть важливу інформацію про структуру об'єктів у візуальній сцені. Одним з коефіцієнтів, який задовільно представляє особливості СЗЛ, є індекс структурної подібності, SSIM, вперше введений Wang'ом і Bovik'ом в 2002 році [22]. Цей коефіцієнт, природно, є привабливим коефіцієнтом, оскільки він передбачає просту факторизацію трьох множників (компонент), в якій перші дві компоненти є функціями вибіркового середнього та вибіркового дисперсії, а третя компонента відповідає вибіркового коефіцієнту кореляції між зображеннями, тобто структурна частина індексу порівнює середні та дисперсії і обчислює лінійну асоціацію між двома зображеннями.

Детальна інформація щодо обґрунтування саме такої формули для обчислення SSIM міститься в розділі 4-му пояснювальної записки. В свою чергу, деталі, що стосуються особливостей обчислення і практичного застосування SSIM наведено в 5-му розділі цієї роботи. В даному розділі буде розглянуто питання доцільності використання індексу SSIM для оцінювання якості стиснених JPEG зображень.

Для проведення експериментальних досліджень на мові програмування R було створено скрипт, який здійснює основні операції, що притаманні JPEG компресії. А саме: розбиття початкового зображення на блоки фіксованого розміру 8x8 пікселів. Такі блоки не повинні перетинатись, але мають повністю покривати початкове зображення. По відношенню до отриманих блоків меншого розміру здійснювалось пряме, а для реконструкції зображення – зворотне, двовимірне дискретне косинусне перетворення (ДКП). Отримані ДКП коефіцієнти кантувались за

допомогою базової, рекомендованої в стандарті JPEG [12], таблиці квантування, що відповідає показнику якості $QF = 50$ (див. таб. 6.1).

Таблиця 6.1 – Базова таблиця квантування для показника якості 50

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

Аби отримати таблиці квантування для інших показників якості, елементи базової таблиці масштабуються за формулою

$$T_s[i] = \left\lfloor \frac{SF \cdot T_b[i] + 50}{100} \right\rfloor, \quad (6.1)$$

$$\text{де } SF = \begin{cases} 5000 / QF, & \text{якщо } QF \leq 50, \\ 200 - 2 \cdot QF, & \text{якщо } 50 < QF \leq 100, \end{cases} \quad T_b[i] \text{ -- відповідні}$$

елементи таблиці 6.1, що підлягають масштабуванню.

Для порівняння (оцінки якості) стиснених з різними показниками якості зображень із оригінальним зображенням було використано дві метрики : SSIM, яка обчислювалась відповідно до формули (4.15), та пікове відношення сигнал/шум (Peak Signal to Noise Ratio, PSNR), яке обраховувалось за формулою

$$PSNR = 10 \lg \frac{255^2}{\sigma^2}, \quad (6.2)$$

де 255 – максимальне значення яскравості пікселів зображення, а σ^2 - дисперсія шуму компресії

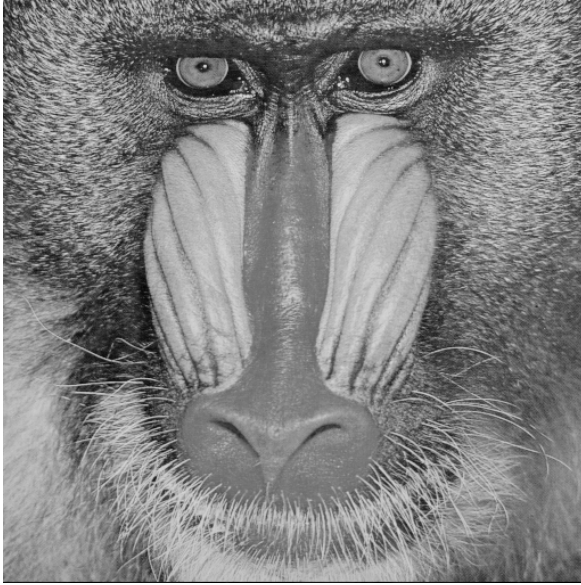
$$\sigma^2 = \frac{1}{N \cdot M} \sum_{i=1}^N \sum_{j=1}^M (I[i, j] - I^*[i, j])^2, \quad (6.3)$$

тут N , M - висота і ширина зображення в пікселях, $I[i, j]$ - піксель оригінального зображення з координатами i та j , а $I^*[i, j]$ - відповідний піксель стисненого зображення.

PSNR є метрикою, що традиційно використовується для оцінки якості стиснених зображень і може бути легко обчислена за наявності двох зображень: оригінального і спотвореного в наслідок стиснення або інших перетворень зображення. Наприклад, шляхом додавання білого гаусового шуму. Для обчислення значення SSIM використовувалась відповідна функція пакету SpatialPack (детальніше див. [26]).

Мета дослідження полягала в порівнянні двох метрик SSIM та PSNR, що їх було використано саме для порівняння оригінального і стисненого зображень. Також було приділено увагу іншим видам спотворення аніж компресія, конкретно – додавання до зображення адитивного білого гаусового шуму.

Для проведення експерименту використовувались два напівтонові тестові зображення: “baboon” (рис. 6.1 а) та “boats” (рис. 6.1 б) розміру 512x512 бітів кожне. Зображення “baboon” має в своєму складі багато дрібних деталей, тобто є текстурним за характером. Натомість, зображення “boats” містить в своєму складі суттєві області монотонності, наприклад, небо або борти човнів. Зазначені зображення відтворюють основні особливості типових зображень, які створюються людиною. Саме тому їх було обрано в якості тестових.



а)



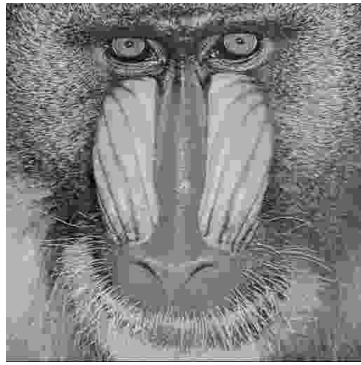
б)

Рисунок 6.1 – Тестові зображення формату pgm: а)“baboon”; б)“boats”

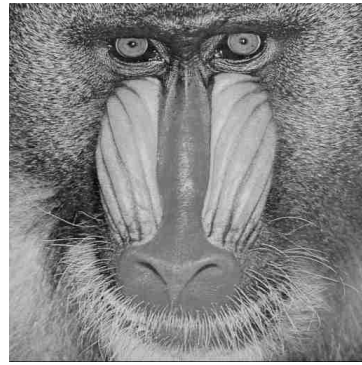
Тестові зображення піддавались компресії, що відповідає значенням показника якості $QF=5(5)95$, тобто показник якості змінювався із кроком 5 від 5 до 95. Загалом було отримано 19 стиснених копій кожного з тестових зображень. Приклади стиснених копій зображень “baboon” і “boats” наведено на рис. 6.2 та рис. 6.3 відповідно.

Аналіз зображень з рис. 6.2 і 6.3 свідчить про те, що візуальна помітність спотворень, які виникають в наслідок компресії, швидко спадає із збільшенням значення показника якості QF . На тестурному за своєю природою зображенні “baboon” спотворення є менш помітними навіть за дуже низьких значень показника якості. Останнє не так для зображення “boats”, що містить в своєму складі значні області монотонності, де значення пікселів залишаються майже незмінними.

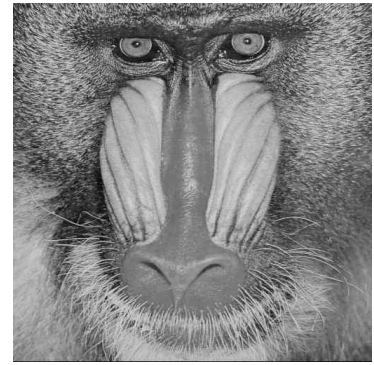
Порівняємо між собою здатність SSIM та PSNR до відображення змін у якості стиснених зображень. Застосуємо для обчислення SSIM формулу (4.15) безпосередньо до пікселів всього зображення (як стисненого, так і оригінального). Тобто спочатку ми не будемо слідувати рекомендаціям щодо обчислення індексу SSIM.



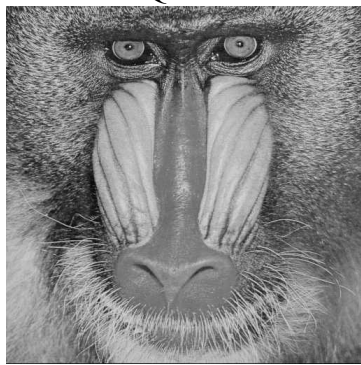
QF=5



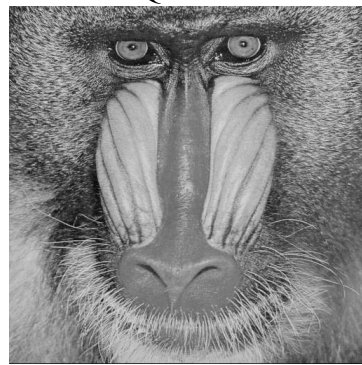
QF=20



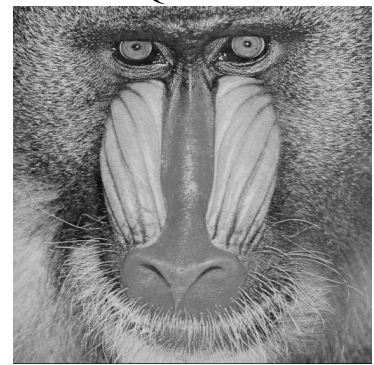
QF=35



QF=50



QF=65



QF=80

Рисунок 6.2 - Приклади стиснених копій зображення “baboon”



QF=5



QF=20



QF=35



QF=50



QF=65



QF=80

Рисунок 6.3 - Приклади стиснених копій зображення “boats”

На рис. 6.4 наведено залежність оціненого значення індексу SSIM від PSNR, які були обраховані для кожної з 19 стиснених копій зображень “baboon” і “boats” (приклади яких розміщено на рис. 6.2 та 6.3).

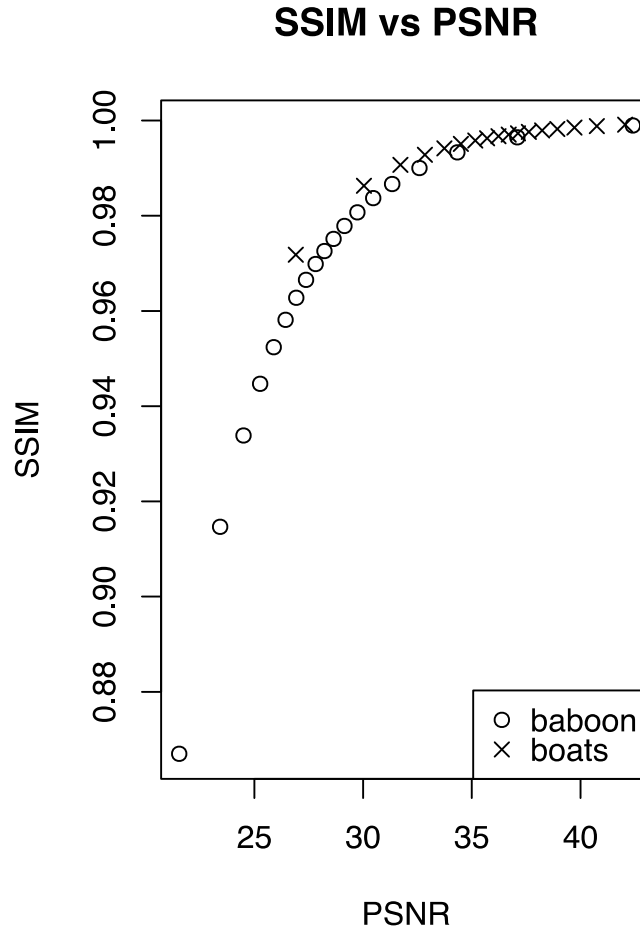


Рисунок 6.4 – Залежність значень індексу SSIM від значень PSNR, обрахованих для стиснених з різними показниками якості тестових зображень

З аналізу даних, представлених на рис. 6.4, робимо висновок, що значення метрики PSNR змінюються в досить широкому діапазоні від 20 дБ до 40 дБ. Зауважимо, що значенням $PSNR \geq 40$ дБ відповідає найбільша якість зображення, коли спотворення виявляються важко помітними якщо їх взагалі вдається локалізувати на зображенні яке проходило обробку. Натомість для зображення “boats” значення SSIM змінюються від 0,97 до майже 1, тобто в дуже малому діапазоні. Отже

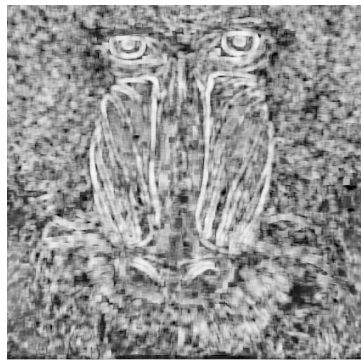
стиснені із низьким показником якості зображення мають бути віднесені до дуже подібних до оригіналу. Аналогічна ситуація спостерігається і для зображення “baboon”, хоча значення SSIM і змінюються у трохи більшому діапазоні. Крім того, компресії зображення “baboon” з показником якості $QF=5$ відповідає суб’єктивно краща копія аніж це має місце для зображення “boats”. Однак, індекс SSIM у зображення “baboon” при $QF=5$ є значно меншим ніж індекс SSIM у зображення “boats” при тому ж значенні $QF=5$. При цьому різниця між відповідними значеннями PSNR є не такою великою (не більше від 6 дБ).

Таким чином, доходимо висновку, що PSNR краще за SSIM, який було обчислено по відношенню до всього зображення, без розбиття на менші частини, відображає зміни якості у стиснених за алгоритмом JPEG зображеннях.

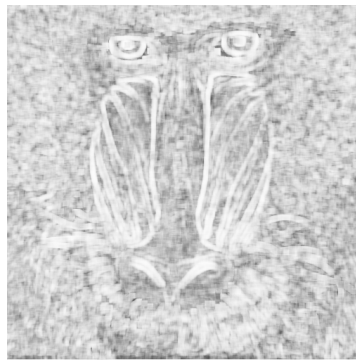
Дещо кращих результатів вдається досягти, якщо виконувати рекомендації до обчислення SSIM, які було розміщено в попередньому розділі. А саме: обчислювати локальні значення індексу SSIM за допомогою ковзного вікна розміру 8×8 пікселів. Результуюче значення індексу SSIM тоді обчислюється шляхом усереднення по всіх отриманих локальних значеннях. На рис. 6.5 та 6.6 показані карти якості, що відповідають стисненим із різними показниками якості зображенням “baboon” і “boats”.

Аналізуючи карти якості на рис. 6.5 і рис. 6.6 звертаємо увагу, що карти якості вказують на те, що зображення “boats” має кращу якість в порівнянні з зображенням “baboon” (більше світлих областей). Хоча це і суперечить інтуїції: спотворення як правило є менш помітними на текстурних зображеннях.

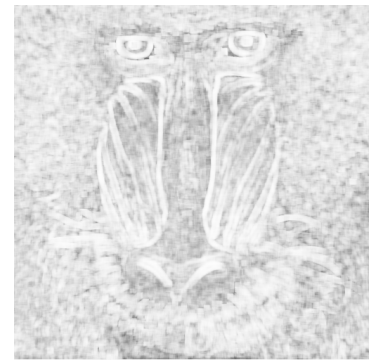
На рис. 6.7 наведено залежності індексу структурної подібності SSIM, обраховані по відношенню до всього зображення в цілому (globally applied) і з застосуванням технології ковзного вікна (locally applied) від відношення PSNR для зображень “baboon” та “boats”.



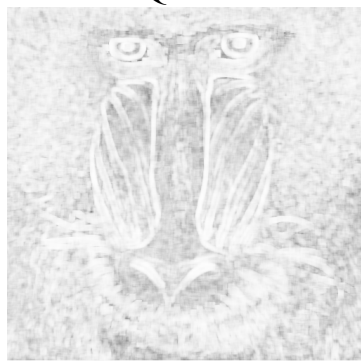
QF=5



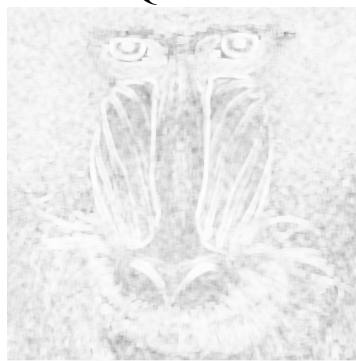
QF=20



QF=35



QF=50

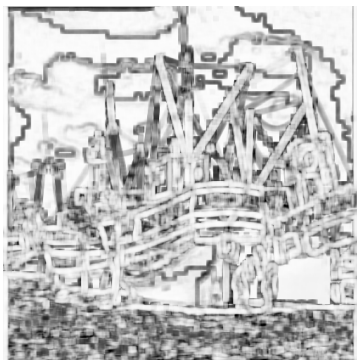


QF=65

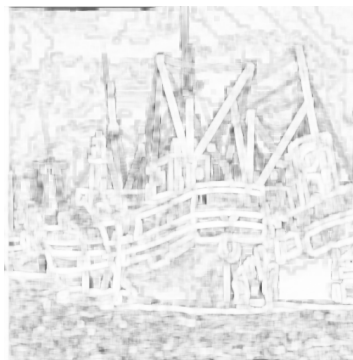


QF=80

Рисунок 6.5 - Приклади карт якості для стиснених копій зображення “baboon”



QF=5



QF=20



QF=35



QF=50



QF=65



QF=80

Рисунок 6.6 - Приклади карт якості для стиснених копій зображення “boats”

Цікаво проаналізувати вплив різних типів шумів на значення індексу SSIM. Зокрема шумів квантування (вже досліджено) та шумів у вигляді адитивного білого шуму. Фіксуючи значення PSNR, обчислимо величину середньоквадратичного відхилення σ_g еквівалентного білого гаусового шуму з нульовим середнім як

$$\sigma_g = 255 \cdot 10^{-\frac{PSNR}{20}}. \quad (6.4)$$

Будемо додавати відповідний білий гаусів шум до тестових зображень таким чином, аби величина PSNR залишалась тією самою, як при компресії зображень (тобто дисперсія нормального шуму і дисперсія шуму квантування співпадатимуть). На рис. 6.8 наведено залежності індексу SSIM від PSNR для зображень “baboon” та “boats”.

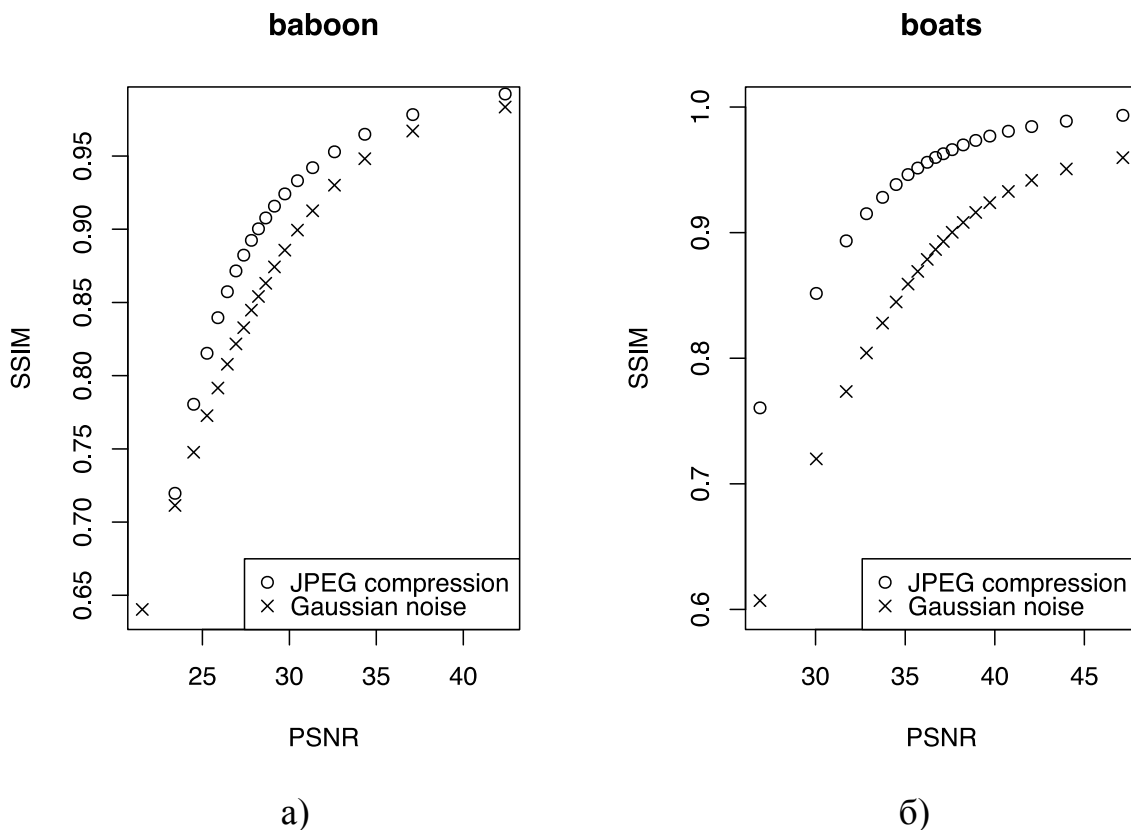
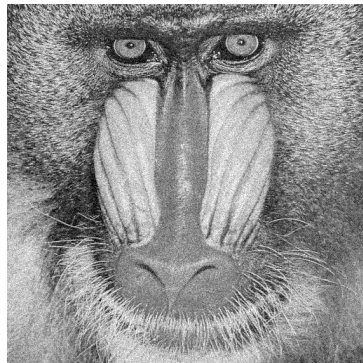
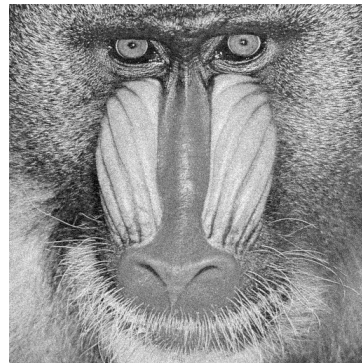


Рисунок 6.8 - Залежності індексу SSIM, обраховані і з застосуванням технології ковзного вікна від відношення PSNR для зображень: а) “baboon” та б) “boats” та різних типів шуму: JPEG компресії, або нормального

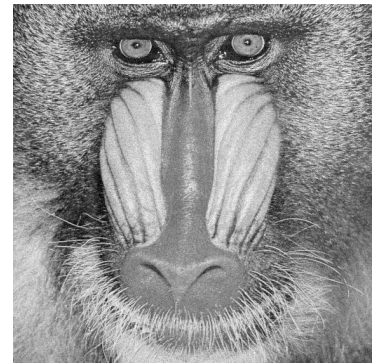
На рис. 6.9 та 6.10 показані приклади зашумлених тестових зображень, в свою чергу на рис. 6.11 і 6.12 дано відповідні карти якості.



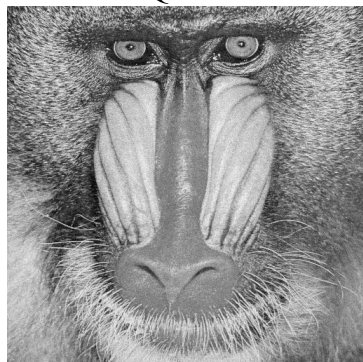
QF=5



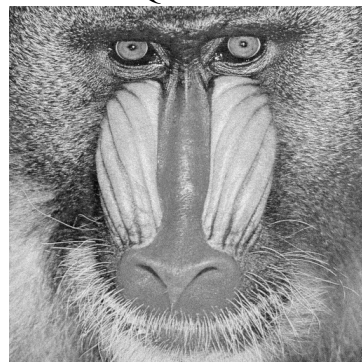
QF=20



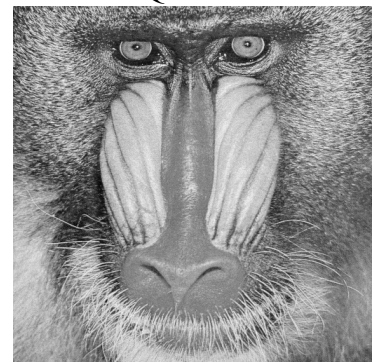
QF=35



QF=50



QF=65



QF=80

Рисунок 6.9 - Приклади зашумлених копій зображення “baboon”



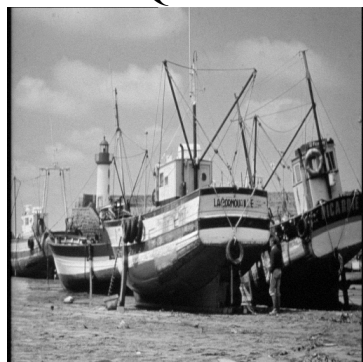
QF=5



QF=20



QF=35



QF=50

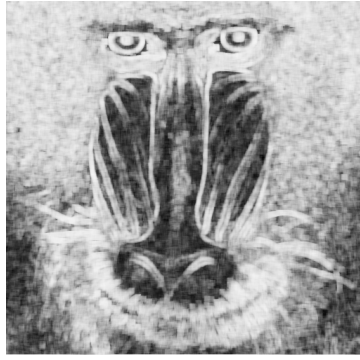


QF=65

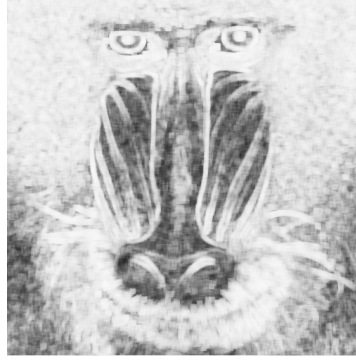


QF=80

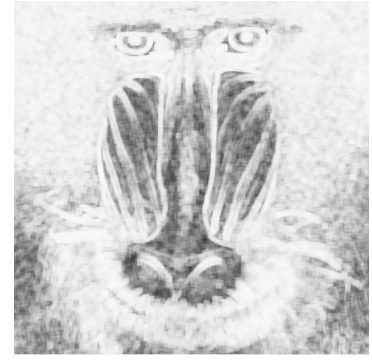
Рисунок 6.10 - Приклади зашумлених копій зображення “boats”



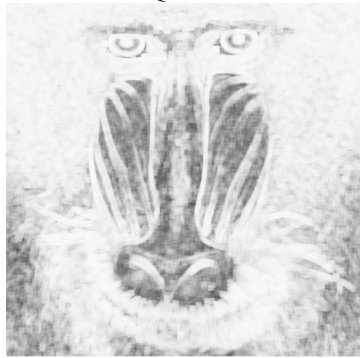
QF=5



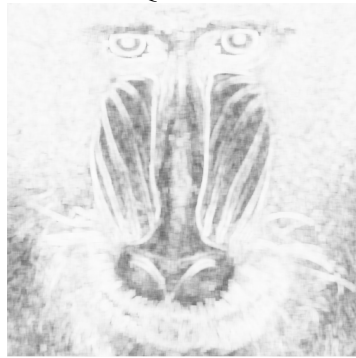
QF=20



QF=35



QF=50



QF=65



QF=80

Рисунок 6.11 - Приклади карт якості для зашумлених копій зображення “baboon”



QF=5



QF=20



QF=35



QF=50



QF=65



QF=80

Рисунок 6.12 - Приклади карт якості для зашумлених копій зображення “boats”

Аналіз залежностей на рис. 6.8 дає можливість стверджувати, що значення індексу структурної подібності SSIM є більш чутливими (спостерігаємо менші значення) до додавання білого шуму, аніж до JPEG компресії. Цей висновок також підтверджують карти якості, що їх зображено на рис. 6.11 і 6.12. З іншого боку, безпосередній візуальний аналіз прикладів стиснених (див. рис. 6.2 та 6.3) і зашумлених (див. рис. 6.9 та 6.10) зображень, вказує на те, що стиснені зображення (особливо монотонні) є менш візуально приємними в порівнянні із зашумленими. Останнє знову ж таки свідчить на користь метрики PSNR як суттєво простішої в обчислювальному плані, але досить точної.

ВИСНОВКИ

В роботі розглянуто основні ідеї та алгоритми структурного підходу до оцінки якості зображень. Зокрема, описано концепцію, алгоритм обчислення індексу SSIM, а також алгоритм оцінки якості зображення на основі синтезу зображень у просторі зображень. Продемонстровано, що спотворення зображення вздовж різних напрямків у просторі зображень мають різне для сприйняття значення. Структурний підхід намагається відокремити напрямки, пов'язані зі структурними спотвореннями, від напрямків з неструктурними спотвореннями. Цей поділ дає нову систему координат в просторі зображення. Нова система координат не є фіксованою, як у традиційних системах декомпозиції зображення (наприклад, Фур'є та вейвлет-перетворення), а адаптована до структур зображення, що лежать в його основі.

З точки зору побудови систем оцінювання якості зображення, більшість традиційних методів, заснованих на СЗЛ, базуються на філософії «знизу вгору», яка намагається змоделювати функції відповідних компонентів СЗЛ і об'єднати їх разом, в надії, що об'єднана система зможе передбачити поведінку всієї СЗЛ. Ефективність цих методів залежить від того, наскільки добре розуміється СЗЛ і наскільки точно може бути реалізоване моделювання. На противагу цьому, структурні підходи базуються на філософії «зверху вниз», яка починається з самого верхнього рівня - моделювання гіпотетичної функціональності всієї СЗЛ. Підхід «зверху вниз» здатний призвести до значного спрощення алгоритму, але залежить від правильності основної гіпотези. Зокрема, основне припущення структурного підходу полягає в тому, що СЗЛ добре пристосована до вилучення структурної інформації з візуальної сцени, і тому міра структурних спотворень повинна добре передбачати якість сприйняття зображення. Багато експериментів продемонстрували, що це припущення справджується.

ПЕРЕЛІК ПОСИЛАНЬ

1. Wu, H. R., & Rao, K. R. (Eds.). (2017). Digital video image quality and perceptual coding. CRC Press.
2. Chou, C. H. (1995, April). Adaptive transform coding of images based on removing just noticeable distortion. In Visual Communications and Image Processing'95 (Vol. 2501, pp. 607-618). SPIE.
3. Hecht, S. (1924). The visual discrimination of intensity and the Weber-Fechner law. *The Journal of General Physiology*, 7(2), 235.
4. Schreiber, W. F. (1986). *Fundamentals of Electronic Imaging Systems*, volume 15 of Springer Series in Information Sciences.
5. Shen, D. F., & Wang, S. C. (1996, September). Measurements of JND property of HVS and its applications to image segmentation, coding, and requantization. In *Digital Compression Technologies and Systems for Video Communications* (Vol. 2952, pp. 113-121). SPIE.
6. Mannos, J., & Sakrison, D. (1974). The effects of a visual fidelity criterion of the encoding of images. *IEEE transactions on Information Theory*, 20(4), 525-536.
7. Sharma, D., & Netravali, A. (1977). Design of quantizers for DPCM coding of picture signals. *IEEE Transactions on communications*, 25(11), 1267-1274.
8. Limb, J., & Rubinstein, C. (2003). On the design of quantizers for DPCM coders: A functional relationship between visibility, probability and masking. *IEEE Transactions on Communications*, 26(5), 573-578.
9. Schäfer, R. (1983). Design of adaptive and nonadaptive quantizers using subjective criteria. *Signal Processing*, 5(4), 333-345.
10. Glassner, A. S. (1995) *Principles of Digital Image Synthesis*. San Francisco, CA: Morgan Kaufman.

11. Leger, A. M., Omachi, T., & Wallace, G. K. (1991). JPEG still picture compression algorithm. *Optical Engineering*, 30(7), 947-954.
12. Information Technology — Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines. Technical report. (1993) ISO/IEC JTC1 10918-1. ITU-T Rec. T.81.
13. Joint Photographic Experts Group. (2000). JPEG2000 image coding system. ISO/IEC FCD 15444-1.
14. Christopoulos, C., Skodras, A., & Ebrahimi, T. (2002). The JPEG2000 still image coding system: an overview. *IEEE transactions on consumer electronics*, 46(4), 1103-1127.
15. Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4 (12), 2379-2394.
16. Ruderman, D. L. (1996). The statistics of natural images, *Network: Computation in Neural Systems*, 5, 517-548.
17. Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual review of neuroscience*, 24(1), 1193-1216.
18. Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. *Sensory Communication*, W. A. Rosenblith, ed., MIT Press, Cambridge, 1(01), 217-233.
19. Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4), 600-612.
20. Wang, Z., Bovik, A. C., & Simoncelli, E. P. (2005). Structural approaches to image quality assessment. *Handbook of image and video processing*, 7(18).
21. Wang, Z., & Bovik, A. C. (2001). Embedded foveation image coding. *IEEE Transactions on image processing*, 10(10), 1397-1410.
22. Wang, Z., & Bovik, A. C. (2002). A universal image quality index. *IEEE signal processing letters*, 9(3), 81-84.

23. Wang, Z., Bovik, A. C., & Lu, L. (2002, May). Why is image quality assessment so difficult?. In 2002 IEEE International conference on acoustics, speech, and signal processing (Vol. 4, pp. IV-3313). IEEE.

24. Wang, Z., & Simoncelli, E. P. (2004, June). Stimulus synthesis for efficient evaluation and refinement of perceptual image quality metrics. In Human Vision and Electronic Imaging IX (Vol. 5292, pp. 99-108). SPIE.

25. Kornblum, J. D. (2008). Using JPEG quantization tables to identify imagery processed by software. *Digital Investigation*, 5, S21-S25.

26. Vallejos, R., Osorio, F., & Bevilacqua, M. (2020). Spatial relationships between two georeferenced variables. Springer. [https://doi.org/10, 1007, 978-3](https://doi.org/10.1007, 978-3).