

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ОДЕСЬКИЙ ДЕРЖАВНИЙ ЕКОЛОГІЧНИЙ УНІВЕРСИТЕТ

Б.В.ПЕРЕЛИГІН

ЦИФРОВА ОБРОБКА ЗОБРАЖЕНЬ

Конспект лекцій

Одеса
Одеський державний екологічний університет
2023

УДК 004.67:528.71:528.77
П 27

Перелигін Б.В.

П27 Цифрова обробка зображень: конспект лекцій. Одеса : Одеський державний екологічний університет, 2023. 186 с.

ISBN 978-966-186-235-6

В конспекті лекцій розглянуті поняття про безперервне й цифрове зображення та про одержання цифрового зображення з безперервного, викладені принципи дії основних видів пристроїв, які дозволяють одержати цифрові зображення, наводяться способи і методи цифрової обробки зображень, особливості обробки кольорових зображень, космічних знімків.

Даний конспект лекцій призначений для бакалаврів спеціальності 122 Комп'ютерні науки які навчаються за освітньою програмою Комп'ютерні науки. Може використовуватися бакалаврами спеціальності 193 Геодезія та землеустрій для поглиблення знань в галузі моніторингу навколишнього середовища.

УДК 004.67:528.71:528.77

Рекомендовано методичною радою Одеського державного екологічного університету Міністерства освіти і науки України як конспект лекцій

(протокол 4 від 02.03. 2023 р.)

ISBN 978-966-186-235-6

© Б.В.Перелигін, 2023

© Одеський державний екологічний університет, 2023

ЗМІСТ

	стор.
Вступ	5
1 Поняття про зображення	6
1.1 Математична модель безперервного зображення.....	8
1.2 Зорова система людини.....	10
1.3 Моделі кольорового зображення.....	14
2 Цифрове зображення	18
2.1 Дискретизація безперервного зображення.....	18
2.2 Квантування безперервного зображення.....	19
2.3 Кодування й типи цифрових зображень.....	23
2.4 Стискання зображень.....	24
2.5 Формати подання цифрових зображень і відео.....	28
3 Одержання цифрових зображень	32
3.1 Елементарний сенсор.....	33
3.2 Сканер.....	37
3.3 Цифрові фотоапарати.....	38
3.4 Одержання зображень за допомогою ультразвуку.....	44
3.5 Рентгенівські зображення.....	45
3.6 Томографічне зображення.....	47
3.7 Зображення, одержувані при дистанційному зондуванні Землі.....	51
4 Обробка цифрових зображень на основі їхніх геометричних перетворень	61
4.1 Класифікація методів цифрової обробки зображень.....	61
4.2 Геометричні перетворення зображень.....	64
5 Перетворення зображень точкового типу	81
5.1 Логічні операції над зображеннями.....	81
5.2 Лінійні перетворення яскравості напівтонових зображень.....	82
5.3 Арифметичні операції над зображеннями.....	87
5.4 Нелінійна корекція яскравості зображень.....	88
5.5 Перетворення яскравості на базі гістограми зображення..	95
5.6 Локально-адаптивна обробка зображень.....	97
6 Перетворення локального типу	101
6.1 Типи шумів.....	102

6.2	Фільтрація зображень.....	104
6.3	Низькочастотні фільтри.....	107
6.4	Високочастотні фільтри.....	110
6.5	Нелінійна фільтрація.....	115
6.6	Внесення ефектів у зображення.....	120
7	Операції математичної морфології із зображеннями.....	124
7.1	Математична морфологія бінарних зображень.....	124
7.2	Математична морфологія напівтонових зображень.....	133
8	Обробка кольорових зображень.....	135
8.1	Корекція кольорів на зображеннях.....	135
8.2	Автоматичне обчислення порога бінаризації кольорового зображення.....	138
8.3	Підвищення чіткості кольорового портрета.....	140
9	Сегментація зображень.....	143
9.1	Сегментація зображень за яскравістю на основі граничної обробки.....	144
9.2	Статистичний підхід до текстурної сегментації.....	149
9.3	Сегментація методом морфологічного водорозділу.....	153
10	Цифрова обробка космічних знімків.....	154
10.1	Особливості зображень на космічних знімках у різних ділянках спектра.....	154
10.2	Етапи і зміст обробки космічних знімків.....	159
10.3	Програмні засоби цифрової обробки супутникових зображень.....	182
	Використані джерела.....	185

ВСТУП

Зображення – це відбиття природних, виникаючих у природі сигналів, тобто хвиль різної довжини, які можуть бути зареєстровані зоровою системою людини або приладами. Причому око людини сприймає хвилі в досить вузькій ділянці електромагнітного спектра, який називається видимим. Все видиме оком, а також реєстроване приладами, споконвічно являє собою безперервні сигнали, які утворюють аналогові зображення. Вони описуються безперервними функціями, а сучасні комп'ютери обробляють інформацію в цифровому виді. Тому обробка зображень є одним із самих затребуваних напрямків сучасної інформатики.

Інтерес до методів цифрової обробки зображення виростає із двох основних областей її застосування, якими є підвищення якості зображення для поліпшення його візуального сприйняття людиною й обробки зображення для їхнього зберігання, передачі й подання в автономних системах машинного зору.

Таким чином, цифрова обробка зображення тісно пов'язана з розвитком цифрової обчислювальної техніки. Для цифрових зображень потрібно таку велику пам'ять й обчислювальну потужність, що прогрес в області цифрової обробки зображень у значній мірі визначається розвитком комп'ютерів і допоміжних технологій для зберігання, відображення й передачі даних.

Перші комп'ютери з потужністю, достатньою для виконання завдань цифрової обробки зображень, з'явилися на початку 1960-х років. Народження того, що ми сьогодні називаємо цифровою обробкою зображення, простежується з моменту виникнення таких машин і початку вивчення Землі з космосу, зокрема з аналізу зображень поверхні. Паралельний прогрес у цих областях пустив у хід потужний потенціал ідей цифрової обробки зображення.

У всіляких галузях науки й техніки є безліч завдань, пов'язаних з обробкою зображень. Це, наприклад, аналіз й інтерпретація аерофотознімків і космічних знімків у науках про Землю, у метеорології, астрономії, дослідженні природних ресурсів; аналіз оптичних рентгенівських, теплових, радіографічних й інших зображень у медичній діагностики, промислової дефектоскопії, наукових дослідженнях. Аналіз зображення й полів, зареєстрованих методом оптичної, радіо- і акустичної голографії.

У зв'язку із цим у вузівських програмах з'являються дисципліни, спрямовані на вивчення принципів обробки зображень, причому пріоритетна увага приділяється цифровим методам, які приваблюють своєю гнучкістю.

1 ПОНЯТТЯ ПРО ЗОБРАЖЕННЯ

Зображення – це відбиття природних, виникаючих у природі сигналів, тобто хвиль різної довжини, які можуть бути зареєстровані зоровою системою людини або приладами. Причому око людини сприймає хвилі в досить вузькій ділянці електромагнітного спектра, що називається видимим. Все видиме оком, а також реєстроване приладами, споконвічно являє собою безперервні сигнали, що утворюють аналогові зображення. Вони описуються безперервними функціями, а сучасні комп'ютери обробляють інформацію в цифровому виді. Тому обробка зображень є одним із самих затребуваних напрямків сучасної інформатики.

Процес реєстрації відбитого безперервного сигналу й перетворення його в цифрове подання називається одержанням цифрового зображення. Перетворення також називають обробкою зображень, вона виконується за певними алгоритмами за допомогою програм, що реалізують ці алгоритми.

Слід зазначити, що в науково-технічній літературі часто використовуються загальнонаукові терміни, які не мають однозначного визначення, а словники також дають різні тлумачення цих понять. Наприклад, метод й алгоритм – синоніми, а в іншому джерелі це різні поняття. В об'єктно-орієнтованому програмуванні поняття метод деградувало до визначення: метод – це функція або процедура, яка належить якомусь класу або об'єкту. Слід зазначити, що інформаційні технології – це область людської діяльності, яка балансує між наукою (зокрема, математикою) і технікою (конкретною реалізацією наукових ідей), тому тут не зовсім строгі визначення також представляють якийсь науково-технічний симбіоз.

Базу інформаційних технологій становлять дані й інформація. Пояснимо, у чому різниця. Нехай є множина фактів, вони можуть бути представлені як дані у формі фраз, зображень, звукозапису й т.п. **Дані** – це будь-які відомості, представлені в придатному для розуміння й обробки вигляді. Основними типами даних є графічні (відео, фото, креслення) і символні (числа, букви). Форми подання даних: аналогова (напис на папері), дискретна (фільм на кіноплівці має дискретність по часовій шкалі) і цифрова (пісня у форматі MP3). Ті самі цифрові дані можна записати в різних форматах, наприклад, зображення у форматі JPEG, число у форматі із плаваючою комою, бінарні дані у вигляді нулів й одиниць. **Інформація** – це дані, у яких усунута надмірність їхнього опису, тобто це оброблені дані. Опис того самого факту може бути представлено даними різного типу (текстом, звуком, зображенням). Інформацію на відміну від даних можна виміряти відповідно до теорії К. Шеннона. Основна форма подання інформації – цифрова. У цифровому виді дані можуть бути оброблені на електронній обчислювальній машині (ЕОМ). Вони представляються

нулями й одиницями, а їхня інтерпретація залежить від формату подання, тобто той самий набір нулів й одиниць може бути фрагментом зображення, пісні або книги.

У результаті абстрагування даних одного або декількох класів виникають **знання**. Наприклад, якщо нагріти алюмінієву деталь, то вона розплавиться. Сталева деталь теж розплавиться, але при більшій температурі. Це знання, отримані в результаті спостережень, осмислення різних фактів й їхнього узагальнення.

У контексті інформаційних технологій варто усвідомити наступний набір понять: **метод, методика, технологія**. **Науковий метод** – це сукупність способів одержання нових знань і методів рішення задач. **Метод** пояснює, чому і як з набору даних можна одержати інші дані або вирішити певну задачу. Наприклад, відповідно до молекулярно-кінетичної теорії в алюмінієвому сплаві при нагріванні до N градусів змінюється молекулярна структура й він стає рідким, тому його можна залити у форму, понизити температуру до N градусів. Зв'язки між молекулами відновлюються й можна одержати твердий сплав заданої форми. **Метод в інформатиці** – це узагальнений спосіб рішення задач певного класу (який обґрунтовує чому і як у межах цього класу можна застосовувати даний метод з гарантованим одержанням результату). Наприклад, метод декомпозиції – основний метод паралельних обчислень: задача розбивається на більш прості підзадачі, вирішивши які, можна одержати рішення вихідної задачі шляхом комбінації рішень окремих підзадач. **Методика** – це перелік дій для одержання деякого результату, але на відміну від методу вона не пояснює, чому треба робити саме так (наприклад, методика навчання плаванню). При застосуванні методики досягнення результату не гарантовано, це поняття частіше використовується в сфері навчання. Так, застосовуючи деяку методику, можна й не домогтися бажаного результату, але для досягнення одного результату можуть використатися різні методики. На відміну від методики **технологія** гарантує одержання запланованого результату при виконанні відповідних умов (наприклад, технологія виробництва чавуну). **Інформаційна технологія** описує чітко впорядковану послідовність застосування засобів одержання й зберігання, обробки, передачі й одержання цифрових даних для створення деякого інформаційного продукту.

Звідси **теорія інформації** – це математична теорія, що включає в себе опис й оцінки методів добування, передачі, зберігання й класифікації інформації, а **інформатика** – область людської діяльності, пов'язана з перетворенням цифрових даних засобами обчислювальної техніки й взаємодії цих засобів із середовищем застосування. Ключове поняття інформатики – **алгоритм** – це набір точних інструкцій, що описують порядок дій для рішення задачі за кінцевий час. Алгоритм може базуватися

на деякому методі, але він не пояснює, чому треба робити так, а не інакше, і носить чисто прикладний характер. Алгоритм, як правило, не орієнтований на мову програмування й операційну систему, тобто це набір абстрагованих інструкцій універсального призначення. Комп'ютерна програма – це послідовність інструкцій, закодованих певною мовою програмування, призначена для виконання за допомогою ЕОМ у певній операційній системі, тобто це реалізація алгоритму. Один алгоритм можна реалізувати різними способами, навіть програмуючи на одній мові.

1.1 Математична модель безперервного зображення

Будемо називати зображенням функцію двох змінних $f(x,y)$, визначену в деякій області S площини Oxy і яка має деяку множину своїх значень. Наприклад, звичайну чорно-білу фотографію можна представити як $f(x, y) \geq 0, 0 \leq x \leq a, 0 \leq y \leq b$, де $f(x,y)$ – яскравість (іноді називана оптичною щільністю або ступенем білизни або тоном) зображення в точці (x, y) ; a – ширина; b – висота кадру.

З урахуванням особливостей функції f звичайно виділяють наступні класи зображень (рис. 1.1):

- 1) напівтонові,
- 2) бінарні,
- 3) лінійні,
- 4) точкові.

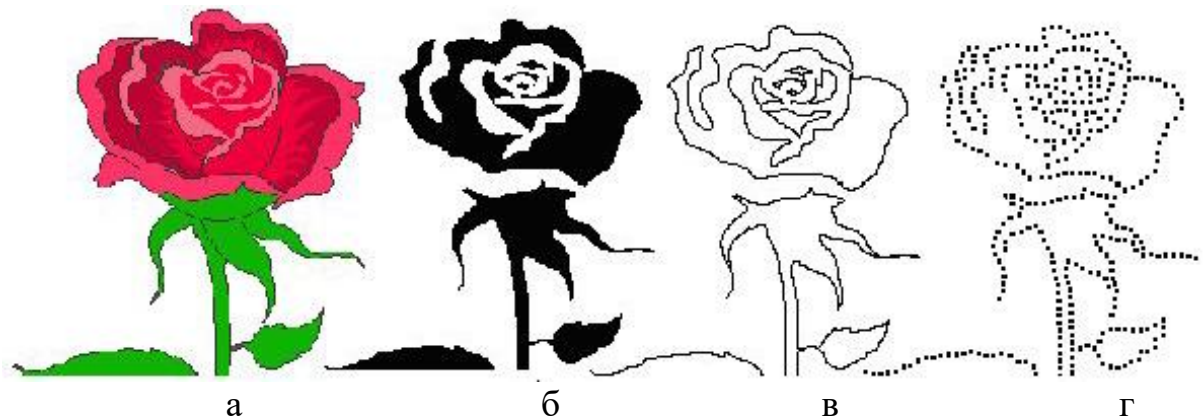


Рисунок 1.1 – Види зображень: а – напівтонове; б – бінарне; в – лінійне; г – крапкове

Прикладами напівтонових (сірих) зображень (рис. 1.1, а) можуть служити звичайні чорно-білі фотографії. Множина значень функції в S може бути дискретним ($f \in \{f_0, \dots, f_l\}, l > 1$) або безперервним ($0 \leq f \leq f_{max}$). В один клас з напівтоновими звичайно поєднують кольорові фото- і телевізійні зображення.

Експериментально встановлено, що практично будь-які кольори, видимі людиною (але не всякий), можуть бути представлені як сума певних значень m_1 , m_2 , m_3 трьох лінійно-незалежних так званих основних кольорів:

$$f = m_1 \cdot E_1 + m_2 \cdot E_2 + m_3 \cdot E_3.$$

Відповідно до прийнятої системи RGB основними є кольори, що відповідають монохроматичним випромінюванням з довжиною хвилі $\lambda_1 = 0,7$ мкм (червоний – R), $\lambda_2 = 0,5461$ мкм (зелений – G), $\lambda_3 = 0,4358$ мкм (синій – B). Тут множина значень функції f являє собою тривимірний простір інтенсивності (m_1 , m_2 , m_3) основних кольорів. Якщо $m_1 = m_2 = m_3 = m$, то залежно від величини m виходять різні відтінки сірих кольорів – від чорного до білого. Якщо $m_1 \neq m_2 \neq m_3$, то зображення пофарбоване в різні кольори.

При космічній й аерозйомці застосовують багатоспектральну фотографічну, телевізійну і сканувальну апаратуру, спектральні канали можуть лежати як у видимому, так й у невидимих УФ й ІЧ діапазонах спектра; число таких каналів може досягати $n = 5 \dots 10$ і більше. Тут множина значень функції f являє собою n -мірний простір інтенсивності (m_1, m_2, \dots, m_n). При візуалізації таких зображень на моніторах ЕОМ кожному вектору із цього простору можна привласнити який-небудь видимий людським оком колір, не обов'язково основний. У результаті виходить аеро- або космічне зображення в умовних кольорах (говорять – у псевдокольорах).

Людське око не розрізняє відтінки сірого зображення, яскравість яких відрізняється друг від друга менш чим на 1...2 %, але добре розрізняє кольори. За допомогою комп'ютера різним рівням яскравості напівтонового зображення можна зіставити деяку кольорову шкалу й пофарбувати зображення в так названі неправдиві (ложні) кольори. Правильно підібрана шкала дозволяє операторові візуально виділяти об'єкти, які мало відрізняються за яскравістю.

Бінарні (дворівневі) зображення (рис. 1.1, б). Тут множина значень функції $f \in (0, 1)$. Звичайно 0 – це чорні кольори, а 1 – білий. Прикладом бінарного зображення є карти водойм, карти плоских деталей на конвеєрі й т.д. Лінійні зображення (рис. 1.1, в). Типовий приклад – карта ізоліній, зображення берегової лінії, різні контури. Зображення являє собою одну криву або сімейство кривих. Точкові зображення (рис. 1.1, г). Типовий приклад – кадр ділянки зоряного неба. Зображення являє собою k точок з координатами (x_i, y_i) і яскравістю $f_i, i = 1, \dots, k$.

1.2 Зорова система людини

Наше око (як і спостережна апаратура) сприймає деякий сигнал, відбитий у певній ділянці електромагнітного спектра.

Спектр – це сукупність значень або їхній розподіл за яким-небудь параметром спостережуваної величини. Оскільки першими були отримані оптичні спектри, то дотепер найбільше часто термін «спектр» вживається для опису розподілу потоку випромінювання або часток за довжинами хвиль або енергії.

Електромагнітні випромінювання підрозділяються: на радіохвилі, інфрачервоне випромінювання, видиме оком світло, ультрафіолетове випромінювання, рентгенівське випромінювання та гамма-випромінювання. Між ними немає чітких границь, області цих видів випромінювання можуть частково перекриватися. Шкала, у якій вони вимірюються, називається електромагнітним спектром (рис. 1.2). Виміри частіше здійснюють в метрах.

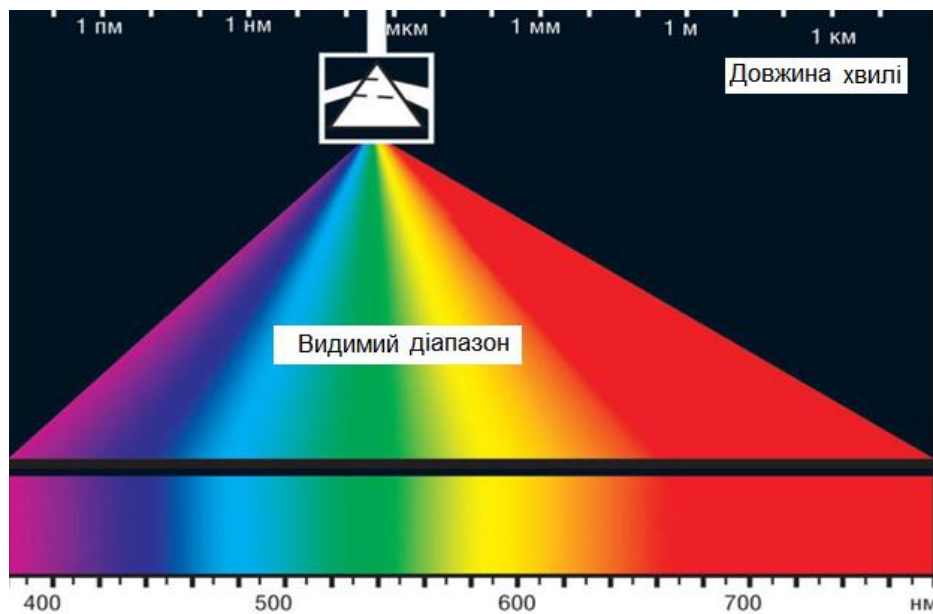


Рисунок 1.2 – Шкала електромагнітного спектра й видимого світла

Око людини сприймає тільки частина відбитих хвиль, які попадають у так званий видимий діапазон електромагнітного спектра. До 80 % інформації людина одержує у вигляді зображень, тому важливо розуміти механізми зорового сприйняття.

Зорова система – це оптико-біологічна бінокулярна система, яка еволюційно виникла у тварин і здатна сприймати випромінювання видимої ділянки електромагнітного спектра, створюючи зображення у вигляді сенсорного відчуття положення предметів у просторі, їхніх розмірів і кольорів.

Око людини забезпечує адаптивне настроювання сприйняття зображень відповідно до зовнішніх умов. Око являє собою систему лінз, які формують на сітківці зменшене й перевернене зображення зовнішнього миру (рис. 1.3), містить у собі роговицю, кришталік, склоподібне тіло, за яким лежить сприймаюче світло сітківка.

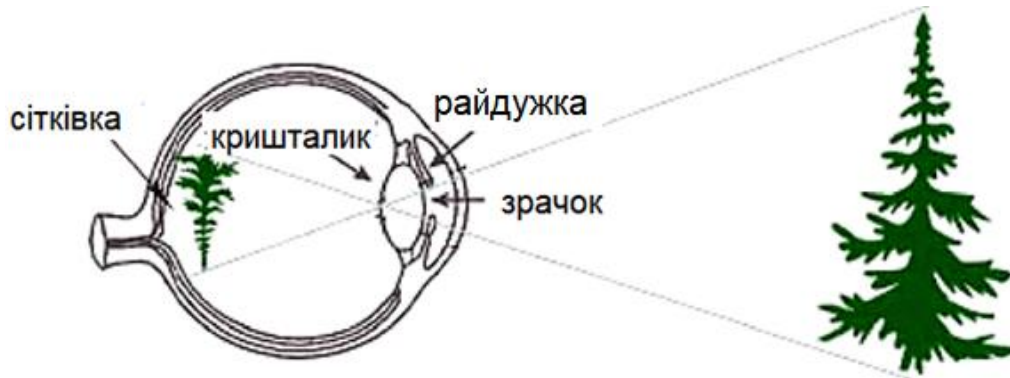


Рисунок 1.3 – Схема відображення видимого предмета на сітківці ока

Важливою характеристикою зорової системи є **чутливість**, тобто здатність реагувати на зовнішні зміни. Вона характеризується верхніми й нижніми абсолютними порогоми. Існує кілька видів чутливості. **Світлова чутливість** характеризує властивість ока реагувати на мінімальний світловий потік. Однак слід зазначити, що ймовірність розпізнавання мінімального світлового потоку залежить від різних факторів, наприклад, кута зору. Зорова система по-різному реагує на випромінювання, які рівні по потужності, але мають різні довжини хвиль, тобто належать різним ділянкам спектра. Така чутливість називається **спектральною чутливістю** або чутливістю до кольорового тону. Людське око сприймає хвилі електромагнітного спектра в діапазоні від 380 до 780 нм. Крім того, зорова система характеризується **чутливістю до насиченості кольорів**. Здатність ока розрізняти мінімальні розходження яскравості суміжних областей зображення характеризується **контрастною чутливістю**.

Чутливість зорової системи наведених типів не є постійною, а залежить від багатьох факторів, зокрема умов освітлення. Наприклад, при переході з темної кімнати у світлу потрібно якийсь час для відновлення світлочутливості ока. Цей процес називається яскравісною адаптацією ока.

Колірорідчуття описують три основні характеристики: **світлота, кольірний тон і насиченість**.

У тварин і людини органами зору є очі, а зір бінокулярний, він забезпечує об'ємне подання зображення. В оці людини містяться два типи світлочутливих клітин (рецепторів): високочутливі **палички** й менш чутливі **колбочки** (рис. 1.4). Вони перетворюють світлові роздратування в нервові збудження. Палички функціонують при світлі слабкої

інтенсивності, у сутінках. Зорові відчуття, одержувані при цьому, безбарвні. Колбочки функціонують вдень і при яскравому освітленні. Їхня функція – визначення відчуття кольоровості.

Палички – це периферичні відростки світлочутливих клітин сітківки ока, названі так за свою витягнуту циліндричну форму. У сітківці ока людини міститься близько 120 млн. паличок довжиною 0,06 мкм і діаметром 0,002 мкм. Палички чутливі до світла завдяки наявності в них специфічного пігменту – родопсину. Під дією світла відбуваються дуже швидкі перетворення й знебарвлення зорового пігменту. Чутливість палички достатня, щоб зареєструвати влучення на них навіть двох-трьох фотонів. Щільність розміщення паличок на різних ділянках сітківки ока нерівномірною й може становити від 20 до 200 тис. на кв. мм. Причому на периферії сітківки їхня щільність вище, ніж у її середині, що визначає їхню участь у нічному й периферійному зорі. У центрі сітківки в жовтій плямі палички практично відсутні. Палички мають цікаву особливість. У зв'язку з тим, що всі вони містять той самий світлочутливий пігмент родопсин, їхня спектральна характеристика залежить від рівня освітлення. При слабкому освітленні максимум поглинання родопсину становить близько 500 нм (спектр сутінкового неба) і при цьому палички відповідальні за сутінковий зір, коли кольори предметів нерозрізнені.

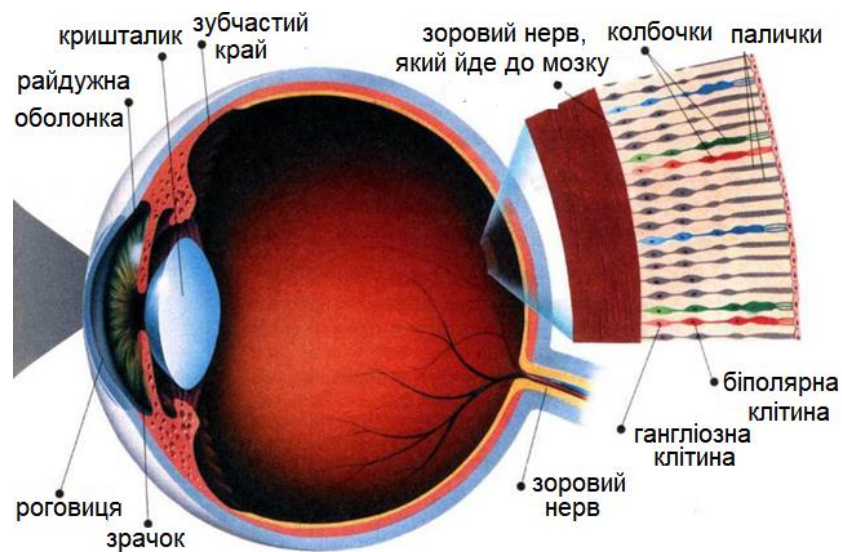


Рисунок 1.4 – Будова ока людини

До іншого типу фоторецепторів відносяться колбочки, названі так за свою конічну форму. Вони чутливі до світла завдяки наявності в них пігменту йодопсину. У свою чергу, йодопсин складається з декількох зорових пігментів. На сьогоднішній день добре досліджені два з них: хлоролаб (чутливий до жовто-зеленої області спектра) і еритролаб (чутливий до жовто-червоної частини спектра). У сітківці ока у дорослої

людини зі 100 % зором налічується близько 6...7 млн колбочок довжиною близько 50 мкм і діаметром від 1 до 4 мкм. Колбочки приблизно в 100 разів менш чутливі до світла, чим палички, але набагато краще сприймають швидкі зміни світла. Розрізняють три види колбочок за чутливістю до різних довжин хвиль світла (кольорам) (табл. 1.1, рис. 1.5).

Таблиця 1.1 – Характеристики різних видів колбочок

Тип колбочок	Сприймані довжини хвиль, нм	Максимум чутливості, нм
<i>S</i>	380.....500	420
<i>M</i>	450.....630	534
<i>L</i>	500.....750	564

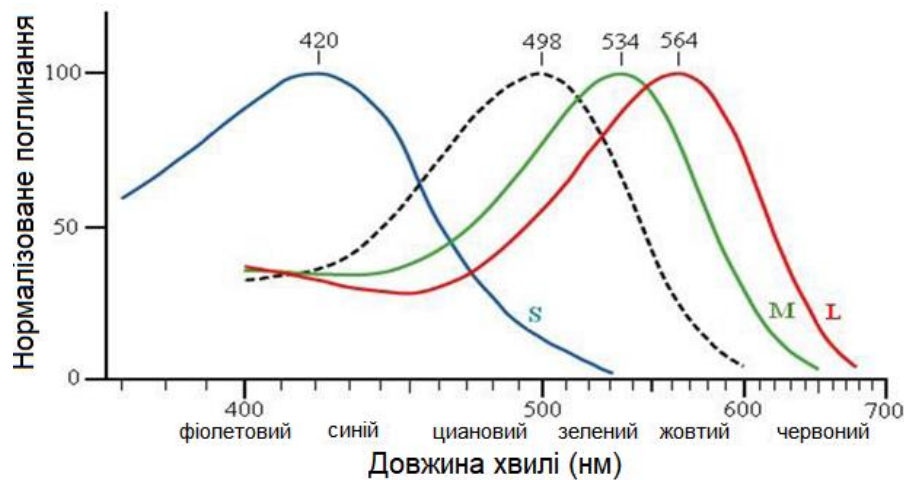


Рисунок 1.5 – Графіки світлочутливості колбочок людського ока *S*, *M*, *L*-типів (пунктиром показана сутінкова, чорно-біла сприйнятливність паличок)

Колбочки *S-типу* (Short – короткохвильовий спектр) чутливі у фіолетово-синій, *M-типу* (Medium – середньо-хвильовий) – у зелено-жовтій й *L-типу* (Long – довго-хвильовий) – у жовто-червоної областях спектра. Наявність трьох видів колбочок і паличок дозволяє людині розрізняти кольори.

Водойма здається синьою тому, що молекули води поглинають всі промені видимого спектра, крім синіх. Це помітно тільки при його глибині більше метра, тому калюжі не виглядають синіми. У нічний час, коли потік фотонів недостатній для нормальної роботи колбочок, зір забезпечують тільки палички, тому вночі людина не може розрізняти кольори.

Розподіл типів колбочок у сітківці нерівномірний: «сині» колбочки перебувають ближче до периферії, у той час як «червоні» й «зелені» розподілені випадковим образом. (це стало відомо в 1970-х рр. і було підтверджено більше детальними дослідженнями на початку ХХІ ст.)

Відповідність типів колбочок трьом «основним» кольорам забезпечує розпізнавання людиною тисяч відтінків кольорів. Криві спектральної чутливості трьох видів колбочок частково перекриваються (рис. 1.5), що сприяє явищу метамерії. Дуже сильне світло збуджує всі три типи рецепторів і тому сприймається як випромінювання сліпучо-білих кольорів. Рівномірне роздратування всіх трьох елементів, що відповідає середньозваженому денному світлу, викликає відчуття білих кольорів.

Відповідно до трикомпонентної теорії кольорів за колірний зір людини й мавп відповідають гени, що кодують світлочутливі білки опсини трьох типів, які реагують на різні довжини хвиль, що є достатнім для колірного сприйняття. У більшості ссавців таких генів тільки два, тому вони мають чорно-білий зір. Якщо у людини два білки, кодовані різними генами, виявляються занадто схожі або один з білків не синтезується, у нього розвивається дальтонізм.

1.3 Моделі кольорового зображення

Уперше трикомпонентну теорію колірного зору описав М.В. Ломоносов в 1756 р. в роботі «Про походження світла». Ще через сто років її розвинув німецький вчений Г. Гельмгольц.

Для опису кольорів розроблено багато моделей. Термін «колірна модель» означає абстрагований опис кольорів у вигляді набору трьох чисел, названих колірними координатами. Множина припустимих відтінків кольорів й метод інтерпретації цих даних визначають колірний простір.

Більшість кольорів виходять змішуванням декількох основних кольорів. Тому для опису кольорів використовується поняття колірної моделі як способу подання великої кількості відтінків кольорів за допомогою розкладання їх на прості складові. Виникає питання: навіщо потрібні різні колірні моделі? Одна із причин – випромінюваний і поглинений колір описуються по-різному.

В 1931 р. після серії експериментів з оцінки сприйняття кольорів людиною Міжнародна комісія з освітлення розробила стандарт CIE1931XYZ, який описував колірний простір, що вміщає в себе всі сприймані людиною кольори. Однак цей простір (для стислості XYZ) було істотно нелінійним.

З появою кольорового телебачення й кольорової печатки виникла необхідність у розробці нових моделей, орієнтованих на різні додатки.

Найбільш популярною колірною моделлю на сьогоднішній день є RGB (Red – червоний, Green – зелений, Blue – синій). У ній всі три кольори рівноправні й незалежні. Якщо три параметри RGB мають нульові значення, вони описують чорні кольори, якщо максимальні значення – білий. Ця модель гарна для коліровідтворення на пристроях із чорним

базовим тлом, наприклад, на випромінювальному моніторі. RGB – це адитивна колірна модель. Відображення реальних кольорів на моніторі залежить від виробника монітора й налаштувань.

Якщо базове тло біле, модель RGB застосовувати важко (наприклад, білий папір поглинає кольори). У цьому випадку зручніше використати субтрактивну модель, наприклад, CMY (Cyan – голубий, Magenta – пурпурний, Yellow – жовтий). «Субтрактивна» означає «що віднімається», тобто з білого віднімаються основні кольори. Вона є додатковою до моделі RGB і фактично визначає негативні відтінки, тобто від максимального значення яскравості в моделі RGB (звичайно це 255) віднімаються власно відтінки кольорів: $C=255-R$, $M=255-G$, $Y=255-B$. Цю модель активно використовують у поліграфії при друці на білому папері. Теоретично рівна кількість фарби $C=M=Y$ повинна давати сірі кольори, однак при друці такі кольори на папері мають бурі відтінки, крім того, витрачаються зайві фарби. Тому на практиці модель доповнюють чорними кольорами, і називається вона CMYK. Для переходу з моделі CMY у модель CMYK використовують наступні співвідношення:

$$\begin{aligned}K &= \min(C, M, Y); \\C &= C - K; \\M &= M - K; \\Y &= Y - K.\end{aligned}$$

Кожне число в моделі CMYK визначає відсоток фарби базових кольорів, які становлять колірну комбінацію. Наприклад, для одержання кольорів хакі змішують 30 частин блакитної фарби, 45 частин пурпурної, 80 – жовтої й 5 – чорної. Це можна позначити в такий спосіб: CMYK = (30, 45, 80, 5) або C30M45Y80K5. Важливо відзначити, що числове значення фарби в CMYK не описує відтінок кольорів. Цифри – це набір даних, використовуваних у друкарському процесі для формування зображення. Реальні кольори залежать від багатьох факторів: розміру пікселя на фотовиводі, розтискування фарби, якості паперу, вологості в цеху, спектральних характеристик джерела освітлення й ін.

В 1976 р. Міжнародною комісією з освітлення була розроблена колірна модель CIE Lab, що є міжнародним стандартом. Використовувані в ній величини означають: L – світлота (змінюється від 0 до 100, тобто від самого темного до найяснішого); a – червона й зелена кольоровості; b – жовта й синя кольоровості. У ній значення світлоти відділене від значення хроматичної складової кольори (тону й насиченості). При розробці Lab переслідувалася мета створення моделі колірного простору, зміна кольорів у якому буде максимально лінійним з погляду людського сприйняття, тобто однакова зміна значень координат кольору в різних областях колірного простору повинна робити однакове відчуття зміни кольору. Це

дозволяє математично коректувати нелінійність сприйняття кольору людиною. Модель Lab однозначно визначає кольори, тому вона застосовується в обробці зображень у якості проміжного колірному простору, через яке відбувається конвертування даних між іншими колірними просторами (наприклад, з моделі RGB сканера в СМΥК друкувального пристрою). При цьому подання кольорів в Lab активно використовується у кольорокорекції, можна окремо впливати на яскравість, контраст зображення і його кольори. Недоліком моделі Lab є використання кубічних коренів при перерахуванні.

Модель HSV (Hue – тон, Saturation – насиченість, Value – значення) зручна для створення зображень художниками. У ній координатами кольорів є: Hue – колірний тон, звичайно варіюється в межах $0^{\circ} \dots 360^{\circ}$, де 0° – червоний, 120° – зелений, 240° – синій (тон кольорів задається в градусах); Saturation – насиченість, змінюється в межах $0 \dots 100$ або $0 \dots 1$; Value – яскравість, задається в межах $0 \dots 100$ або $0 \dots 1$.

Колірне охоплення означає діапазон кольорів, які можуть бути представлені колірною моделлю пристрою щодо кольорів, сприйманих людським оком і моделлю Lab. Різні колірні охоплення мають електронно-променева трубка монітора й телевизор, поліграфічні фарби й око людини. На рис. 1.6 показано порівняння колірних охоплень людського ока (модель Lab), монітора (модель RGB) і друкувальної машини (модель СМΥК) у площині, обумовленої осями $x = X/(X + Y + Z)$ і $y = Y/(X + Y + Z)$ у просторі XYZ.

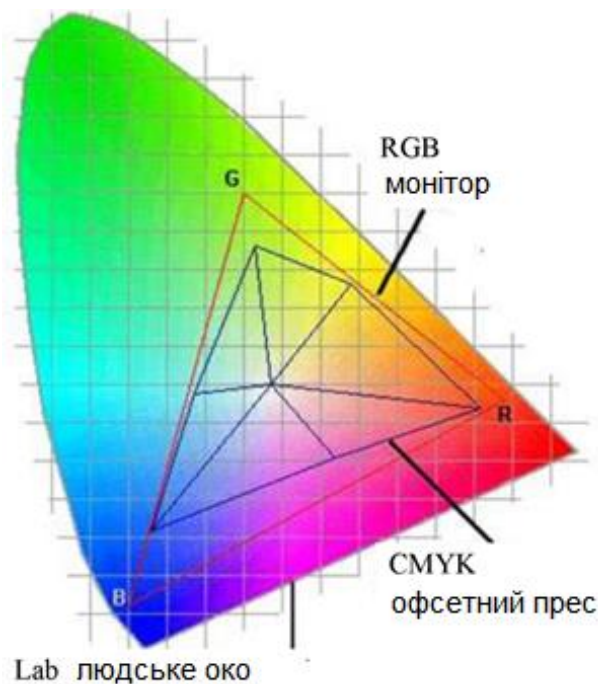


Рисунок 1.6 – Колірне охоплення різних пристроїв

Різні колірні моделі використовуються в різних додатках: XYZ й Lab – для опису сприйняття кольору людським оком; адитивні моделі – для одержання кольорів на моніторі (наприклад, RGB, sRGB); поліграфічні моделі – для одержання кольорів на папері при використанні різних фарб і поліграфічного устаткування (наприклад, CMYK); моделі, не пов'язані з фізичними процесами устаткування, орієнтовані на передачу даних; математичні моделі – для кольорокорекції без зв'язку з устаткуванням (наприклад HSV). Математично колірні моделі можна однозначно перетворити одне в одну, але пристрої відображення мають різні технічні можливості й колірні охоплення.

Паралельно із трикомпонентною розроблялася опонентна теорія кольорів. Вона виходила з того, що в мозок надходить інформація не про червоний (R), зелений (G) і синій (B) кольори (як у теорії кольорів Гельмгольца), а про різницю яскравості білого (I_{\max}) і чорного (I_{\min}) кольорів, зеленого й червоного ($G-R$), синього й жовтого ($B-Y$). Наприклад, жовтий колір ($Y = R + G$) – це сума червоного й зеленого кольору. Параметри кольорів у цій моделі описуються трьома коефіцієнтами

$$K_{\text{grey}} = I_{\max} - I_{\min}; K_{\text{gr}} = G - R; K_{\text{brg}} = B - R - G,$$

де K_{grey} , K_{gr} , K_{brg} – коефіцієнти для будь-якого типу освітлення.

Опонентна теорія в цілому краще пояснює той факт, що люди сприймають кольори предметів однаково при різних джерелах освітлення, у тому числі при різному кольорі джерел світла в одній сцені (це властивість колірної адаптації). Дві теорії не зовсім погоджені одна з одною. Однак більшість вчених припускає, що на рівні сітківки діє трьохстимульна теорія (її основний представник – модель RGB), потім інформація обробляється й у мозок надходять дані, що вже погодяться з опонентною теорією. Слід також зазначити, що в будь-якій моделі (за винятком CMY) кольори описуються трьома параметрами.

Контрольні запитання

1. Що таке зображення ?
2. Що таке дані ?
3. Що таке інформація ?
4. Що таке метод в інформатиці ?
5. Що таке методика ?
6. Що таке технологія ?
7. Що таке інформаційна технологія ?
8. На які класи поділяються зображення ?
9. Що таке напівтонове зображення ?

10. Що таке бінарне зображення ?
11. Що таке лінійне зображення ?
12. Що таке крапкове зображення ?
13. Як побудоване око людини ?
14. Що таке світова чутливість ока ?
15. Що таке спектральна чутливість ока ?
16. Що таке контрастна чутливість ока ?
17. Що таке чутливість ока до насиченості кольорів ?
18. Що таке колірна модель зображення ?
19. Які існують колірні моделі ?

2 ЦИФРОВЕ ЗОБРАЖЕННЯ

2.1 Дискретизація безперервного зображення

У природі всі процеси протікають безупинно. Зорова система людини також сприймає безперервні сигнали. У сучасній техніці навпаки, всі частіше дані представлені в цифровій формі.

Процес перетворення аналогового сигналу, що описує зображення, у цифрову форму складається із трьох етапів: *дискретизації, квантування і кодування*. Сигнал, до якого застосовані операції дискретизації й квантування, називається цифровим. Після операції дискретизації, але без квантування одержуємо дискретний сигнал (його значення змінюються в безперервному діапазоні).

Дискретизація – це перетворення безперервної функції в дискретну, тобто значення функції визначені в окремих (дискретних) точках. При цьому самі значення можуть змінюватися в безперервній шкалі. Інакше кажучи, множина припустимих значень дискретної функції може бути нескінченним (рис. 2.1, вертикальна вісь).

Заміну безперервного зображення дискретним можна виконати різними способами. Розглянемо основні особливості прямокутної дискретизації.

Нехай $x(t_1, t_2)$ – безперервне зображення, розглянуте як функція від двох змінних; $x(i_1, i_2)$ – відповідне йому дискретне зображення, отримане з безперервного шляхом розбивки області, у якій визначена функція x , на фрагменти, як правило, прямокутної форми. Зв'язок між функціями визначається виразом

$$x(t_1, t_2) = x(i_1 \Delta t_1, i_2 \Delta t_2),$$

де $\Delta t_1, \Delta t_2$ – вертикальний і горизонтальний інтервали дискретизації.

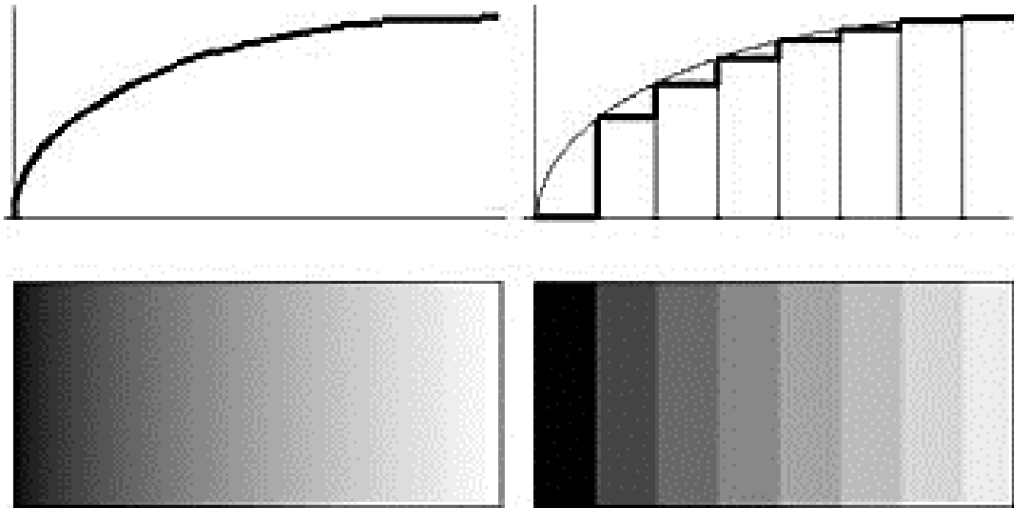


Рисунок 2.1 – Рівномірна дискретизація безперервної функції яскравості

Нескінченна множина точок з координатами (t_1, t_2) області, у якій визначена безперервна функція яскравості, замінюється кінцевою множиною з координатами (i_1, i_2) . Звичайно $\Delta t_1 = \Delta t_2$ і цей параметр Δt визначає відрізнення зображення. У техніці дискретизація визначається кількістю точок на дюйм (DPI) або пікселів на дюйм (PPI) і дорівнює цілому від ділення дюйма на Δt .

Основне питання при заміні безперервного зображення дискретним складається у визначенні умов, при яких така заміна мінімізує втрату інформації, що міститься в безперервному зображенні. Втрати відсутні, якщо, маючи у своєму розпорядженні дискретний сигнал, можна повністю відновити вихідний безперервний сигнал. З математичної точки зору питання полягає у відновленні безперервного сигналу у двомірних проміжках між вузлами, у яких його значення відомі, тобто в здійсненні двомірної інтерполяції.

На рис. 2.2 зображені відбитки пальця при різних інтервалах дискретизації. Праве зображення сильно розфокусоване й невірно передає структуру відбитка, воно непридатне для розпізнавання.

2.2 Квантування безперервного зображення

При цифровій обробці зображень безперервний діапазон значень функції яскравості поділяється на три піддіапазони і всі значення яскравості, які потрапили в один піддіапазон, округляються до єдиного значення, наприклад, середини цього діапазону. Така процедура називається *квантуванням за яскравістю*, і вона перетворить безперервну множину значень функції x у дискретну множину значень $x_{k\delta}$. Інакше кажучи, функція приймає кінцеву множину значень $\{r_1, \dots, rL\}$. Ці значення

називаються рівнями квантування. У загальному випадку перетворення виражається східчастою функцією (рис. 2.3).

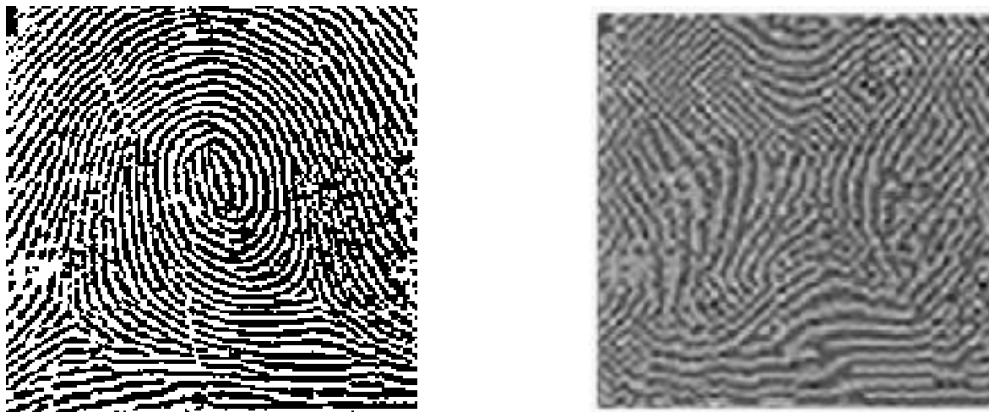


Рисунок 2.2 – Вплив інтервалу дискретизації на зображення відбитка пальця

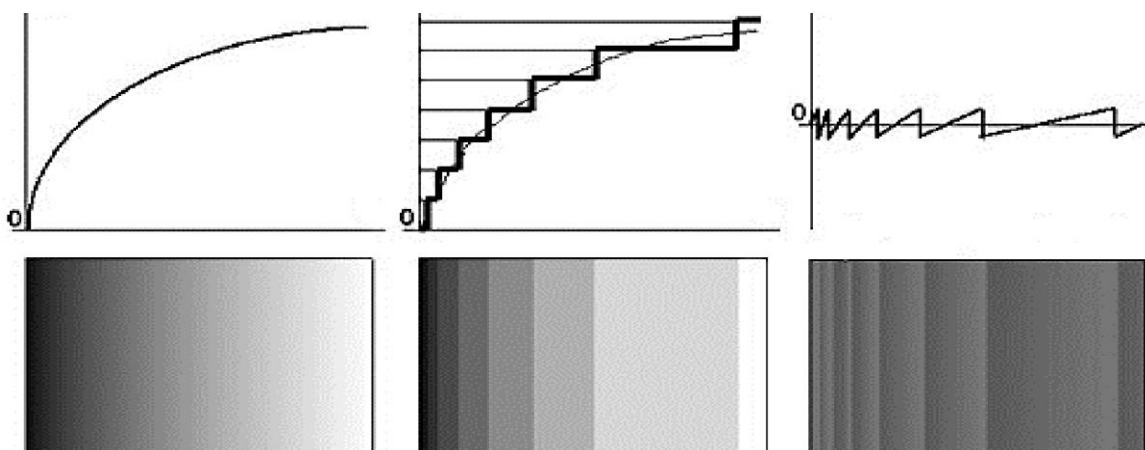


Рисунок 2.3 – Безперервний сигнал (ліворуч); функція, що описує його рівномірне квантування (у центрі); функція похибки квантування (праворуч)

Якщо значення яскравості x попадає в діапазон $d_j < x \leq d_{j+1}$, то воно замінюється на число, яке дорівнює рівню квантування r_j , де $d_j, j = \overline{1, L + 1}$, – пороги квантування. При цьому динамічний діапазон значень яскравості всього зображення обмежений і дорівнює $[d_1, d_{+1}]$. Для простоти кодування цифрового зображення замість значення r_j використовується цілочисельний номер діапазону j , починаючи з нуля. Таким чином, значення яскравості, використовувані в описі кольору, наприклад, у моделі RGB, є абстрактними безрозмірними величинами.

Не слід плутати процедури квантування й дискретизації та відповідно крок квантування із кроком дискретизації. При дискретизації

виконується просторова розбивка області зображення, при квантуванні розбивається множина значень яскравості.

Квантування можна виконати різними способами. Розглянемо кілька варіантів розбивки колірного простору.

Нехай $I(x, y)$ – яскравість пікселя вихідного зображення з координатами (x, y) ; $I_{new}(x, y)$ – яскравість пікселя нового зображення; NI – кількість відтінків у вихідному зображенні (звичайно $NI = 256$); N – кількість елементів у новій палітрі (наприклад, $N = 16$).

Алгоритм рівномірного квантування:

Крок 1. Організувати цикл по рядках і стовпцям зображення.

Крок 2. У циклі для кожного пікселя (x, y) обчислити нове значення яскравості $I_{new}(x, y) = I(x, y)/(NI/N)$. Результат округлити до цілого.

Для кольорового зображення повторити обчислення кроку 2 для кожного із трьох кольорів RGB. Розіб'ємо колірний простір на рівні частини уздовж кожної осі. Наприклад, у напрямку синьої або зеленої осі розіб'ємо куб на вісім частин, а в напрямку червоної – на чотири. У підсумку виходить 256 відтінків кольорів, рівномірно розподілених у просторі RGB. Кожен відтінок можна закодувати цілим індексом від 0 до 255, описавши в таблиці, який колір (тобто значення RGB) він представляє, одержимо індексоване або палітрове кольорове зображення. Колірна палітра – це фіксований набір кольорів і відтінків, описаний номерами (індексами).

Перевагою цього алгоритму є простота й висока швидкість. Недолік – малоконтрастні зображення будуть мати малу кількість відтінків, у місцях плавних переходів виникають помилкові контури.

Квантування методом медіанного перетину. Розглянемо алгоритм, що сформує палітру так, щоб кожне її значення відповідало рівній кількості значень відтінків у вихідному зображенні. Це досягається шляхом послідовної розбивки колірного простору на паралелепіпеди зі сторонами, паралельними осям колірного простору RGB.

Алгоритм медіанного квантування колірної палітри:

Крок 1. Задати N – розмір палітри. Знайти мінімальний паралелепіпед у просторі RGB, що включає всі значення відтінків кольорів пікселів вихідного зображення: R_{min} , R_{max} , G_{min} , G_{max} , B_{min} , B_{max} . Далі розбити паралелепіпед на частини.

Крок 2. Виконати так названу адаптивну розбивку, яка складається з наступних кроків: вибір самої довгої сторони паралелепіпед, сортування значень уздовж обраного напрямку, знаходження медіанного значення серед множини значень уздовж обраного напрямку, поділ паралелепіпед по знайденій медіані на дві частини. У такий спосіб вийдуть два паралелепіпеди, які містять приблизно однакову кількість пікселів. Повторювати попередню процедуру для кожного з паралелепіпедів доти, поки не сформується N паралелепіпедів, де N дорівнює заданому розміру

палітри. Якщо на якомусь етапі виявиться, що потрібно розділити паралелепіпед, що містить лише одне значення, то порожній паралелепіпед, що вийшов, варто приєднати до найбільшого паралелепіпеда.

Крок 3. Заповнити палітру або центральні значення RGB паралелепіпедів, або середніми арифметичними кольорів пікселів, що потрапили в дані паралелепіпеди, округливши їх до цілих.

Крок 4. Для кожного пікселя визначити, у який паралелепіпед він потрапив, і замінити його значення RGB параметрами паралелепіпеда, обчисленими на кроці 3.

Результатом роботи цього алгоритму є нерівномірне квантування. Воно краще попереднього передає відтінки, при цьому швидкість роботи даного алгоритму висока.

Квантування зображень методами кластеризації. Квантування можна виконувати методами, які роблять кластеризацію кольірних значень пікселів зображення для побудови палітри. У загальному випадку кластеризація – це процес розбивки об'єктів на групи (кластери) на основі загальних властивостей об'єктів. Стосовно до квантування зображень це означає розбивку всіх відтінків заданого зображення (у моделі RGB це буде куб) на кластери так, що усередині кожного кластера перебувають лише близькі значення. У результаті також одержимо нерівномірне квантування. Один із самих популярних алгоритмів кластеризації – це кластеризація методом *K*-середніх. Алгоритм будує *K* кластерів, розташованих на максимально більших відстанях друг від друга. Вибір числа *K* може базуватися на результатах попередніх досліджень або інтуїції. Загальна ідея алгоритму: задане фіксоване число *K* кластерів спостереження зіставляється кластерам так, що середні в кластерах максимально відрізняються друг від друга.

Алгоритм квантування кольори методом K-середніх.

Вихідні дані: кольорове зображення в просторі RGB, число кластерів *K*, максимальне число ітерацій *N*.

Крок 1. Визначити початкові центри кластерів. Вибір координат центрів можна виконати різними способами: випадково, вибрати значення відтінків R, G, B перших *K* пікселів і т.д. Для кольорів кожного пікселя знайти найближчий кластер, обчисливши відстань між ним і центрами всіх кластерів. У результаті кольори кожного пікселя будуть віднесені до одного з *K* кластерів.

Крок 2. Обчислити нові координати центрів всіх *K* кластерів, якими вважаються середні арифметичні значень R, G, B всіх пікселів, які входять у кластер. Кольори всіх пікселів зображення заново перерозподілити по *K* кластерах, обчисливши відстані між ними й знайденими центрами кластерів.

Крок 3. Процес обчислення центрів кластерів і перерозподілу

кольорів повторювати доти, поки не буде виконана одна з умов:

- 1) координати центрів всіх кластерів перестають змінюватися, тобто кольори належать тим же кластерам, що й на попередній ітерації;
- 2) число ітерацій досягає максимального заданого числа ітерацій N .

Недоліком даного методу є те, що він здатний ефективно формувати лише опуклі кластери. Перевагою – висока швидкість роботи. Стосовно до квантування зображень даний метод показує дуже гарні результати.

2.3 Кодування й типи цифрових зображень

У цифровій обробці зображень прийнято вважати, що зображення представляється матрицею цілих чисел розміром $N \times M$, де значення кожного елемента відповідає певному рівню квантування його енергетичної характеристики, або яскравості. Це так звана піксельна система координат.

У цьому випадку поняття піксель (англ. pixel – скорочено від picture cell – елемент зображення) має значення найменшого логічного елемента двомірного цифрового зображення.

Таким чином, **цифрове зображення** можна представити у вигляді **матриці**, елементи якої – пікселі – містять цілочисельні значення з деякої палітри. Нагадаємо, що ці значення безрозмірні, оскільки мають абстрактний характер.

За способом формування зображення діляться на **векторні й растрові**. Векторним називається зображення, описане у вигляді набору графічних примітивів, воно рисується лініями на графічних пристроях виводу. Растрове зображення являє собою двомірний масив, елементи якого містять інформацію про кольори. Воно орієнтовано на растрові пристрої відображення. У **цифровій обробці** використовуються **растрові зображення**, тому **векторні не розглядаємо**.

Растрові зображення можна поділити на чотири типи: **бінарні, напівтонові, палітрові, повнокольорові**. Їхня основна характеристика – **глибина кольорів**. Цей термін означає кількість біт, використовуваних для опису кольорів одного пікселя, і в літературі він позначається як bpp (від англ. bits per pixel).

Вибір типу зображення залежить від розв'язуваної задачі. Основні типи цифрових зображень та їхніх характеристик представлені в табл. 2.1.

Таблиця 2.1 – Основні типи цифрових зображень

Тип зображення	Основний діапазон значень яскравості	Глибина кольорів, біт	Кількість відтінків
Бінарне	[0, 1]	1	2
Напівтонове	[0, 255]	8	4-256
Палітрове (індексоване)	[0, 255]	8	2-256
Повнокольорове	[0, 255] ³	24	16 777 216

Елементи бінарного (дворівневого) зображення описуються одним бітом, що приймає значення 0 (чорний) і 1 (білий). Ці значення виходять у результаті обробки (бінарізації) напівтонових або кольорових зображень. Бінарні зображення зручні при передачі зображень документів у факсимільному зв'язку.

Напівтонове зображення складається з пікселів, які можуть приймати значення від 0 до N , де $N=2^k$. Це один з найпоширеніших типів зображень. У більшості випадків використовується значення $k=8$ біт (тобто один байт) на піксель.

Повнокольорові зображення містять інформацію про яскравість кольорів (звичайно R, G, B), тобто його можна розглядати як три напівтонових зображення.

У палітрових (індексованих) зображеннях значення пікселя, тобто індекс, фактично є посиланням на осередок палітри, що являє собою таблицю значень RGB для цього індексу. Для опису таких зображень необхідно зберігати палітру (таблицю 256 кольорів) і, власно, зображення у вигляді матриці індексів (один байт на піксель). При використанні нерівномірних алгоритмів квантування кольорів зорова система людини практично не розрізняє зображення, представлені в повнокольоровому та індексованому варіантах.

2.4 Стискання зображень

Зображення займають великий об'єм пам'яті, їхнє компактне зберігання й передача припускають стиск вихідних даних. Наприклад, при дистанційному зондуванні Землі (ДЗЗ) за допомогою супутників стискання даних представляє серйозну проблему. По супутникових каналах може передаватися до 2100 Мб/с. Тому алгоритми стиску насамперед орієнтовані на усунення надмірності двовимірних багатоспектральних даних таким чином, щоб мати можливість, якщо буде потреба, повністю відновити вихідне зображення.

Коефіцієнт стискання – основна характеристика алгоритму стискання. Він визначається як відношення обсягу вихідних нестиснутих даних до обсягу стиснутих.

Методи стискання зображень можна розбити на дві групи:

- стискання без втрат, що дозволяє точно відновлювати вихідне зображення;
- стискання із втратами (або стискання із регульованим ступенем перекручувань), зображення відновлюється приблизно.

Алгоритми стискання без втрат засновані на виключенні статистичної надмірності. Стискання без втрат означає, що процес оборотний, тобто значення яскравості (кольори) можна відновити при

декодуванні в первісному виді. До них відносять:

- кодування довжин серій (алгоритм RLE);
- метод групового кодування (алгоритм LZW);
- метод Хаффмана;
- арифметичне кодування й ін.

Одним з найпростіших алгоритмів стискування без втрат є алгоритм RLE (Run Length Encoding – групове кодування). Стиск в RLE відбувається за рахунок того, що у вихідному зображенні зустрічаються ланцюжки однакових байтів, які можна ощадливо кодувати. Різні модифікації алгоритму RLE реалізовані в утилітах, що записує зображення в різних форматах (PCX, GIF, TIFF, PMP). Слід зазначити, що багато сучасних форматів підтримують запис даних (не тільки зображень) з використанням декількох алгоритмів стискування або без них. Наприклад, формат TIFF 6.0 може зберігати зображення з використанням алгоритмів стискування RLE-PackBits, RLE-CCIT, LZW, Хаффмана з фіксованою таблицею кодів, JPEG-Lossless або без стискування. Алгоритм LZW (від прізвищ авторів Lempel-Ziv-Welch) є більше складним й універсальним алгоритмом (відносно RLE), що стискає зображення за рахунок пошуку, виділення й кодування однакових підланцюжків у потоці байтів. Алгоритм Хаффмана забезпечує стискування за допомогою попереднього обчислення частоти появи однакових байтів у зображенні. Говорячи про кращі на сьогодні алгоритмах стискування зображень без втрат, відзначимо, що в 1994 р. був прийнятий стандарт ISO/JPEG-Lossless. Він орієнтований на повнокольорові 24-бітові або 8-бітові напівтонові зображення. На 8-бітових тестових зображеннях декількох категорій (космічних, медичних, фото- і комп'ютерних графічних, змішаних) середній коефіцієнт стискування алгоритмом JPEG-Lossless дорівнювався двом.

Методи стискування із втратами засновані на тім, що в зображенні втримується інформація, слабко сприймана людським оком. Такий вид надмірності називається психофізіологічною надмірністю. Її зниження складається в зменшенні об'єму несуттєвої інформації, яку людське око не сприймає в силу психофізіологічних особливостей. При цьому частина інформації безповоротно губиться. Відновити вихідні значення яскравості або кольори всіх пікселів при декодуванні неможливо. Наприклад, око людини не розрізняє кольори дрібних деталей зображення. Отже, можна зберігати тільки яскравість деталей, а кольоровість відкинути й у такий спосіб зменшити об'єм файлу, що описує зображення. Цей метод застосовується в телебаченні. Розбивка безперервного відеопотоку на кадри також являє приклад виключення психофізіологічної надмірності.

Алгоритм стискування зображень – JPEG. Об'єднана група експертів в області фотографії (The Joint Photographic Experts Group – JPEG) в 1993 р. у рамках Міжнародної організації по стандартизації (ISO) розробила міжнародний стандарт JPEG для стиску повнокольорових (24-бітових)

зображень. Для нього розроблено чотири варіанти стискування:

- послідовний на базі дискретного косинусного перетворення (ДКП);
- прогресивний із записом коефіцієнтів ДКП, який використовує пірамідалне подання зображень;
- послідовний без втрат (JPEG-Lossless);
- ієрархічний із втратами й без втрат відповідно.

Алгоритм оперує блоками 8×8 пікселів; передбачається, що в них яскравість і кольори міняються порівняно плавно. До матриць значень таких блоків застосовується ДКП, після чого значимими виявляються тільки перші коефіцієнти. Система колірної сприйняття людини слабо розпізнає деякі частоти при колірній зміні картинки, тому можна апроксимувати коефіцієнти ДКП без помітної втрати якості відновленого зображення, використовуючи квантування коефіцієнтів. При такому перетворенні губиться частина інформації, але досягаються більші коефіцієнти стискування. Чому зображення розбивається на блоки 8×8 ? У цифровій обробці даних часто використовується ступінь двійки. Експериментально встановлено, що розбивка 8×8 дає кращий результат стискування, чим розбивка 4×4 . Застосування більшого розміру блоку незначно поліпшує якість відновленого зображення (при рівній кількості коефіцієнтів, які далі зберігають), але збільшує число обчислень. Наприклад, обчислення ДКП при розбивці 64×64 у два рази складніше, ніж для 8×8 . Таким чином, 8×8 – це експериментально обраний компроміс за критерієм якість-складність.

Хоча алгоритм JPEG і є стандартом ISO, формат його файлів не був зафіксований. Тому багато виробників заміняють внутрішні таблиці, рекомендовані стандартом, на свої власні й створюють свої формати JPEG, не завжди сумісні між собою.

Алгоритм JPEG дозволяє стиснути деякі зображення в 10...15 разів без візуально помітних перекручувань. Він орієнтований на стиск фотографій, що містять сцени із плавними переходами яскравості та кольорів. Найбільше поширення JPEG одержав у мережі Інтернет. Він мало придатний для стискування креслень і тексту, які мають різкий контраст між сусідніми пікселями, і що приводить до появи помітних артефактів. Подібні зображення краще зберігати у форматах без втрат, таких як TIFF, GIF або PNG.

JPEG не повинен застосовуватися тоді, коли неприпустимі навіть мінімальні втрати, наприклад, при стискуванні медичних зображень. У таких випадках можна використовувати передбачений стандартом JPEG варіант стиску JPEG-Lossless. Він не використовує ДКП, що забезпечує точний збіг відновленого й вихідного зображень, але його мала ефективність (коефіцієнт стискування нечасто буває більше двох) і відсутність підтримки з боку розроблювачів програмного забезпечення не сприяли популярності JPEG-Lossless.

В алгоритмі JPEG 2000 замість ДКП (застосовуваного в алгоритмі JPEG) використовується вейвлет-перетворення, яке засноване на поданні сигналу у вигляді суперпозиції базових функцій зі спеціальними властивостями, названих вейвлетами. Стиснуте цим алгоритмом, а потім відновлене зображення виходить більше гладким і чітким, а розмір файлу в порівнянні з JPEG при однаковій якості відновленого зображення виявляється меншим. JPEG 2000 повністю вільний від головного недоліку JPEG: завдяки використанню вейвлетів, зображення, відновлені після сильного стискання, не містять артефактів у вигляді блоків пікселів. Алгоритм JPEG 2000, як й JPEG, підтримує так назване прогресивне стискання, що дозволяє в міру завантаження побачити спочатку розмите, а потім усе більш чітке зображення. Артефакти, які виникають при стисканні за алгоритмом JPEG 2000 з високим ступенем компресії, якісно мало відрізняються від артефактів, які виникають при стисканні алгоритмом JPEG. У тих місцях, де оригінальне зображення мало плавні колірні переходи, вони стають ще більш плавними (розмитими), а де були різкі переходи яскравості або кольорів (ділянки зображення з високою контрастністю), виникає характерний артефакт у вигляді яскравого контуру, що обрамляє границю переходу, незначно (на парі пікселів) відступаючого від більше темної ділянки переходу. Розходження в артефактах: немає сітки в 8×8 пікселів; не спотворюються кольори дрібних деталей, що сильно відрізняються за кольорами від тла; артефакти, характерні для JPEG 2000, стають помітними при більших, ніж у випадку JPEG, ступенях стискання.

Алгоритм DjVu (від франц. *deja vu* – вже бачене) реалізує стискання зображень із втратами. Він розроблений спеціально для стискання відсканованих документів (книг, журналів, рукописів і т.п.), де надмірність формул, схем, малюнків і рукописних символів робить трудомістким їхнє повноцінне розпізнавання. Також даний алгоритм є ефективним, якщо необхідно передати: всі нюанси оформлення, наприклад історичних документів, де важливе значення мають зміст, кольори й фактура паперу; дефекти пергаменту (тріщинки, сліди від складання); виправлення, ляпки, відбитки пальців, сліди, залишені іншими предметами й т.д.

Для стиску кольорових зображень в DjVu застосовується технологія, яка розділяє вихідне зображення на три шари: передній план, задній план і чорно-білу (однобітову) маску. Маска зберігається з відрізненням вихідного файлу й саме вона містить зображення тексту та інші чіткі деталі. Відрізнення заднього плану, у якому залишаються ілюстрації й текстура сторінки, за замовчуванням знижується для економії місця. Передній план містить колірну інформацію про маску. Його відрізнення звичайно знижується ще сильніше. Потім задній і передній плани стискаються за допомогою вейвлет-перетворення, а маска – алгоритмом JB2.

Особливістю алгоритму JB2 є те, що він шукає на сторінці повторювані символи й зберігає їхнє зображення тільки один раз. У багатосторінкових документах кожні декілька сторінок, які ідуть підряд, користуються загальним «словником» зображень символів.

Для стискання більшості книг можна обійтися тільки двома кольорами. У цьому випадку використовується всього один шар, що дозволяє досягти рекордного ступеня стискання. У типовій книзі із чорно-білими ілюстраціями, відсканованою з відрізненням 600 dpi, середній розмір сторінки становить близько 15 КБ, тобто приблизно в 100 разів менше, ніж об'єм вихідного зображення. При наявності складного заднього плану виграш об'єму становить звичайно 4...10 разів. Однак при стандартних налаштуваннях в DjVu використовується стискання даних із втратами, тому для особливо важливих документів застосовуються алгоритми стискання без втрат: PNG, JPEG 2000, TIFF і т.п. В DjVu можна використати й стискання даних без втрат за допомогою утиліти CJB2.

В основі алгоритму DjVu лежать кілька базових алгоритмів:

- алгоритм відділення тексту від тла на відсканованому зображенні;
- алгоритм вейвлет-стискання тла IW44;
- алгоритм стиску чорно-білих зображень JB2;
- універсальний алгоритм стискання ZIP;
- алгоритм розпакування «за запитом»;
- алгоритм «маскування» зображень.

Фрактальне стискання зображень. Деякі недоліки відновлення зображення після стискання алгоритмом JPEG пов'язані з тим, що при стисканні не враховується специфіка зображених об'єктів, тобто не виявляються їхня структура й характерні ділянки. Врахування структури зображеного об'єкта лежить в основі фрактального методу стискання зображень, який запропонував в 1988 р. Б. Мандельброт. Фрактал – це структура, виділена при аналізі зображення й яка має схожу форму незалежно від її розмірів. Наприклад, у зображенні крони дерева, фрактал може представляти листя. Тому зображення можна формувати, збираючи об'єкти, називані фракталами. У свою чергу, окремий фрактал може бути описаний деяким стандартним образом. Основу фрактального підходу становить постулат про те, що існує набір афінних коефіцієнтів, які описують обертання, стискання, розширення, перекручування форми й зсув об'єктів, які складають зображення.

2.5 Формати подання цифрових зображень і відео

Зображення потрібно ефективно зберігати і експортувати. Для цього створені різні описи для зберігання зображень у комп'ютерних файлах – формати. Існує багато різних форматів графічних файлів, тобто файлів, що містять зображення. Деякі з них стандартизовані.

Формат – це специфікація структури даних, записаних у комп'ютерному файлі. Тип формату часто вказується в розширенні імені файлу, відділеному крапкою. Наприклад, закінчення імені «.tif» використовують для позначення файлів, що містять зображення, описане у форматі TIFF.

Для багатьох форматів існують опубліковані специфікації, де докладно описана структура файлів даного формату й те, як програми повинні кодувати дані для запису в цей формат і декодувати їх при читанні. Найпоширеніші формати і їхні особливості представлені в таблиці 2.2.

Таблиця 2.2 – Формати графічних файлів, їхній опис і область застосування

Формат	Особливості	Застосування
RAW	Матриця цифрових даних, отриманих із ПЗС-матриці фотокамери	Стискання без втрати якості
BMP	Зберігає дані із глибиною кольорів до 48 біт і максимальним розміром 65 535×65 535 пікселів без стискання	Підтримується ОС Windows
GIF	Використовує 256 відтінків кольорів, ефективно стискає однотонні кольорові області, зберігаючи деталі зображення без втрат. Може містити прозорі області	Текст, логотипи, векторні ілюстрації, анімовані малюнки
TIF	Зберігає фотографії в колірних просторах (RGB, CMYK, YCbCr, CIE Lab й ін.) із глибиною кольорів 8, 16, 32 й 64 біт	Поліграфія; GeoTIFF використовується для зберігання в одному файлі декількох спектральних супутникових зображень
JPEG	Підтримує до 16 млн кольорів. Стискає дані із втратами, може внести перекручування в малюнок, особливо який містить текст або тонкі лінії	Фотографії. Не годиться для малюнків, які містять прозорі ділянки, дрібні деталі або текст
PNG-24	Підтримує 16 млн кольорів. Подібно JPEG зберігає яскравість і відтінки кольорів у фотографіях, подібно GIF й PNG-8 зберігає деталі зображення (наприклад, у векторних малюнках і логотипах)	Фотографії; малюнки, що містять прозорі ділянки; малюнки з більшою кількістю кольорів і чітких країв зображень

Формат TIFF (Tagged Image File Format) споконвічно був створений як універсальний формат для зберігання сканованих зображень із колірними каналами. Його головна перевага – переносимість на різні платформи. На сьогоднішній день TIFF є одним з найпоширеніших і надійних, його підтримують практично всі програми, так чи інакше пов'язані із графікою. У цьому форматі доступний весь діапазон колірних моделей – від монохромної до RGB, CMYK і додаткових кольорів Pantone. Крім того, TIFF дозволяє, якщо буде потреба, зберігати різноманітну

додаткову інформацію про зображення (контури, альфа-канали й т.п.). У форматі TIFF може бути використаний стиск за алгоритмом LZW.

TIFF – це кращий алгоритм при зберіганні зображень із високою якістю, призначених для подальшого аналізу й обробки, імпорту растрової графіки у векторні програми й видавничі системи.

Незалежний від апаратного забезпечення формат GIF (Graphics Interchange Format) був розроблений в 1987 р. (модифікація GIF87a) фірмою CompuServe для передачі растрових зображень по мережах. В 1989 р. специфікація формату була дороблена (модифікація GIF89a), додані підтримка прозорості й анімації зображень. Формат GIF використовує LZW-стискання і має обмежену палітру в 256 кольорів. Його можна назвати форматом зберігання зображень без втрат. Однак це твердження справедливо тільки в тому випадку, якщо вихідне зображення має від 2 до 256 кольорів. Якщо ж кількість кольорів більше 256, то частина з них буде змінена. GIF використовується як один з основних растрових форматів для подання графічних даних в Інтернеті.

Формат JPEG описує структуру даних у файлі після стискання зображень однойменним алгоритмом JPEG і призначений для зберігання напівтонових і кольорових фотографій з високим коефіцієнтом стискання без помітної для ока людини втрати якості. При цьому, як правило, зображення не може бути повністю відновлене у вихідному виді. При збереженні зображення у форматі JPEG можна управляти параметром якості, вимірюваним у відносних одиницях, найчастіше від 0 (максимальний стиск) до 100 (максимальна якість). Формат JPEG варто використовувати для зберігання повнокольорових зображень (у першу чергу фотографій), не призначених для подальшої обробки. Не рекомендується його застосовувати для зберігання синтезованих зображень із невеликою кількістю кольорів (менш 256) і з різкими переходами між кольорами, тому що це може привести до прояву помітних оком дефектів.

Формат BMP (Windows Device Independent Bitmap, від BitMa – бітова карта) є растровим форматом ОС Windows. Він дозволяє описувати структуру зображення як в індексованому виді (з палітрою до 256 кольорів), так і повнокольорові RGB-зображення (із глибиною кольорів 24 біта). Можливе застосування алгоритму стискання RLE. Використання даного формату обмежене тим, що він не підтримується графічними програмами під управлінням більшості операційних систем, крім Windows.

Формат DjVu оптимізований для зберігання й передачі документів по мережі таким чином, що сторінку можна переглядати ще до завершення завантаження файлу. DjVu-файл може містити текстовий шар (це дозволяє здійснювати повнотекстовий пошук по файлу), а також вбудований інтерактивний зміст й активні області – посилання (це дозволяє реалізувати зручну навігацію в DjVu-книгах).

Цифрове відео базується на технологіях запису, обробки, передачі, зберігання, відтворення візуальних й аудіовізуальних даних у цифровому поданні. Не варто плутати поняття «цифрове відео» із цифровим телебаченням. Цифрове телебачення визначає стандарти передачі відео- і аудіосигналу від передавача до телеприймача, використовуючи при цьому цифрову модуляцію, тобто припускає передачу цифрового відео на відстань за допомогою супутникових, наземних, мобільних або кабельних мереж.

Відеосигнал формується з послідовності кадрів – окремих зображень. Системи реєстрації відеосигналу бувають одноматричні й трьохматричні. Для одержання кольорового відеозображення застосовують спеціальну RGB-фільтрацію елементів зображення, щоб на кожен елемент доводилося по три значення кольорів – червоного, зеленого й синього. Такий метод застосовується в одноматричних відеокамерах і фотоапаратах з підтримкою відеозапису.

Для професійної відеозйомки застосовуються трьохматричні системи, де оптичний кольороподіл здійснюється за допомогою дихроїдної призми. Кожне з кольороподілених зображень попадає на окрему матрицю, з якої зчитуються елементи кадру, потім формується відеосигнал окремого компонента. Після гамма-корекції сигналів R, G, B проводиться їхнє перетворення в модель YCbCr. Отримані компоненти Y, Cr, Cb квантуються з розрядністю 8 або 10 біт. Але для передачі компонента Y при 8-бітному кодуванні з 256 доступних рівнів використовуються тільки 220 (від 16-го до 235-го), а інші – для сигналів синхронізації. При 10-бітному кодуванні використовується 877 рівнів; для кольорних компонентів Cr, Cb – 225 рівнів в 8-бітній системі й 897 рівнів в 10-бітній.

Для цифрових відеостандартів прийняте базове співвідношення 4:2:2, що означає, що кольорорізницеві компоненти Cr, Cb передаються із просторовим відрізненням у два рази меншим, чим по яскравісному сигналу Y, тому що людське око більш чутливе до зміни яскравості, чим кольору. При цьому частота дискретизації для яскравісного сигналу Y встановлюється рівної 13,5 МГц, що у два рази більше, ніж для кольорорізницевих сигналів Cr й Cb – 6,75 МГц. Для сигналів телебачення високої чіткості встановлені частоти дискретизації сигналів яскравості (74,25 МГц) і кольоровості (37,125 МГц). Для європейського відеостандарту PAL розмір кадру становить 720×576 пікселів, для північноамериканського стандарту NTSC – 720×480, для відео DVD 720×480, для відео високої чіткості (HD 720 p) – 1280×720, а для стандарту HDTV (Full HD) – 1920×1080, для 4K 3840×2160 точок. Мінімальний показник, при якому рух буде сприйматися плавним, – 16 кадрів у секунду. У кіно стандартна частота зйомки й проєкції становить 24 кадра в секунду. Системи телебачення PAL і SECAM використовують 25 кадрів в секунду, а система NTSC – 30 кадрів в секунду.

Стандарт розкладання визначає параметри телевізійної розгортки, застосовуваної для перетворення двовимірного зображення в одновимірний потік даних. Розгортка може бути прогресивною (порядковою) або черезрядковою. При прогресивній розгортці всі горизонтальні рядки зображення відображаються по черзі один за одним, при черезрядковій – кожен кадр розбивається на два напівкадри, кожний з яких містить парні або непарні рядки.

Для розміщення 60-секундного ролика із частотою 30 кадрів у секунду, відрізненням 720×576 пікселів й 16-бітною глибиною кольорів буде потрібно біля півтора гігабайт без урахування звукової доріжки. Існує кілька десятків алгоритмів стискання, які використовують різні варіанти компресії. Для кодування відеосигналу використовують формати відео MPEG-2, MPEG-4, H.264, DV, AVC й ін. При цьому, застосовуючи певний алгоритм перетворення даних, можна стиснути відео різними програмними засобами, що дає на виході різні результати. Ці відмінності визначаються кодеком – програмою, що здійснює алгоритм стискання вихідних сигналів. Кожен кодек використовує власний алгоритм, який впливає на якість і швидкість кодування. Фільм тривалістю 1,5 години, записаний у форматі DV, займає 20 Гб, в MPEG-2 – 4,7 Гб, в MPEG-4 – 700 Мб.

Контрольні запитання

1. Що таке дискретизація зображення ?
2. Що таке квантування зображення ?
3. Що таке кодування зображення ?
4. Які існують способи формування зображення ?
5. Які існують типи растрових зображення ?
6. Які існують методи стискання зображення ?
7. Які алгоритми і методи стискання зображень відносять до алгоритмів без втрат ?
8. Які алгоритми і методи стискання зображень відносять до алгоритмів з втратами ?
9. Які існують формати подання зображення ?
10. Що таке формат файлу зображення ?

3 ОДЕРЖАННЯ ЦИФРОВИХ ЗОБРАЖЕНЬ

Цифрові зображення можуть бути отримані за допомогою реєстрації не тільки звичних для ока картин у видимому діапазоні електромагнітного спектра, але й за допомогою реєстрації невидимих сигналів. Розглянемо основні типи пристроїв реєстрації та способи перетворення сигналів у

цифрові зображення.

3.1 Елементарний сенсор

Принцип створення зображення однаковий й у цифровій камері, і в плівковій: обидві забезпечують фіксацію об'єкта зйомки, використовуючи енергію світла, який впливає на світлочутливий матеріал. Елементарна частка електромагнітного випромінювання називається фотоном. Розглянемо, як фотон стає електроном.

Пристрої реєстрації зображень в оптичному діапазоні спектра складаються з безлічі світлочутливих елементів, названих пікселями. Піксель складається з р-підкладки, покритої прозорим діелектриком, на який нанесений електрод, який пропускає світло. Світло формує потенційну яму (рис. 3.1).

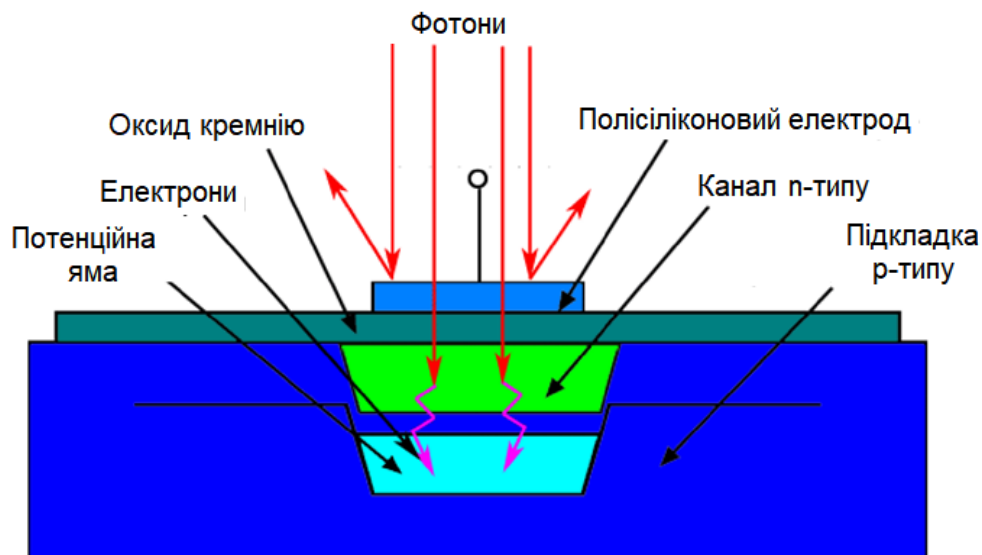


Рисунок 3.1 – Перетин пікселю ПЗЗ-матриці

Для фотона, який потрапив на піксель, можливі три варіанти розвитку подій: він зрикошетить від поверхні, буде поглинений товщею напівпровідника, проб'є наскрізь робочу зону напівпровідника. Розроблювачі прагнуть створювати сенсори, у яких втрати від рикошету й пробивання були б мінімальні.

У приладах із зарядовим зв'язком перетворення фотона в електрон відбувається в результаті поглинання світлового кванта кристалічною решіткою напівпровідника з виділенням електронів. Це може бути або пари «електрон+дірка», або одиничний електрон – останнє відбувається при використанні домішок у напівпровіднику. Носії, які утворилися, заряду до моменту зчитування необхідно зберегти. Для цього кремнієва підкладка р-типу оснащується каналами з напівпровідника n-типу, над

якими з полікристалічного кремнію виготовляються прозорі для фотонів електроди. Після подачі на такий електрод електричного потенціалу в збідненій зоні під каналом n-типу створюється потенційна яма, призначення якої – зберігати заряд, що виникає в результаті внутрішнього фотоефекта. Чим більше фотонів упаде на світлочутливий елемент (піксель) і перетвориться в електрони, тим вище буде заряд, накопичений ямою. Даний заряд одержав назву фотострум, його значення досить мало й після зчитування вимагає обов'язкового посилення.

Не вся поверхня пікселю є світлочутливою, тому над ним часто встановлюється мікролінза для збільшення кількості накопичувальних фотонів (рис. 3.2). Розмір одного пікселя дорівнює 3...8 мкм.

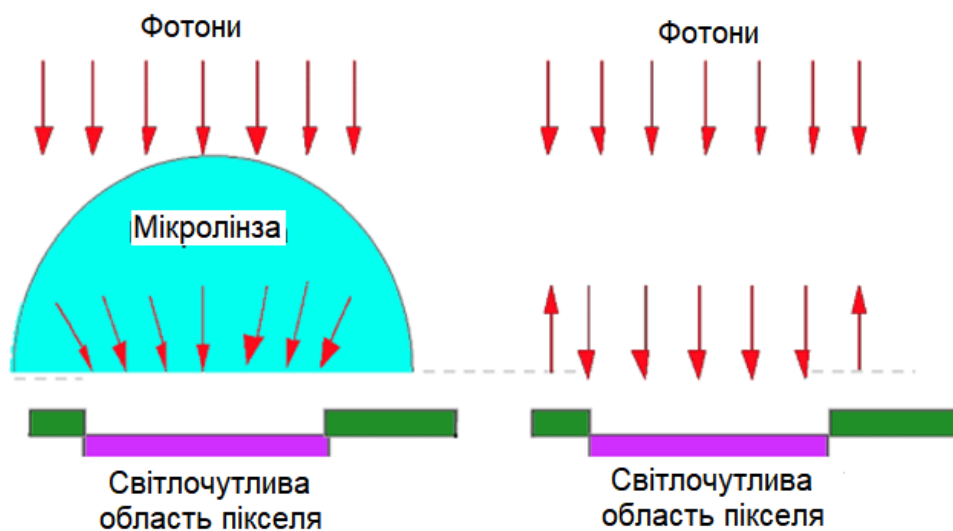


Рисунок 3.2 – Мікролінза підвищує світлочутливість пікселя

Зчитування фотострумів світлочутливих елементів здійснюється так називаними послідовними регістрами зсуву, які перетворюють рядок зарядів на вході в серію імпульсів на виході. Дана серія являє собою аналоговий сигнал.

У техніці використовуються пристрої різного типу, у яких реалізовані два способи збору фотострумів з індивідуальних сенсорів:

ПЗЗ – прилади із зарядовим зв'язком;

КМОН – прилади на базі симетричних комплементарних металооксидних напівпровідників.

Перевагами ПЗЗ-приладів є:

- низький рівень шумів;
- високий коефіцієнт заповнення пікселів (до 100 %);
- висока ефективність (відношення числа зареєстрованих фотонів до числа, що потрапили на світлочутливу область сенсора) – 95%, а людське око має квантову ефективність порядку 1%;
- високий динамічний діапазон (відношення струму насичення до

- середнього квадратичного шумового струму, дб);
- гарна чутливість в інфрачервоному діапазоні спектра.

До недоліків ПЗЗ-приладів ставляться:

- складна система зчитування сигналу;
- високий рівень енергоспоживання (до 2...5 Вт);
- дорожче КМОН у виробництві.

Переваги КМОН-приладів:

- висока швидкодія (реєстрація до 500 кадрів у секунду);
- низьке енергоспоживання (в 100 разів у порівнянні із ПЗЗ);
- дешевше й простіше ПЗЗ у виробництві;
- перспективність технології (на одному кристалі можна реалізувати всі необхідні додаткові схеми: аналого-цифрові перетворювачі, процесор, пам'ять, тобто фактично закінчену цифрову камеру на одному кристалі).

Недоліки КМОН-приладів:

- низький коефіцієнт заповнення пікселів, що знижує чутливість (ефективна поверхня пікселя до 75%, інше займають транзистори);
- високий рівень шуму (навіть при відсутності освітлення через фотодіод тече значний струм), боротьба з яким ускладнює технологію;
- невисокий динамічний діапазон.

У ПЗЗ-сенсорах сигнал зчитується послідовно з кожного осередку, рядок за рядком. Тому зробити наступний знімок можна лише тоді, коли попередній уже повністю сформований. КМОН улаштовані інакше: сигнал з кожного осередку зчитується індивідуально. Для кожного осередку задані координати в матриці й можна одержати індивідуальний доступ до окремого осередку.

Пікселі можна поєднувати в рядки (лінійки) або матриці. Кожен елемент уловлює різну кількість фотонів, формує аналоговий сигнал, який можна перетворити в масив, або матрицю цифрових значень. У такий спосіб формують напівтонові цифрові зображення.

Для одержання кольорових зображень використовують сенсори зі світлофільтрами. Найпоширенішим варіантом є використання мозаїки кольірних фільтрів (мозаїки Байера), де світлочутливий елемент накритий фільтром, який пропускає фотони, що відповідають діапазонам спектра червоного, зеленого й синього кольорів. Найбільш відомим (але не єдиним варіантом) є мозаїка Байера (рис. 3.3). Із чотирьох сусідніх сенсорів два реєструють зелені кольори, інші два – червоні й сині (блоки RGGB).

У результаті матриця, що описує зображення, складається з мозаїки пікселів трьох основних кольорів різної інтенсивності. Зелених кольорів реєструється більше, тому що це відповідає особливостям людського зору. Для відновлення відсутніх кольорів у кожному пікселі застосовуються

спеціальні алгоритми інтерполяції мозаїчного зображення (рис 3.4).

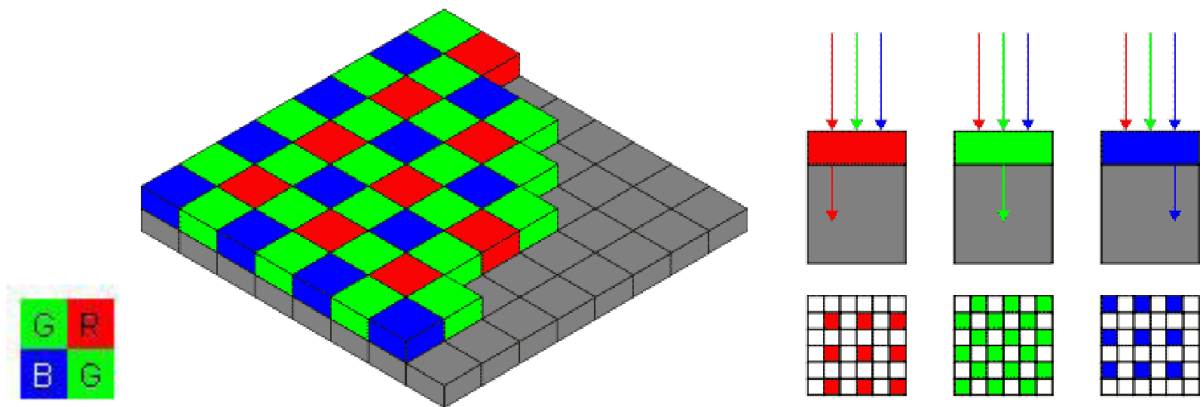


Рисунок 3.3 – Мозаїка фільтрів Байера

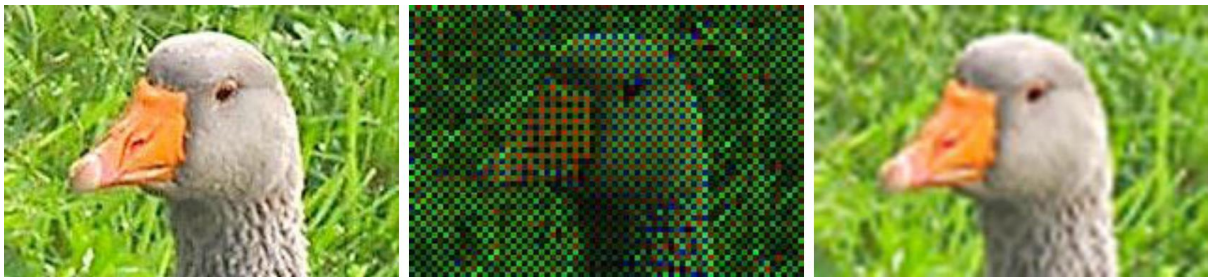


Рисунок 3.4 – Реальне зображення (ліворуч), мозаїка з фільтрів Байера (у центрі), інтерпольоване (праворуч)

У КМОН-матриці світлофільтри не використовуються в принципі. Поділ світлового потоку на складові кольори відбувається за рахунок того, що світло з різною довжиною хвилі проникає в шар кремнію на різну глибину. Наприклад, хвилі синіх кольорів мають найменшу проникаючу здатність, а червоного – найбільшу. Фотоелемент кожного пікселя складається із трьох шарів (рис. 3.5). Товщина цих шарів розрахована так, що вони пропускають промені з довжинами хвиль певного діапазону. При цьому до нижнього шару доходять фотони з довжиною хвилі, що відповідає червоним кольорам, до середнього – зеленим, до верхнього – синім. Електрони, які утворюються при цьому, і «дірки» накопичуються в трьох потенційних ямах – по одній на кожен шар. Таким чином, для кожного з пікселів матриці є дані по кожному з колірних діапазонів. Це дозволяє відмовитися від мозаїчних схем і пов'язаних з ними втрат даних про кольори. Однак при розподілі фотонів по шарах частина з них неминуче поглинається при переході з одного шару в іншій. У результаті чутливість матриці послаблюється.

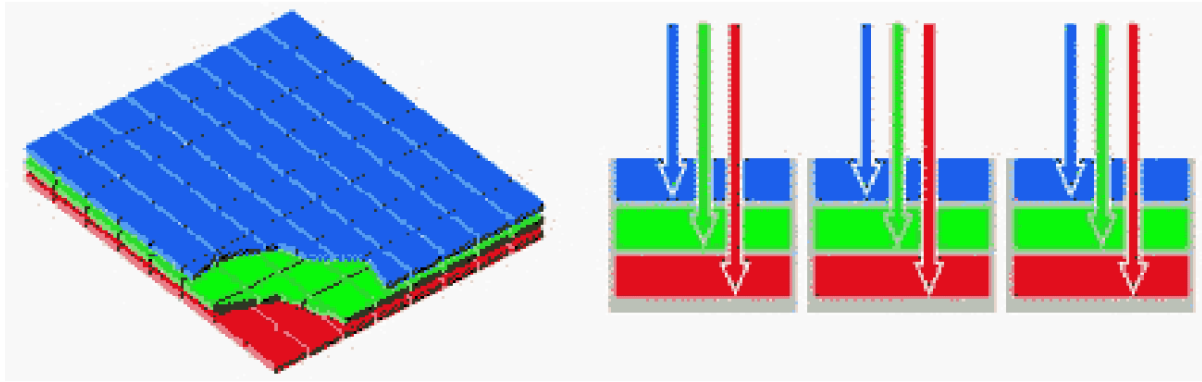


Рисунок 3.5 – Багатошарові сенсори

Відзначимо, що форма сенсора, називаного пікселем, може бути не тільки прямокутною (рис. 3.6), але й L-образною, і шестикутною (розробка Super CCD компанії Fujifilm, 1999 р.).

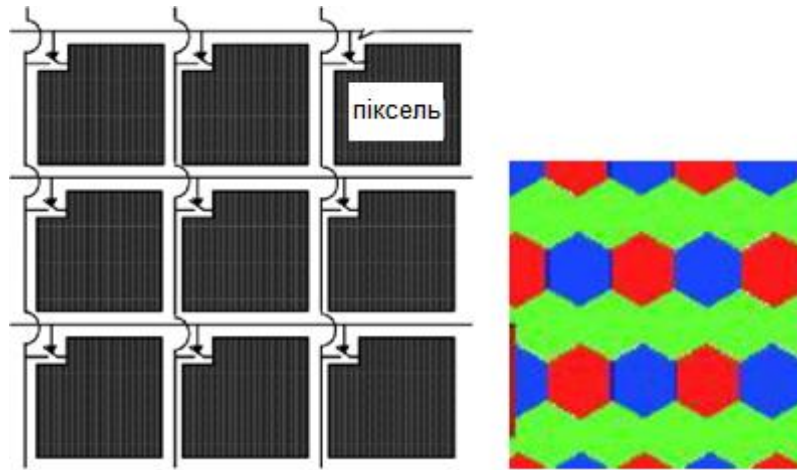


Рисунок 3.6 – Пікселі-сенсори не прямокутної форми

3.2 Сканер

Сканер – це пристрій, який за допомогою аналого-цифрового перетворювача створює цифрове зображення деякого документа. Для цього використовуються штучне освітлення й відбиття у видимому діапазоні електромагнітного спектра. Процес одержання цього зображення називається скануванням. Сканери бувають ручні, рулонні, планшетні й проєкційні. Різновидом проєкційних сканерів виступають слайди-сканери, призначені для сканування фотоплівки. У високоякісній поліграфії використовуються барабанні сканери, як світлочутливий елемент у них застосовується фотоелектронний помножувач.

Головною особливістю сканера є формування зображення ПЗЗ-сенсорами, вбудованими в лійку. В основному сканери відрізняються

способом пересування ПЗЗ-лінійки або сканованого документа один відносно одного. Найпоширенішими є планшетні сканери.

Принцип роботи однопрохідного планшетного сканера (рис. 3.7) полягає в тому, що уздовж сканованого документа, розташованого на прозорому нерухомому склі, рухається сканувальна каретка із джерелом світла. Відбите світло через оптичну систему сканера (що складається з об'єктива й дзеркал або призми) попадає на ПЗЗ-сенсори, розташовані лінійно.

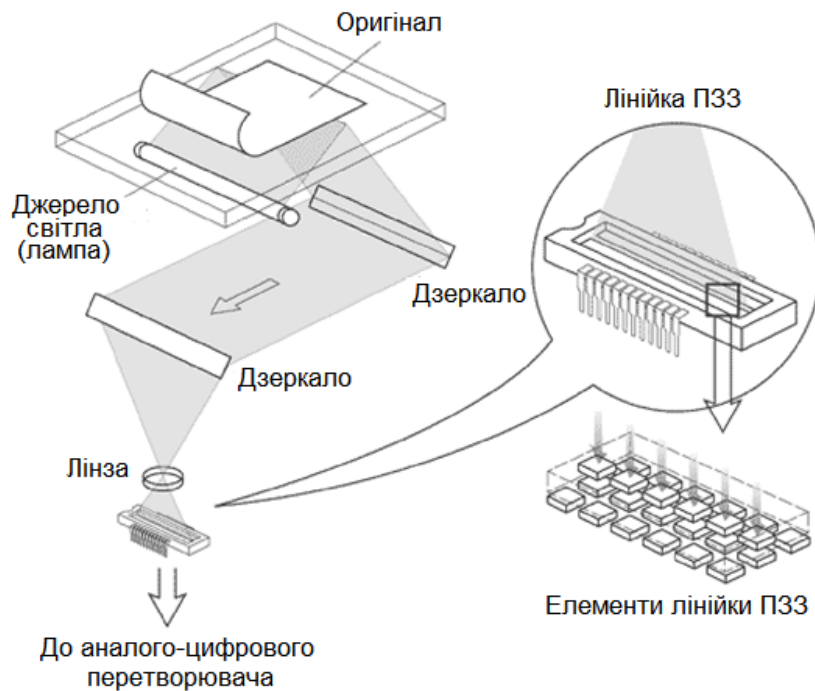


Рисунок 3.7 – Схема побудови планшетного сканера

При реєстрації використовується ПЗЗ-лінійка, яка створює один рядок зображення. Для одержання зображення у вигляді матриці каретка зміщується на фіксовану відстань і реєструє наступний фрагмент документа у вигляді рядка пікселів.

Розрізняють оптичне і механічне відрізнення сканера. Наприклад, якщо на ньому зазначене відрізнення 300×1200 dpi то його оптичне відрізнення дорівнює 300 dpi, а механічне (крок зсуву каретки) – 1200 dpi.

3.3 Цифрові фотоапарати

Цифровий фотоапарат за конструкцією схожий зі звичайним фотоапаратом, але замість фотоплівки, на якій фіксується зображення, у нього є спеціальні датчики (фотоматриці): матричні ПЗЗ або КМОН. Ці датчики перетворюють проєктоване на них об'єктивом зображення в цифрову форму.

Дві якісні характеристики, які безпосередньо впливають на результат зйомки, – це динамічний діапазон сенсора й розрядність подання кольорів. Перша із цих характеристик відбиває здатність матриці передавати світлові відтінки, друга ставиться не тільки до сенсора, але й до аналого-цифрового перетворювача, який переводить електричні сигнали з виходів матриці в цифровий код.

Динамічний діапазон визначає кількість відтінків сірого (рівнів яскравості), які здатний розрізнити сенсор цифрової камери між абсолютно чорним й абсолютно білим кольорами. Чим ширше динамічний діапазон, тим вище подібність реєстрованого зображення з оригіналом.

Слід зазначити, що дотепер самий широкий динамічний діапазон є у негативної фотоплівки. Тому, незважаючи на досягнення цифрових технологій, при демонстрації фільмів у кінотеатрах все ще використовуються плівкові проектори.

Існує кілька варіантів конструкції цифрових фотоапаратів: камери із заднім розгорненням, трьохкадрові камери й однокадрові камери з однієї або трьома фотоматрицями. У поточний час в основному застосовуються однокадрові фотокамери.

Камера із заднім розгорненням (рис. 3.8) робить сканування в площині зображення. Принцип роботи такої камери схожий на технологію, реалізовану в сканерах з порядковим зчитуванням інформації.

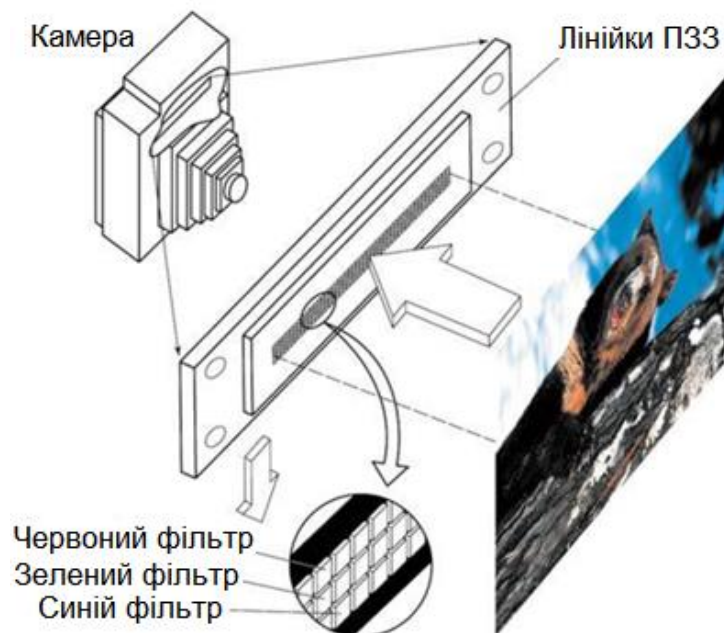


Рисунок 3.8 – Принцип дії камери із заднім розгорненням

Сканувальна голівка, яка містить лінійку світлочутливих ПЗЗ, переміщується з невеликим кроком уздовж задньої фокальної площини камери поперек зображення, реєструючи за кожен крок один рядок

пікселів. Камери, у яких застосовується такий принцип, дозволяють одержувати зображення з високим відрізненням, але час експозиції може досягати декількох хвилин, що робить техніку заднього розгорнення непридатною для зйомки рухомих об'єктів, або при роботі зі спалахом. Під час сканування затвор камери залишається відкритим, тому необхідно застосовувати постійне освітлення, тому що ні спалах, ні стробоскоп у цьому випадку непридатні.

Трьохкадрова камера (рис. 3.9) призначена для реєстрації кольорових зображень нерухомих об'єктів. Як світлочутливий датчик використовується двовимірна матриця ПЗЗ.

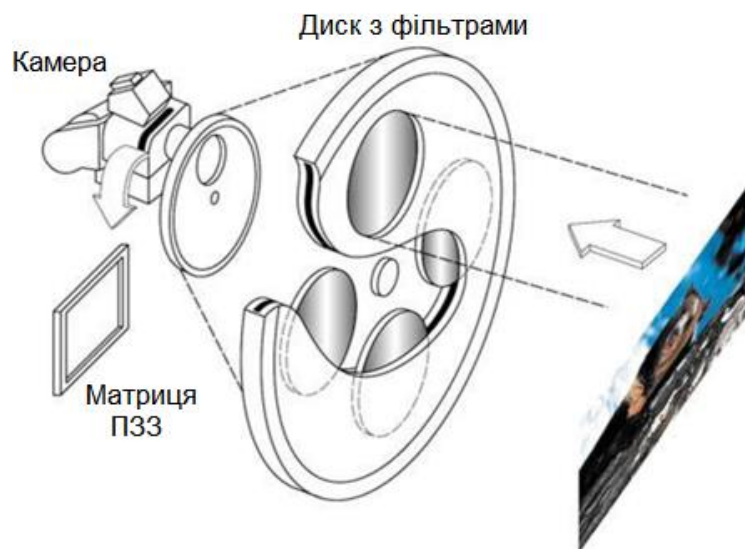


Рисунок 3.9 – Принцип дії трьохкадрової камери

Плоскі двовимірні матриці мають набагато менше відрізнення, чим лінійні. Кожен елемент матриці формує один елемент (піксель) зображення. Експозиція здійснюється з такою швидкістю, що можна користуватися освітленням від звичайного спалаху. Для реєстрації кольорового зображення потрібно зробити три окремих знімки через три світлофільтри (червоний, зелений і синій). Між експонуваннями диск із фільтрами повертається таким чином, щоб у момент зйомки перед матрицею перебували червоний, зелений або синій світлофільтри.

Технологія, реалізована в однокадрових камерах з однією матрицею, забезпечує високу швидкість цифрування зображення, але характеризується більше низьким відрізненням і гіршою передачею кольору, чим «багатознімальна» технологія. Такі камери називаються також камерами з обчисленням кольорів.

Як й у трьохкадровій, в однокадровій камері з однією матрицею (рис. 3.10) застосовується плоска матриця, але дані про кольори реєструються не через окремі фільтри, а через нанесений на поверхню матриці ПЗЗ

плівковий фільтр, що складається із червоних, зелених і синіх елементів. Дані про кожен елемент зображення реєструються тільки в одному із трьох кольорів, наприклад, у червоному. Для додавання до нього належних часток зеленого й синього програма обробки інтерполює дані про кольори сусідніх точок. Оскільки потрібно всього одна експозиція, однокадрові камери забезпечують зйомку рухомих об'єктів.

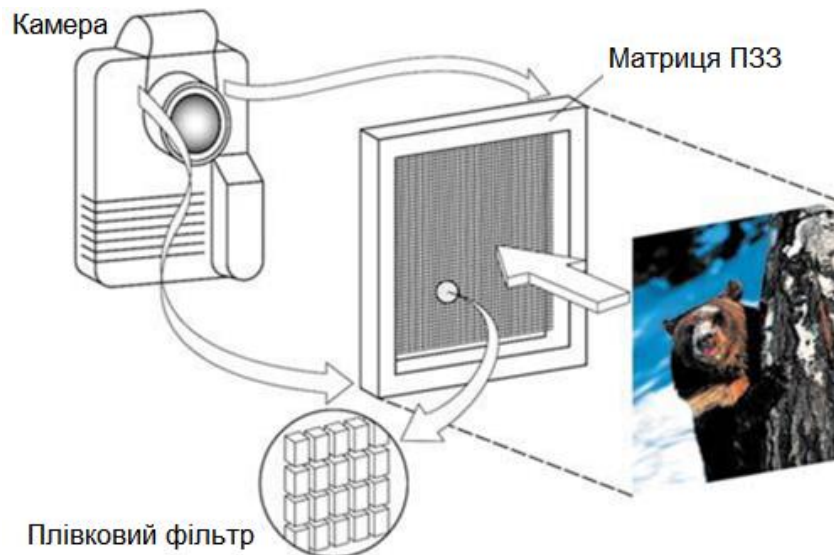


Рисунок 3.10 – Принцип дії однокадрової камери з однією матрицею

Принцип дії однокадрової камери із трьома матрицями (камери для однокадрової кольорової зйомки) полягає в оптичному поділі нахожого світла призмою на червону, зелену й синю складові, причому кожна з них направляється на свою матрицю (рис. 3.11). В одних моделях кожна матриця реєструє свій колір, в інших поєднуються плоска матриця, на поверхню якої нанесені червоні й синій плівкові фільтри, і дві додаткові матриці із зеленими фільтрами. У другому випадку інтерполяція проводиться тільки по двох кольорах, що приводить до підвищення якості зафіксованого зображення. Недолік цього способу реєстрації – відносно невелика розрізняльна здатність.

При використанні тришарових (ХЗ)-фотоматриць (рис. 3.12) компанії Foveon кольороподіл на основні кольори RGB проводиться в товщі напівпровідникового матеріалу пошарово із застосуванням фізичних властивостей кремнію, які полягають у тім, що зі збільшенням довжини світлових хвиль росте й глибина їхнього проникнення в кремній (рис. 3.12,а).

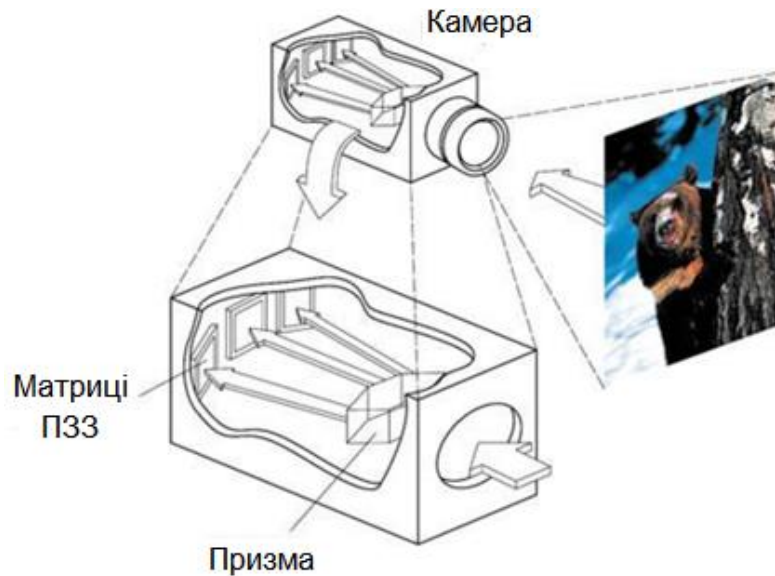


Рисунок 3.11 – Принцип дії однокадрової камери із трьома матрицями

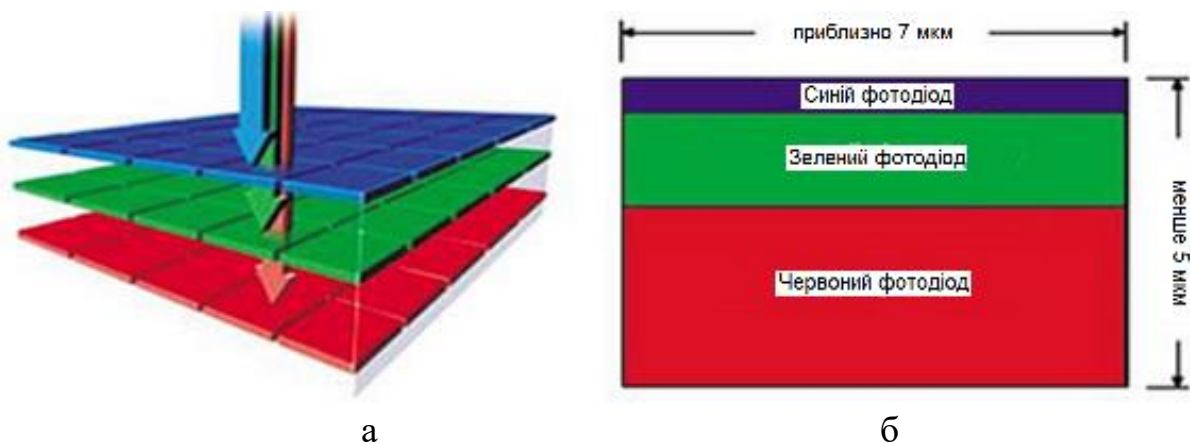


Рисунок 3.12 – Тришарова фотоматриця: а) принцип дії; б) структура шарів

Фотодіоди, створені чередуванням зон проникнення світла, розміщуються один під іншим на певних глибинах для вловлювання фотонів синього, зеленого і червоного кольору. Синя частина спектра поглинається верхнім шаром (товщина 0,2 мкм), зелена – середнім (0,4 мкм), а червона – нижнім (більше 2 мкм). Товщина кожного шару обрана за результатами експериментальних досліджень глибини проникнення квантів відповідного спектрального діапазону в кремній. Шари, у яких відбувається фотоефект, розділені додатковими тонкими зонами низьколегованого кремнію й мають окремі виходи сигналу. У такий спосіб виходить датчик, який реєструє інформацією про трьохколірні компоненти зображення в одній точці, що точно відповідає координатам формованого пікселя.

Перевагами ХЗ-фотоматриць є те, що вони не вимагають установки

кольорових фільтрів, а, отже, і підвищення співвідношення сигнал-шум, тому що фільтри поглинають приблизно дві третини світлового сигналу, й мають потенційно більш високу розрізняльну здатність.

Недоліками ХЗ-фотоматриць можна назвати невисоку точність передачі кольору, тому що найбільшою мірою вона визначається властивостями кремнію, а також відносно високий рівень цифрового шуму, тому що частина фотонів поглинається в «не своїй» області.

Для одержання якісної фотографії важливіше не кількість мегапікселів фотоапарата, а інші параметри: розміри активної області матриці, її чутливість, динамічний діапазон, точність передачі кольору. Найбільше значення мають перші три параметри, причому в ПЗЗ-матриці всі вони нерозривно зв'язані один з одним.

На рис. 3.13 показане співвідношення розмірів кадру плівки 24×36 мм і ПЗЗ-матриць.

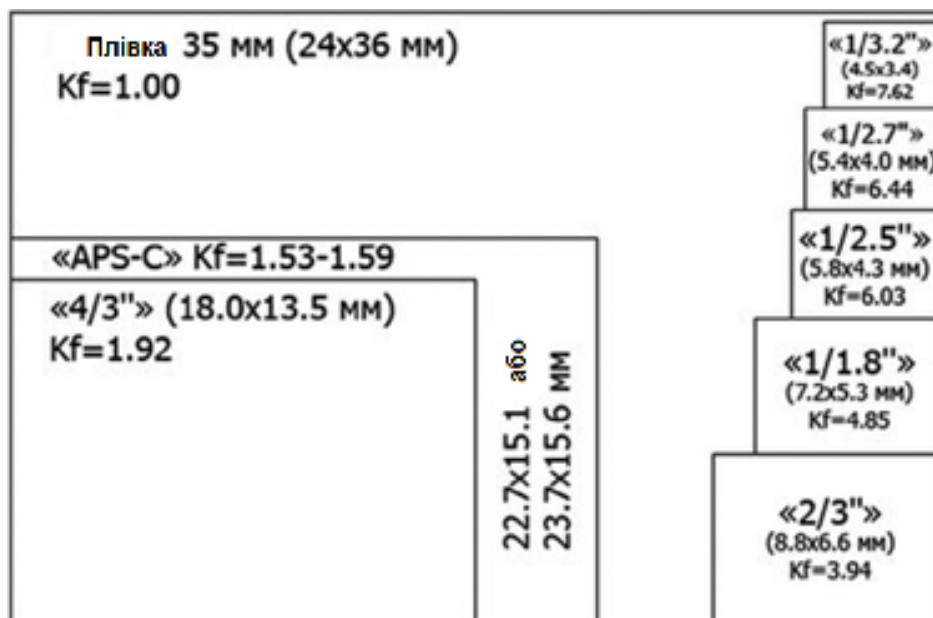


Рисунок 3.13 – Співвідношення розмірів кадру плівки 24×36 мм і ПЗЗ-матриці

Зазначено розміри матриць у дюймах і мм, а Kf – це коефіцієнт, який показує в скільки разів діагональ матриці менше діагоналі кадру плівки. Фізичний розмір матриці прямо пов'язаний з кількістю світла, що попадає на матрицю, тому чим більше матриця, тим якісніше будуть фотографії в умовах поганої освітленості.

Розмір матриці впливає на глибину різкості: менше матриця – більше глибина, але при цьому, наприклад, портретні фотографії виглядають плоскими. Збільшення розміру матриці спричиняє збільшення розміру й вартості фотоапарата.

Відзначимо, що поки цифрові фотоапарати не перевершують

плівкові. Вони не можуть знімати в умовах низької освітленості без застосування джерел штучного світла, а професійна плівка світлочутливістю в 3200 одиниць ISO витягне знімок навіть при світлі звичайної свічі.

3.4 Одержання зображень за допомогою ультразвуку

Ультразвукове дослідження (УЗД) – це неінвазійне дослідження живого організму за допомогою ультразвукових хвиль. Фізичною основою УЗД є п'єзоелектричний ефект.

При деформації монокристалів деяких хімічних сполук під впливом ультразвукових хвиль на поверхні цих кристалів виникають протилежні за знаком електричні заряди, тобто прямий п'єзоелектричний ефект. При подачі на них змінного електричного заряду в кристалах виникають механічні коливання з випромінюванням ультразвукових хвиль. Той самий п'єзоелемент може бути поперемінно приймачем і джерелом ультразвукових хвиль.

У системах медичної ультразвукової діагностики звичайно використовують частоти від 2 до 10 МГц. Роздільна здатність сучасних ультразвукових апаратів досягає 1...3 мм. Будь-яке середовище, у тому числі й тканини організму, перешкоджає поширенню ультразвуку, тобто має різний акустичний опір, величина якого залежить від їхньої щільності й швидкості поширення звукових хвиль. Досягши границі двох середовищ із різним акустичним опором, пучок ультразвукових хвиль перетерплює зміни: одна його частина продовжує поширюватися в новому середовищі, частково поглинаючись нею, інша – відбивається. Коефіцієнт відбиття залежить від різниці величин акустичного опору межуючих одна з одною тканин: чим більше це розходження, тим більше відбиття й амплітуда зареєстрованого сигналу, тим світліше і яскравіше він буде виглядати на екрані апарата.

Повним відбивачем є границя між тканинами й повітрям. Тому що на границі переходу ультразвуку з повітря на шкіру відбувається відбиття 99,99% коливань, то при ультразвуковому скануванні пацієнта необхідне змазування поверхні шкіри водним желе, яке виконує роль перехідного середовища. Відбиття залежить від кута падіння променя (найбільше при перпендикулярному напрямку) і частоти ультразвукових коливань (при більш високій частоті більша частина відбивається). Для дослідження органів черевної порожнини, порожнини малого таза використовується частота 2,5...3,5 МГц, для щитовидної залози – 7,5 МГц.

Генератор працює в імпульсному режимі, посилюючи близько 1000 імпульсів у секунду. У проміжках між генеруванням ультразвукових хвиль пьезодатчик фіксує відбиті сигнали. Відбиті сигнали надходять у підсилювач і спеціальні системи реконструкції, після чого з'являються на

екрані монітора у вигляді відбиття зрізу тіла, пофарбованого в різні відтінки чорно-білих кольорів, тобто напівтонового 2D-зображення (рис. 3.14). Оптимальним є наявність не менш 64 відтінків сірого.

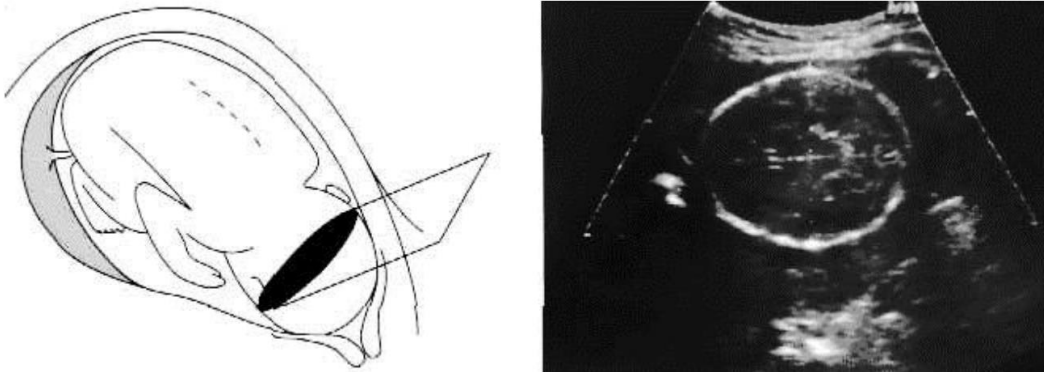


Рисунок 3.14 – Вимір розміру голівки плода: а) схема; б) зображення

У поточний час за допомогою УЗД можна одержати 3D-зображення (статичні тривимірні зображення) у результаті об'ємного сканування (рис. 3.15). УЗД 4D – це рухомі зображення у форматі 3D, які можна спостерігати в режимі реального часу. Як четвертий вимір виступає час.



Рисунок 3.15 – Зображення плода на 20-й тижню вагітності, отримане методом УЗД 3D

3.5 Рентгенівські зображення

Рентгенівські промені охоплюють діапазон електромагнітного спектра з довжиною хвиль від 1 до 0,01 нм. Око людини їх не бачить, але вони здатні проникати через тверді предмети.

Ідея одержання рентгенівського зображення заснована на різному ослабленні випромінювання при проходженні через тканини з різною щільністю й наступною реєстрацією. У результаті виходить зображення всіх тканин, спроектоване на площину (рис. 3.16).



Рисунок 3.16 – Приклад рентгенівського цифрового знімка змії, що проковтнула мишу

У сучасних цифрових апаратах реєстрація вихідного випромінювання проводиться на спеціальну електронну матрицю. У ній накопичується заряд, створюваний при поглинанні рентгенівських променів і його порядковій передачі на підсилювачі заряду під час сканування. Перетворення рентгенівського випромінювання в заряд в аморфному кремнії з метою електронного зчитування може бути реалізовано трьома методами: прямим, методом фотопровідності й сцинтиляційним. Кожен метод має свої переваги й недоліки. У всіх методах заряд перед зчитуванням накопичується протягом періоду кадру.

У прямому методі рентгенівські промені попадають на діод із кремнію, в якому створюються пари «дірка-електрон». Пристрій, який використовує прямий метод перетворення, виготовляється із кристалічного кремнію, але на практиці застосовуються тільки решітки з однієї або двох ліній, при цьому навіть такі решітки відрізняються дорожнечою.

У методі фотопровідності використовуються матеріали з більш високою, чим у кремнію, здатністю поглинати рентгенівське випромінювання. Вони наносяться на безліч провідних пластин для збору заряду, що поставляють разом з накопичувальним конденсатором, а також створюють пари «дірка-електрон» при поглинанні рентгенівських променів. Однак генерований заряд зберігається за межами шару щоб уникнути перехресних перекручувань. Прикладене поле відокремлює заряд і направляє його на розташовану безпосередньо внизу пластину для збору заряду з метою збереження різкості зображення. У цей час селен є єдиним використовуваним у виробництві фотопровідним матеріалом. Він

має відносно низьку здатність поглинати рентгенівське випромінювання й вимагає приблизно 50 електрон-вольт для створення пари «дірка-електрон». Це обмежує як мінімально необхідну дозу, так і величину генерованого сигналу.

У сцинтиляційному методі використовується сполука, яка поглинає рентгенівські промені й перетворює їхню енергію у видиме світло. Ефективний сцинтилятор робить багато фотонів світла для кожного вступного рентгенівського фотона. Сцинтилятори можуть являти собою гранули (наприклад, люмінофори) або кристали (наприклад, йодид цезію).

Практично всі сучасні системи рентгенівської візуалізації використовують моделі з детекторами ПЗЗ. Якість отриманого рентгенівського зображення визначається трьома основними параметрами: напругою, яка подається на рентгенівську трубку, силою струму й часом роботи трубки. Залежно від досліджуваних анатомічних утворень, маси й розмірів об'єкта дослідження ці параметри можуть істотно змінюватися.

Звичайно роздільна здатність досягає 0,5 мм (одна пара ліній, що розрізняють, на міліметр відповідає двом пікселям на міліметр), але можна одержати й більш високе відрізнення – до 26 пар ліній на міліметр, що відповідає 0,02 мм. Кращий на сьогоднішній день ПЗЗ-детектор рентгенівського випромінювання має до 16 мегапікселів і дозволяє одержувати цифрове зображення області 43×43 см.

Недоліки рентгенографії:

- інформативність класичної рентгенографії значно нижче таких методів візуалізації, як комп'ютерна томографія й магнітно-резонансна томографія. Звичайні рентгенівські зображення відбивають проекційне нашарування складних анатомічних структур, тобто їх сумарну рентгенівську тінь на відміну від пошарових серій зображень, одержуваних томографічними методами;
- без застосування речовин, що контрастують, рентгенографія практично неінформативна для аналізу змін у м'яких тканинах.

3.6 Томографічне зображення

Томографія – це метод одержання пошарового зображення внутрішньої структури деякого об'єкта. Томографію застосовують при виробництві деталей відповідального застосування (для авіації, космосу, автомобілебудування), в електроніці й мікроелектроніці, у матеріалознавстві, але в основному в медицині. Томографічне зображення одержують за допомогою пристроїв двох типів: на базі рентгенівського випромінювання і ядерного магнітного резонансу.

Магнітно-резонансна томографія

Магнітно-резонансна томографія (МРТ) – це томографічний метод

дослідження внутрішніх органів і тканин з використанням фізичного явища ядерного магнітного резонансу (ЯМР). Він заснований на вимірі електромагнітного відгуку ядер атомів водню на збудження їх певною комбінацією електромагнітних хвиль у постійному магнітному полі високої напруженості (рис. 3.17).

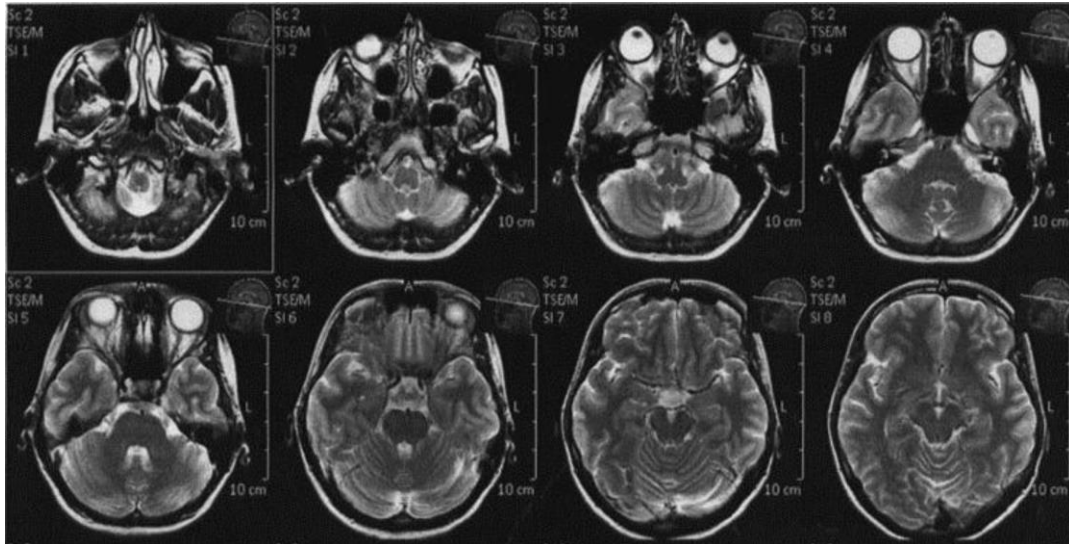


Рисунок 3.17 – МРТ-зображення голови людини

Спосіб і пристрій для ЯМР-томографії цього типу в 1960 р. запропонував В.А. Іванов (термін «ЯМР-томографія» був замінений на МРТ в 1986 р.).

МРТ дозволяє візуалізувати із високою якістю головний, спинний мозок й інші внутрішні органи. Ефект ЯМР виникає при взаємодії між протонами біологічних тканин, постійним або змінним магнітним полем й енергією радіочастотних імпульсів, які випускають котушкою, поміщеною біля досліджуваної частини тіла. Під дією радіочастотних імпульсів протони атомів водню тимчасово переходять на більше високий енергетичний рівень. Повернення протонів до рівноважного стану супроводжується виділенням енергії у вигляді імпульсів певної (резонансної) частоти. Цю енергію можна виміряти за допомогою прийомної котушки. Для формування зображення отримані сигнали обробляються за допомогою перетворення Фур'є.

МРТ дозволяє одержувати зображення в різних площинах, не змінюючи положення хворого. Зображення в кожній площині формується за 5...10 хв. На відміну від комп'ютерної томографії, зміна положення хворого під час дослідження спотворює всі зображення.

Комп'ютерна томографія

Ідея комп'ютерної томографії (КТ) полягає в комп'ютерній реконструкції зображення поперечного зрізу тіла на підставі аналізу

поглинання рентгенівського випромінювання.

Трубка випускає тонкий віялоподібний пучок рентгенівських променів перпендикулярно довгої осі тіла. Цей пучок може бути досить широким й охоплювати весь діаметр тіла. Товщину пучка можна регулювати від 1 до 10 мм, при цьому змінюється товщина досліджуваного зрізу тканини. Пучок рентгенівських променів, який пропускається через пацієнта, фіксується не плівкою, а системою спеціальних детекторів. Рентгенівські фотони генерують у детекторах електричні сигнали. Чим сильніше інтенсивність первинного променя, який досяг детектора, тим сильніше електричний сигнал.

Фіксуючи інтенсивність пропущеного випромінювання, можна розрахувати ослаблення первинного променя. КТ- детектори приблизно в 100 разів чутливіші за рентгенівську плівку при визначенні розходжень в інтенсивності випромінювання. Випромінювання реєструються по шкалі Хаунсфілда (HU) від -1024 до +3071.

Для одержання комп'ютерної томограми (зрізу) на обраному рівні необхідно виконати наступні операції: 1) формування рентгенівського променя необхідної ширини (колімація); 2) сканування пучком рентгенівського випромінювання, здійснюваного рухом (обертальним і поступальним) навколо нерухомого об'єкта пристрою «випромінювачі-детектори»; 3) вимір випромінювання й визначення його ослаблення з наступним перетворенням результатів у цифрову форму; 4) комп'ютерний синтез томограми по сукупності даних вимірів, які ставляться до обраного шару; 5) побудова зображення досліджуваного шару й відображення його на екрані монітора.

Повітря або газ (поза залежністю, де вони перебувають) мають значення біля -1000 HU і виглядають чорними, кістки (+400 HU і вище) – білими. Інші органи, тканини й ділянки патологічних змін відображаються відтінками сірого між цими двома крайностями (від 0 до +300). Оскільки відобразити й розрізнити 210 відтінків не можна, при візуалізації даних КТ вибирається діапазон певних значень (вікно) для аналізу потрібних тканин (рис. 3.18).

Якість зображення зростає при збільшенні числа детекторів і кількості реєстрованих проекцій за один оберт трубки й при збільшенні первинної матриці. Однак збільшення кількості реєстрованих проекцій веде до підвищення променевого навантаження, застосування більшої первинної матриці – до збільшення часу обробки зрізу або необхідності встановлювати додаткові спеціальні процесори відеозображення.

Сучасний томограф дозволяє одержувати зрізи товщиною 1...2, 5 й 10 мм зі швидкістю сканування шару 1...3 с, повне дослідження головного мозку займає 2...3 хв.

Відмінності зображень комп'ютерної томографії від магнітно-резонансної полягають у тім, що на звичайному рентгенівському знімку

найкраще помітні кістки, тому КТ використовується для дослідження хребта, підстави черепа й скроневиx костей.

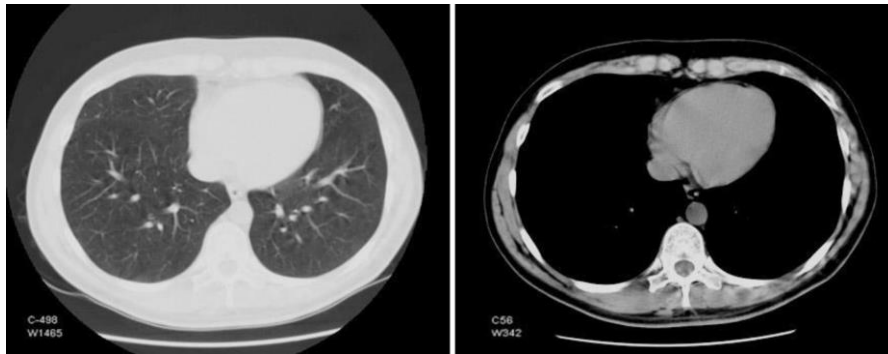


Рисунок 3.18 – Томограма грудної клітки в легеновому (ліворуч) і м'якотканому (праворуч) вікнах шкали Хаунсфілда

Крім того, на КТ-зображенні добре видно свіжу кровотечу, тому цей метод застосовують при дослідженнях хворих із травмами голови, грудної клітини, черевної й тазової порожнин, а також інсультів у ранній стадії. Використання контрастної речовини дозволяє одержати якісне зображення судин, нирок і кишечника. На знімках МРТ кістки теж видні, але краще видні м'які тканини (хрящі, м'язи, зв'язування, мозок і т.д.). МРТ більш чутлива за КТ на ранніх стадіях поразки центральної нервової системи. Тому даний метод показаний при підозрі на розрив зв'язувань і для виключення грижі міжхребцевого диска, а також широко використовується у нейрохірургії й неврології (старі травми головного мозку, інсульти в пізній стадії, підозри на пухлини спинного й головного мозку й т.д.). Різницю в зображенні різних тканин методами МРТ і КТ можна побачити, порівнявши рис. 3.17 і 3.19.

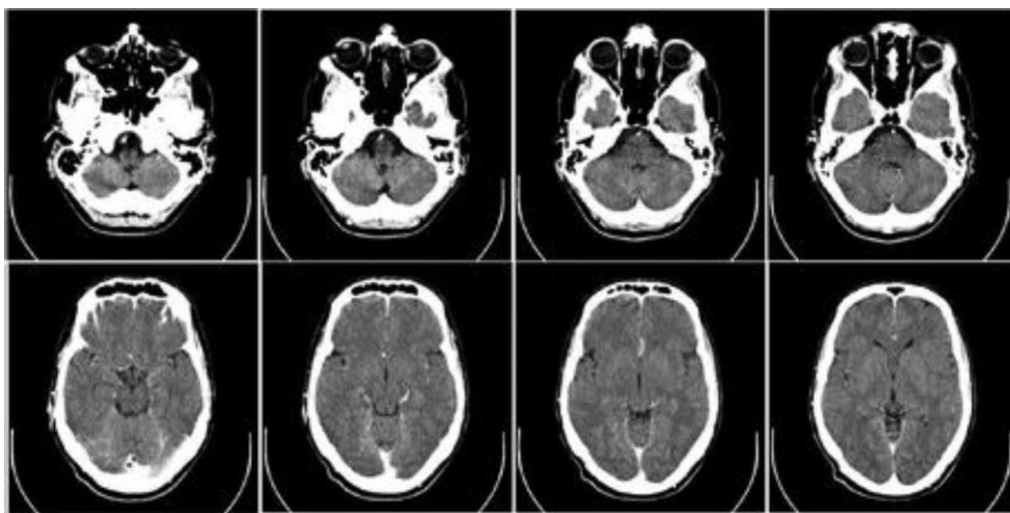


Рисунок 3.19 – КТ-зображення голови людини

3.7 Зображення, одержувані при дистанційному зондуванні Землі

Під дистанційним зондуванням Землі (ДЗЗ) розуміють одержання інформації про стан поверхні Землі по вимірах на відстані без безпосереднього контакту датчиків з поверхнею. Датчики встановлюються на супутниках, літаках і безпілотних літальних апаратах. Діапазон вимірюваних електромагнітних хвиль – від часток мікрометра (видиме оптичне випромінювання) до метрів (радіохвилі). Методи ДЗЗ можуть бути пасивні, тобто з використанням природного відбитого або вторинного теплового випромінювання об'єктів на поверхні Землі, обумовленого сонячною радіацією, і активні – які використовують відбиття, ініційоване штучним джерелом, установленому на літальному апараті. Одержання інформації для ДЗЗ ґрунтується на тім, що об'єкти різних типів (ґрунт, вода, рослинність, будинки, техніка й ін.) по-різному відбивають і поглинають випромінювання в різних діапазонах електромагнітного спектра.

У цей час для ДЗЗ використовуються різні системи для формування зображення досліджуваної місцевості. Їх можна розділити на декілька класів, які розрізняються за спектральним діапазоном використовованого електромагнітного випромінювання, за типом приймача реєстрованого випромінювання й за методом зондування (активний або пасивний):

- фотографічні системи;
- сканувальні системи видимого й інфрачервоного діапазону (телевізійні, сканувальні радіометри та багатоспектральні сканери);
- телевізійні оптичні системи;
- радіолокаційні системи бічного огляду;
- сканувальні СВЧ радіометри.

Атмосфера поглинає частину променів електромагнітного спектра, але в ній є кілька вікон прозорості (рис. 3.20), які пропускають електромагнітні хвилі з мінімальними перекручуваннями. Найчастіше в методах ДЗЗ використається вікно прозорості, яке відповідає оптичному діапазону, що поєднує видиму (380...720 нм), ближню інфрачервону (720...1300 нм) і середню інфрачервону (1300...3000 нм) області.

Отримані знімки класифікуються за просторовим відрізненням (розміру ділянки місцевості, що відповідає одному пікселю):

- низьке (300...1000 м);
- середнє (50...200 м);
- відносно високе (20...40 м);
- високе (10...20 м);
- дуже високе (1...10 м);

– надвисоке (0,3...0,9 м).

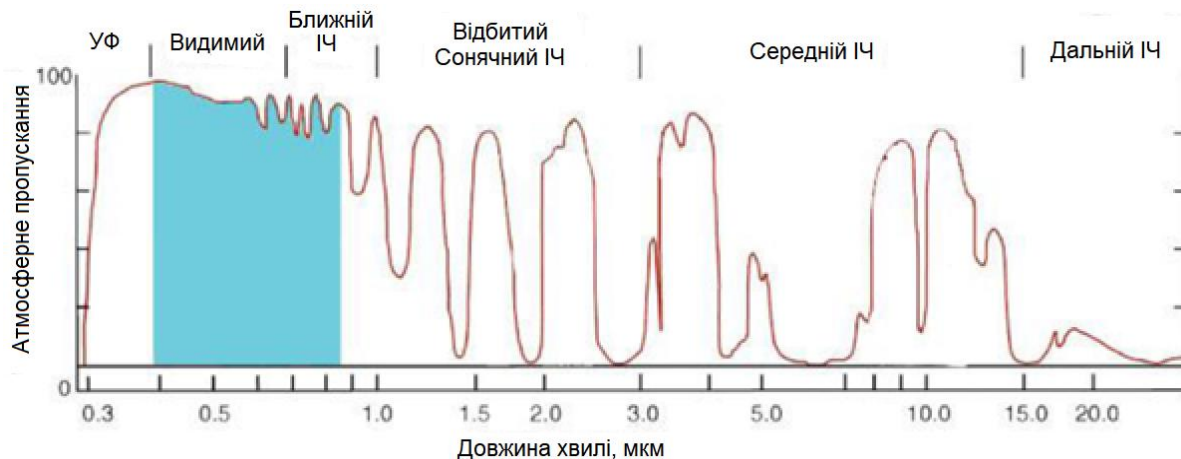


Рисунок 3.20 – Ступінь прозорості атмосфери для різних довжин хвиль спектра

Основні типи зображень, одержувані засобами ДЗЗ.

Мультиспектральні зображення

Мультиспектральні камери – найпоширеніші види сенсорів, установлених на супутниках ДЗЗ. Відбиті хвилі проходять крізь світлофільтри, що виділяють певну ділянку електромагнітного спектра, потім через фокусувальні лінзи, дисперсний пристрій і попадають на ПЗЗ-лінійку, де реєструються у вигляді напівтонового зображення. Число фільтрів визначає кількість каналів мультиспектрального зображення (табл. 3.1). Як правило, всі зображення записуються в один файл у форматі GeoTIFF.

Таблиця 3.1 – Супутники зі знімальними системами високого відрізнєння

Супутник	Ран, мкм	В, мкм	G, мкм	R, мкм	NIR, мкм	Інші	Відрізн., м	Біт
WorldView-3	.45-.80	.4-.5 фіолет. й .45-51	.51- 58 й .585-.625 жовтий	.63-.69 й .705-.745	.77-.895 й .86-1.04	Немає	.31; 1.24	11
WorldView-2	Те ж	Те ж	Те ж	Те ж	Те ж	Немає	.5; 2	11
GeoEye-2	.45-.90	.45-.52	.52-.60	.625-.695	.76-.90	Немає	.25; 1.65	11
Ikonos	Те ж	Те ж	.51-60	.63-.70	Те ж	Немає	1; 4	11
QuickBird	Те ж	Те ж	.52-.60	.63-.69	Те ж	Немає	.61; 2.44	11
OrbView-3	Те ж	Те ж	.52-.60	.625-.695	Те ж	Немає	1; 4	11
Spot-6 і 7	.45-.745	Те ж	.53-.59	.625-.695	.76-.89	Немає	2; 6	12
СІЧ-2 (Україна)	.50-.90	немає	.50-.60	.61-68	.80-.89	1.55-7 мкм	7.8; 39.5	8
БелКА і Канопус-В	.45-.85	.46-.52	.52-.60	.63-.69	.75-.86	Немає	2.1; 10.5	8

Як видно з табл. 3.1, у кожному каналі реєструється напівтонове зображення з великим динамічним діапазоном значень яскравості. Слід зазначити два факти: чим ширше спектральний діапазон відбиваного сигналу, тим вище відрізнення (приклад: панхроматичний режим має відрізнення в 2...4 рази вище спектральних); спектральні діапазони можуть не перетинатися й навпаки, все залежить від використовуваних типів фільтрів.

Гіперспектральні зображення

Ключовим моментом при побудові гіперспектральних систем є спосіб формування тривимірної структури даних з двовимірного масиву (ПЗЗ-матриця) або одновимірного масиву (ПЗЗ-лінійки) фотоприймачів. Пристрій одержання таких зображень називається гіперспектрометром. Найпоширенішими на сьогоднішній день є гіперспектрометри, які в кожен момент часу реєструють вузьку смугу поверхні під собою (рис. 3.21).

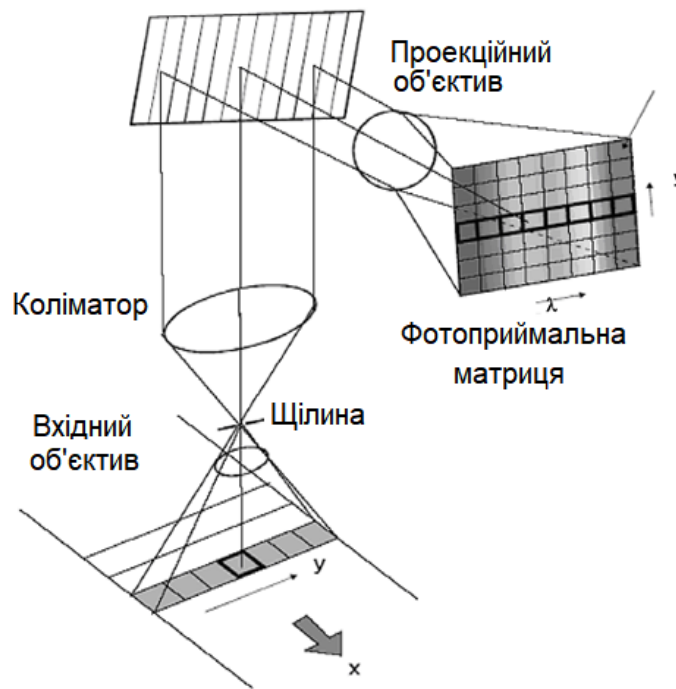


Рисунок 3.21 – Принципова схема гіперспектрометра

Формування зображення вузької ділянки поверхні поперек напрямку польоту виробляється за допомогою щілини на задній фокальній площині вхідного об'єктива.

Гіперкубом називають тривимірний масив даних, два виміри якого відповідають просторовим координатам (x , y), третє – спектральній координаті λ (рис. 3.22). Його одержують при зйомці кадру шляхом фіксації однієї з координат (або просторової, наприклад, x , або спектральної λ). За один кадр фотоприйомна матриця формує двовимірний зріз гіперкуба з фіксованою третьою координатою.

Після колімуєчого об'єктива, відбитий сигнал у паралельних променях попадає на призму. У призмі відбувається його розкладання на хвилі різної довжини, а потім вони проектується на фотоприйомну матрицю. У такий спосіб на матриці формується зріз гіперкуба для фіксованої просторової координати x . Більшість спектрометрів реєструють 100...200 каналів шириною (спектральним відрізненням) по 0,4...2,5 нм кожний у діапазоні 400...2500 нм без проміжків і накладень.

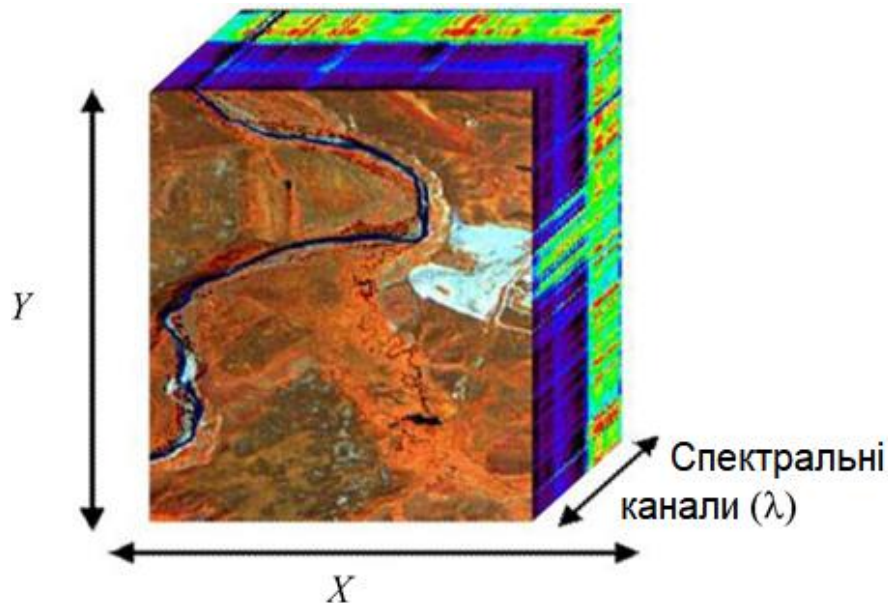


Рисунок 3.22 – Гіперспектрометр сканує простір уздовж осі X (або Y)

Мультиспектральні та гіперспектральні знімки можуть покривати однаковий спектральний діапазон, але відрізняються способами реєстрації зображень. Мультиспектральне є набором з напівтонових зображень однієї ділянки місцевості, зареєстрованих у декількох вузьких діапазонах спектра, виділених шляхом фільтрації. Гіперспектральне одержують шляхом розкладання за допомогою призми сонячного випромінювання, відбитого від ділянки Землі, що відповідає одному пікселю з координатами (x, y) , на лінійку сенсорів, кожний з яких реєструє сигнали у вузькому діапазоні спектра. Всі значення яскравості, зареєстровані лінійкою, привласнюються одному пікселю.

Наприклад, датчик, що формує 20 зображень, створює гіперспектральне зображення, якщо воно являє собою відбиття в діапазоні від 500 до 700 нм, пропущене через призму й зареєстроване 20 сенсорами, кожний з яких охоплює частину цього діапазону, рівну 10 нм завширшки. У той час як датчик, що реєструє 20 дискретних смуг у цьому ж діапазоні, які отримані шляхом фільтрації променів, тобто пропуски й реєстрації хвиль тільки однієї смуги – мультиспектральне. Отримані результати будуть відрізнятися, оскільки при виділенні вузьких діапазонів спектра

використаються фізично різні способи, а фільтрацією неможливо повністю розділити хвилі різної довжини.

Радіолокаційні зображення

Радіолокаційна зйомка заснована на здатності об'єктів по-різному відбивати вузьконаправлені на них радіоімпульси сантиметрового діапазону (0,3...100 см).

Властивості матеріалів впливають на інтенсивність відбитого сигналу. Коефіцієнт діелектричної проникності характеризує здатність матеріалу поглинати електричну енергію й залежить від використовуваної частоти (рис. 3.23). Розходження значень коефіцієнтів для різних матеріалів обумовлює можливість їхнього виявлення за допомогою радарної зйомки (табл. 3.2). Реєструвати сигнали радарного діапазону та одержувати зображення можна в умовах хмарності й у темний час доби.

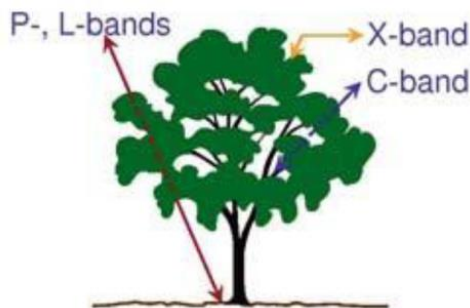


Рисунок 3.23 – При різних частотах сигнал відбивається від різних поверхонь

Нагріті об'єкти випромінюють енергію й у радіодіапазоні, хоча радіотеплове випромінювання по інтенсивності значно уступає тепловому випромінюванню в інфрачервоному діапазоні. Радіотеплове випромінювання, як й інфратеплове, формується поверхневим випромінювальним шаром, товщина якого тим більше, чим довше хвиля випромінювання.

Таблиця 3.2 – Частоти і довжини хвиль, використувані в радіолокації

Назва	Частота, МГц	Довжина хвилі, см
Ka-band	40,000...26,000	0,8...1...1,1
K-band	26,500...18,500	1,1...1...1,7
Ku-band	18,500...12,500	1,7...2...2,4
X-band	12,500...8,000	2,4...3...3,8
C-band	8,000...4,000	3,8...7...7,5
S-band	4,000...2,000	7,5...15...15,0
L-band	2,000...1,000	15,0...30...30,0
P-band	1,000...300	30,0...100,0

Інтенсивність радіотеплового випромінювання характеризується так називаною радіояскравісною температурою – добутком абсолютної температури і коефіцієнта випромінювання. Величина коефіцієнта випромінювання в радіодіапазоні значно варіює залежно від електричних властивостей (а виходить, від вологості й солоності), кристалічної структури та характеру поверхні об'єкта. Найбільше його значення має рослинність (0,9), потім ґрунти (0,8), водна поверхня (0,3) і, нарешті, металеві предмети. Таким чином, розходження в радіояскравісних температурах земних об'єктів визначається їхньою випромінювальною здатністю, а не фізичною температурою. Влітку найбільшу радіояскравісну температуру має рослинний покрив (300 К) і меншу – водні поверхні (100 К). Але самі «холодні» у радіодіапазоні – це металеві предмети.

Активне зондування земної поверхні засновано на вимірі варіацій розсіяного випромінювання. Інформація про поверхні представлена відбитою електромагнітною хвилею. Радіолокатор випускає хвилю. Вона характеризується довжиною і частотою (яка дорівнює відношенню швидкості поширення в середовищі до довжини хвилі, у вакуумі швидкість дорівнює швидкості світла). Крім напрямку хвилі важлива її поляризація. Конструктивно радіолокаційні системи можуть здійснювати випромінювання й прийом на вертикальній і горизонтальній поляризації сигналу (рис. 3.24). Прийнятий у результаті сигнал залежить від типу поляризації при випромінюванні й при прийомі, його тип позначається двома буквами:

- вертикальна поляризація (тип ВВ, або VV);
- горизонтальна (тип ГГ, або GG);
- перехресна (тип ВГ і ГВ, або VG й GV).

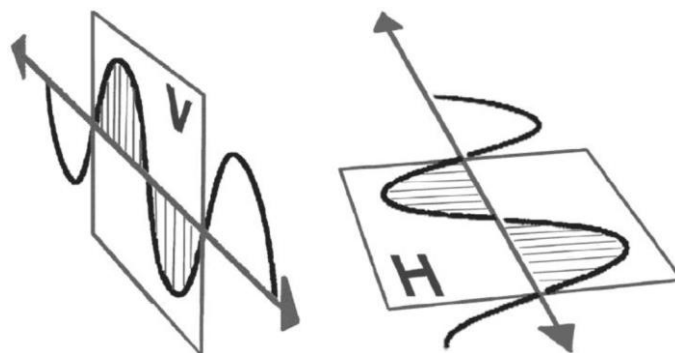


Рисунок 3.24 – Вектор електричного поля при вертикальній V (ліворуч) і горизонтальній H (праворуч) поляризації

Потужність сигналу, що попадає в прийомну антену, залежить від типу поляризації випромінювання та взаємного напрямку поляризації при випромінюванні й прийомі. Коефіцієнт розсіювання водної поверхні при вертикальній поляризації вище, ніж при горизонтальній і перехресній

поляризації. Вертикальна поляризація при випромінюванні й прийомі (тип VV) застосовується при аналізі процесів на водній поверхні. Горизонтальна поляризація (тип GG) – для поділу льоду й води. Перехресні поляризації використовують для виділення на морській поверхні різних об'єктів (кораблів, тріщини у льоді й т.п.).

На рис. 3.25, а чітко видні прояви поверхневих атмосферних і морських процесів, на рис. 3.25, б – відбиття від суден. Вертикально поляризована хвиля взаємодіє з вертикальними стеблами шару рослин, горизонтально поляризована проникає крізь шар. Комбінування різних видів поляризації може поліпшувати детальність класифікації, наприклад, дозволяє визначати різні класи рослинності.

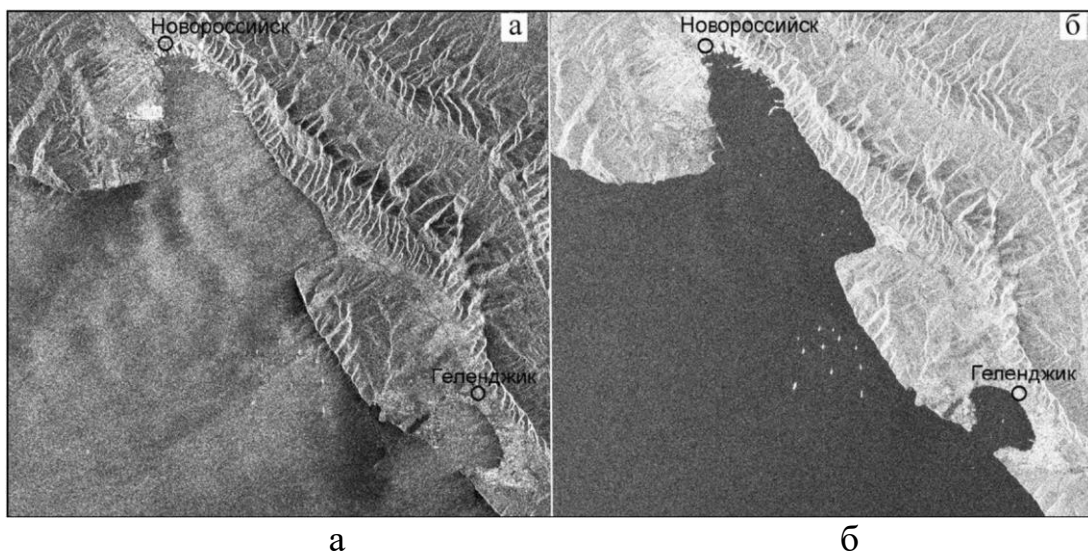


Рисунок 3.25 – Приклади зображень, отриманих методом синтезу апертури із супутника Envisat ASAR: поляризація типу VV (а), типу VG (б)

Відрізнєння при радіолокації залежить від розміру антени і довжини хвилі зондувального сигналу. Радіолокатори, установлені на літаках при відстані до землі порядку 10 км і довжині хвилі 2 см, мають відрізнєння близько 15 м. Якщо подібну апаратуру використати на супутнику, то відстань до Землі складе 500...1000 км і відрізнєння зросте до 1000...2000 м.

Для підвищення розрізняльної здатності використовують метод синтезу апертури, тобто штучне збільшення розміру антени. У цьому методі виробляється когерентна (зі збереженням фази) обробка розсіяних сигналів протягом фіксованого проміжку часу, за який супутник пролітає деяку відстань (рис. 3.26). Ця відстань відіграє роль довжини синтезованої антени. Наприклад, при довжині хвилі 5 см і дальності до Землі 1000 км розрізняльна здатність складе 5 м. Подібні системи називають радіолокаторами із синтезованою апертурою (РСА).

На рис. 3.27 наведені для порівняння зображення однієї місцевості,

zareestrovani v riznikh spektralnykh diapazonakh.

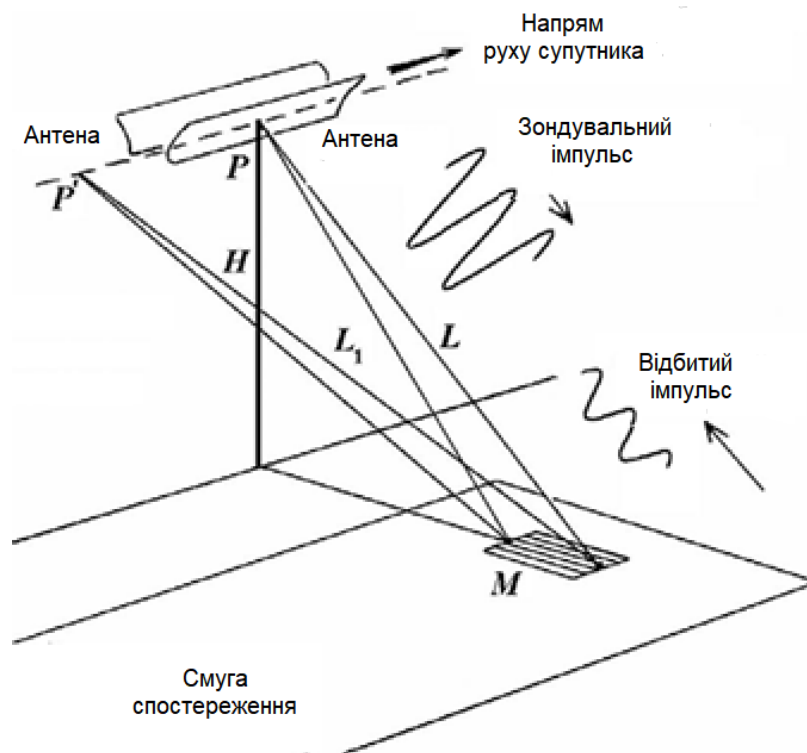


Рисунок 3.26 – Ідея методу синтезованої апертури

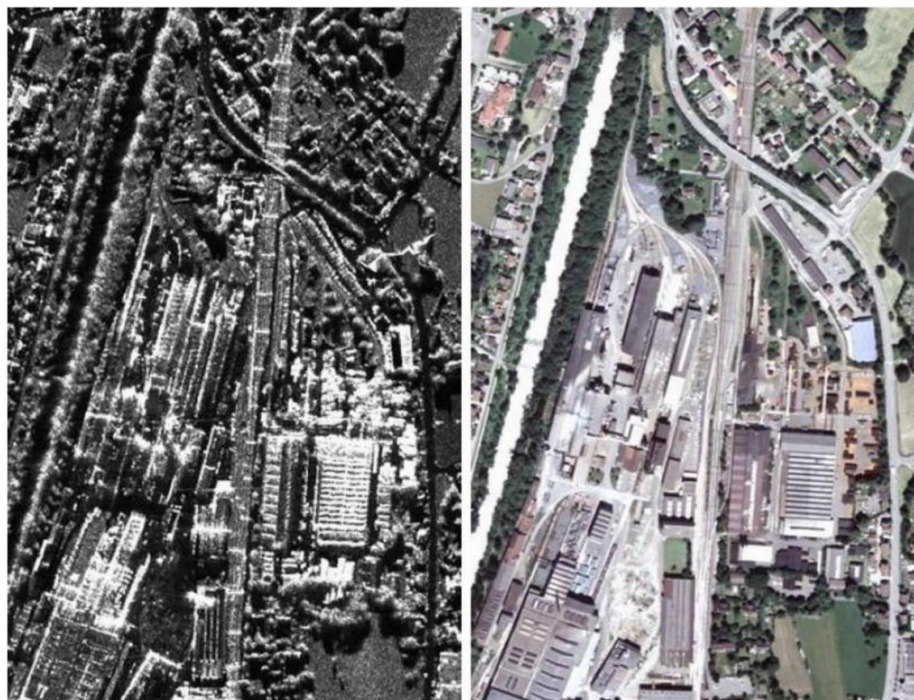


Рисунок 3.27 – Приклади зображень, отриманих методом синтезу апертури (ліворуч) і у видимому діапазоні (праворуч)

На лівому зображенні можна помітити спекл-шум, що представляє посилення яскравості світлих об'єктів у вигляді світлих вертикальних хрестів. Цей вид шуму з'являється тільки на радарних зображеннях у результаті випадкової інтерференції й накладення відбитих сигналів від різних об'єктів. Один з методів зменшення цього шуму – фільтрація, наприклад, медіанна.

При реєстрації відбитих сигналів від гористої місцевості може фіксуватися її перевернуте зображення в силу різниці між пройденими шляхами посланого й відбитого сигналу (рис. 3.28 і 3.29).



Рисунок 3.28 – Перекручування похилої дальності міняє пропорції об'єктів (праворуч реальні пропорції)

Радари здатні вимірювати не тільки амплітуду, але й точний момент коливань (фазу) відбитих хвиль. Системи, що реєструють амплітуди й фази сигналів, називаються когерентними. Що дає знання фази коливань? Припустимо, що із супутника, що перебуває на орбіті висотою 800 км посилає до Землі радіосигнал із частотою 6 ГГц. Довжина хвилі при цьому становить 5 см. Сигнал досягає поверхні Землі й відбивається назад, проходячи загальний шлях в 1600 км, що відповідає дуже великому, але цілому числу довжин хвиль. Фаза коливань при цьому залишається незмінною. Однак якщо відстань до Землі перевищує 800 км хоча б на один сантиметр, то хвиля повинна пройти шлях на два сантиметри більше, що становить 40% від довжини хвилі. У результаті, коли фаза відбитої хвилі досягне супутника, вона буде відставати на 40% періоду коливань, що можна легко зареєструвати. Таким чином, вимір фази дає можливість оцінити відстань до цілі із сантиметровою або навіть міліметровою

точністю.

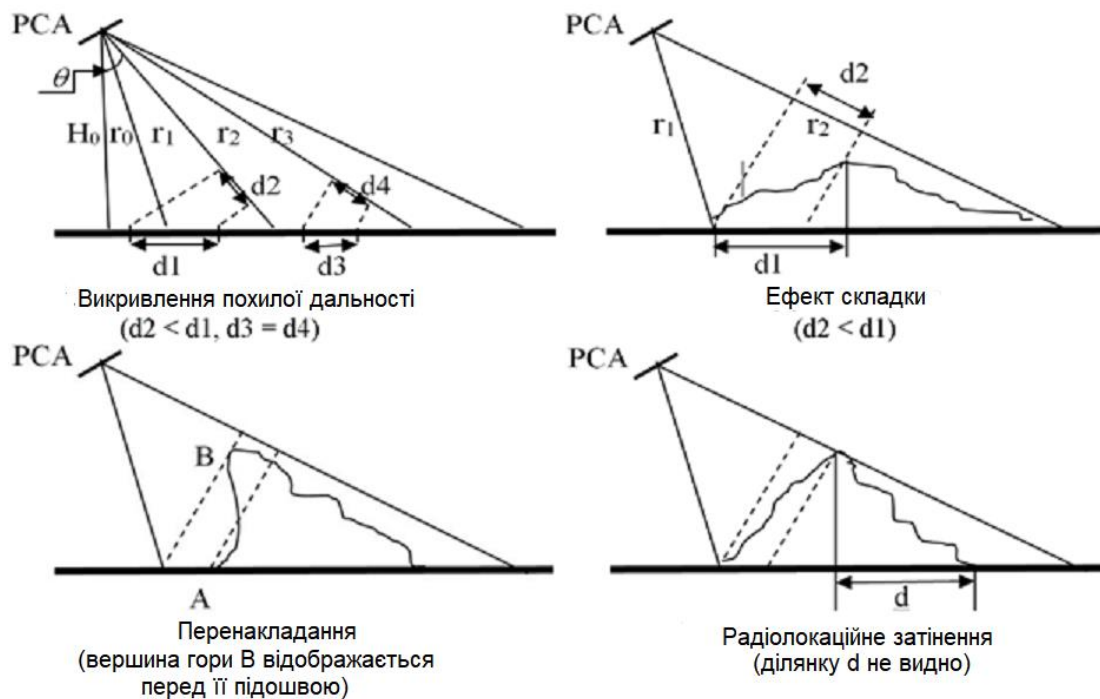


Рисунок 3.29 – Вплив геометрії зйомки і рельєфу на формування радарних зображень

У випадку, якщо два послідовних радарних знімки отримані при зовсім однакових положеннях супутника, то не повинно бути ніякої різниці фаз для будь-якої пари відповідних точок. Якщо сцена на Землі змінилася в проміжку між зйомками хоча б незначно, то фази деяких пікселів будуть відрізнятися. Таким чином, на основі двох радарних знімків можна одержати інтерферограму, яка дозволяє відслідковувати зсув земної поверхні в кілька міліметрів.

З урахуванням викладеного дуже важливо правильно підбирати тип поляризації та довжину хвилі для конкретного проекту, щоб покращити характеристики результату ДЗЗ.

Контрольні запитання

1. Як побудований і як працює елементарний чутливий елемент приладу із зарядовим зв'язком ?
2. Що таке мозаїка Байера ?
3. Як побудований і як працює елементарний чутливий елемент матриці комплементарних металооксидних напівпровідників ?
4. Що таке знання ?
5. Що таке сканер ?
6. Як працює однопрохідний планшетний сканер ?

7. Які існують конструкції цифрових фотоапаратів ?
8. Як працює цифрова фотокамера з заднім розгорненням ?
9. Як працює трьохкадрова цифрова фотокамера ?
10. Як працює однокадрова цифрова фотокамера з однією матрицею ?
11. Як працює однокадрова цифрова фотокамера з трьома матрицями ?
12. Як одержуються зображення за допомогою апарату ультразвукової діагностики ?
13. Як одержуються рентгенівські зображення ?
14. Як одержуються зображення за допомогою магнітно-резонансного томографа ?
15. Як одержуються зображення за допомогою комп'ютерного томографа ?
16. За допомогою яких спостережних систем проводиться дистанційне зондування Землі ?
17. Що таке мультиспектральне зображення ?
18. Що таке гіперспектральне зображення ?
19. Які існують види радіолокаційних зображень ?

4 ОБРОБКА ЦИФРОВИХ ЗОБРАЖЕНЬ НА ОСНОВІ ЇХНІХ ГЕОМЕТРИЧНИХ ПЕРЕТВОРЕНЬ

4.1 Класифікація методів цифрової обробки зображень

Цифрова обробка зображень – це будь-яка зміна даних, представлених у вигляді цифрових зображень. Результатом обробки також є цифрове зображення (рис. 4.1).

Метою обробки зображень є перетворення зображень, орієнтоване на візуальне сприйняття людиною (в основному підвищення якості, корекція кольорів і контрасту, виправлення дрібних перешкод або перекручування зображення) і їхній автоматизований аналіз (наприклад, фільтрація, математична морфологія, сегментація на області певних класів, виділення об'єктів і т.п.).

Серед завдань обробки зображень можна виділити два основні: *покращення та відновлення зображень*.

Головна мета покращення полягає в такій зміні зображення, щоб результат виявився найбільш підходящим з погляду конкретного застосування. Загальної теорії покращення зображень не існує. Наприклад, метод, оптимальний для покращення рентгенівських зображень, як правило, не підходить для обробки кольорових фотографій і навпаки.

Якщо зображення обробляється для візуальної інтерпретації, то користувач буде судити, наскільки гарний застосований метод. Візуальне

оцінювання якості зображення – це суб'єктивний процес, а поняття гарного зображення неможливо формалізувати або оцінити чисельно. Коли метою є обробка зображення для комп'ютерного аналізу, завдання оцінювання простіше. Наприклад, при рішенні завдання розпізнавання символів кращим буде той метод обробки зображення, завдяки якому будуть отримані більш точні результати машинного розпізнавання.

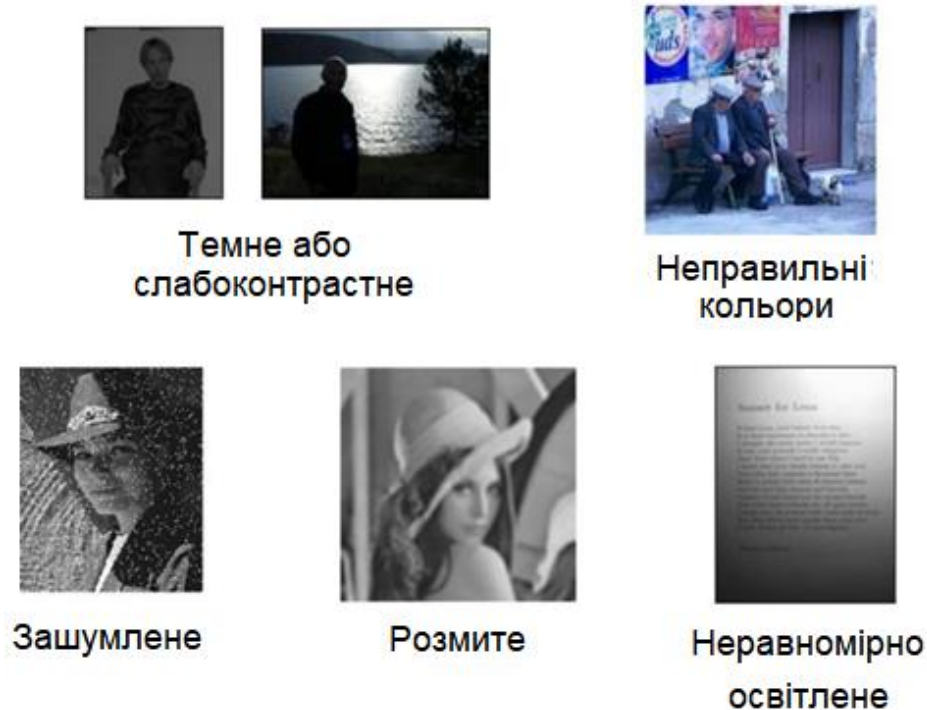


Рисунок 4.1 – Приклади зображень, які вимагають коректування

При відновленні зображення також виконується підвищення його якості, але відомий тип перетворення, яке спотворило зображення (наприклад, змазане зображення автомобіля на фотографії). Незважаючи на перетинання областей застосування методів обох класів, покращення зображень орієнтоване на суб'єктивне сприйняття результату, у той час як відновлення може бути оцінене більш об'єктивно. При відновленні виконується реконструкція зображення, яке було перекручено в процесі реєстрації (наприклад, змазане), за допомогою використання апріорної інформації про явище, яке викликало погіршення зображення. Тому методи відновлення засновані на моделюванні процесів перекручування, застосування зворотних процедур для відтворення вихідного зображення.

Таким чином, покращення зображень виконується як для візуального сприйняття результату людиною, так і для автоматичного аналізу й розпізнавання зображень. У чому різниця? Людське око не відрізняє на напівтоновому зображенні пікселі, що мають значення яскравості, наприклад, 10, 14 і 18, а при автоматичному аналізі треба вирішувати,

утворюють ці пікселі одну однорідну область чи ні.

Методи обробки зображень можна класифікувати різними способами. Зокрема, можна виділити групи методів залежно від їхнього призначення.

Виправлення дефектів зображення, обумовлених:

- шумом (випадковими похибками кольорів в окремих пікселях зображення);
- недостатньою або надлишковою яскравістю;
- недостатньою контрастністю або вузьким динамічним діапазоном зображення;
- неправильним колірним тоном;
- нерізкістю (розфокусуванням);
- перекручуваннями через пил, подряпин на сканованому документі;
- дісторсії (скривленням);
- необхідністю ретуші фотопортретів: усунення прищів, подряпин, шрамів, синців, пор, ластовиння або зменшення їхньої кількості, розгладження зморшок; відбілювання зубів; заміна кольорів волосся, очей; корекція форми фігури й т.п.

Структурне редагування зображень:

- кадрування;
- створення панорам;
- усунення непотрібних деталей зображення;
- фотомонтаж – створення з декількох зображень нового зображення;
- вставка креслень, написів, символів, покажчиків та ін.;
- застосування спецефектів, фільтрів, тіней, текстур.

Підготовка фотографій до публікації (у пресі, інтернеті):

- різні пристрої виводу (монітор, принтер, офсетна друкована машина й т.п.) мають різні можливості колірної охоплення (не будь-які кольори можна відтворити). Виконується корекція кольорів для печатки на папері, визначається кількість фарби кожних кольорів.

Для усунення перекручувань, внесених у процесі реєстрації зображень, використовуються методи геометричної і яскравісної корекції. При обробці космічних знімків застосовують атмосферну корекцію зображень. Окреме місце займають методи обробки й аналізу зображень за допомогою математичної морфології.

Методи обробки умовно розділяють на три рівні:

- низький (не використовуються додаткові знання, моделі й інформація про об'єкти, представлених на зображенні);
- середній (частково використовуються додаткові відомості);
- високий (широко використовуються додаткові відомості й контекст зображення, тобто інформація про зображені об'єкти).

Відповідно до обчислювальної складності, тобто з урахуванням

кількості пікселів, що беруть участь в обчисленні значення яскравості одного пікселя нового зображення, методи ділять на *три класи*:

- точкові (в обчисленнях використовується значення одного пікселя вихідного зображення);
- локальні (в обчисленнях використовуються значення декількох сусідніх пікселів у заданій околиці);
- глобальні (при обчисленні значення пікселя використовуються всі пікселі вихідного зображення).

Обчислювальна складність алгоритмів пропорційна кількості пікселів, які беруть участь в обчисленні нового значення яскравості одного пікселя.

Лінійний функціонал – це лінійне відображення множини функцій у множині чисел із властивостями

$$F[f+g] = F[f] + F[g], \quad F[cf] = cF[f], \quad (4.1)$$

де F – лінійний функціонал; f й g – функції з його області визначення; c – константа.

Всі методи обробки зображень можна розділити на лінійні та нелінійні. Лінійні методи допускають спрощення (і, отже, прискорення) обчислень, а також реалізацію за допомогою операції згортки.

Обробка зображень може здійснюватися рекурсивними й нерекурсивними методами. Рекурсивні методи використовують результат обробки частини сусідніх пікселів, нерекурсивні – не використовують. В обробці зображень більшість методів є нерекурсивними.

Багато методів обробки зображень адаптовані з різних галузей науки: обробки багатомірних сигналів, математики, фізики, фотографії, поліграфії, картографії й ін.

4.2 Геометричні перетворення зображень

Геометричні перетворення полягають у просторовій зміні розташування сукупності пікселів з однієї двовимірної системи координат в іншу. При цьому зображення зручно розглядати у вигляді пікселів, заданих у прямокутній системі координат. При виконанні геометричних перетворень вихідна множина пікселів із цілочисельними координатами (x, y) перетвориться в нову множину (сітка растра трансформується просторово) з координатами (x', y') і збереженням яскравості в цих пікселях. Оскільки цифрове зображення являє собою дискретну множину, коротко опишемо основи геометричних перетворень для безперервних множин, а потім адаптуємо їх для растрових зображень.

При вивченні геометричних перетворень плоских зображень (2D) передбачається, що маємо евклідовий простір с ортонормованою

декартовою системою координат із взаємно ортогональними осями. Тоді кожному пікселю зображення ставиться у відповідність пари (x, y) декартових координат. Їх можна інтерпретувати як двовимірний вектор \mathbf{X} , представлений відрізком із точки $(0, 0)$ до точки (x, y) .

Двовимірні перетворення на площині можна представити, як рух точок, які відповідають множині пікселів. Виділимо лінійні перетворення, описувані матрицями, тобто перетворення, при яких нові координати точки лінійно залежать від вихідних координат цієї точки в такий спосіб: $x' = Tx$.

Для зручності матричного опису і комбінування геометричних перетворень в 2D-площині використовується формалізм «однорідні координати». Точки площини описуються тривимірними векторами виду $(x, y, 1)$, а перетворення – у матричному виді

$$\begin{bmatrix} \bar{x}' \\ \bar{y}' \\ w' \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} \bar{x} \\ \bar{y} \\ w \end{bmatrix}.$$

До лінійних перетворень можна віднести кусочно-лінійні, при яких зображення розбивається на частині, а потім до кожної з них застосовуються різні лінійні перетворення.

Нелінійні перетворення – це всі перетворення, які не задовольняють властивостям лінійності, наприклад, поліноміальні (виконувані за допомогою поліноміальних функцій), синусоїдальні й т.п.

Афінне перетворення

Основні типи афінних перетворень й їхніх матричних поданнях показані на рис. 4.2.

У цілому будь-яке афінне перетворення на площині описується рівняннями $x' = ax+bx+c$, $y' = dx+ey+f$ або представляється в матричному вигляді:

$$\mathbf{x}' = \mathbf{Ax} + \mathbf{c}, \quad \text{де} \quad \mathbf{x}' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} a & b \\ d & e \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c \\ f \end{bmatrix}.$$

В однорідних координатах це перетворення можна описати однією матрицею T або в такий спосіб:

$$\begin{bmatrix} x' & y' & 1 \end{bmatrix} = \begin{bmatrix} x & y & 1 \end{bmatrix} \cdot \begin{bmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ x'_0 & y'_0 & 1 \end{bmatrix}. \tag{4.1}$$

Опишемо параметри базових афінних перетворень.
 Зсув зображення задається співвідношеннями

$$x' = x + c, \quad y' = y + f.$$

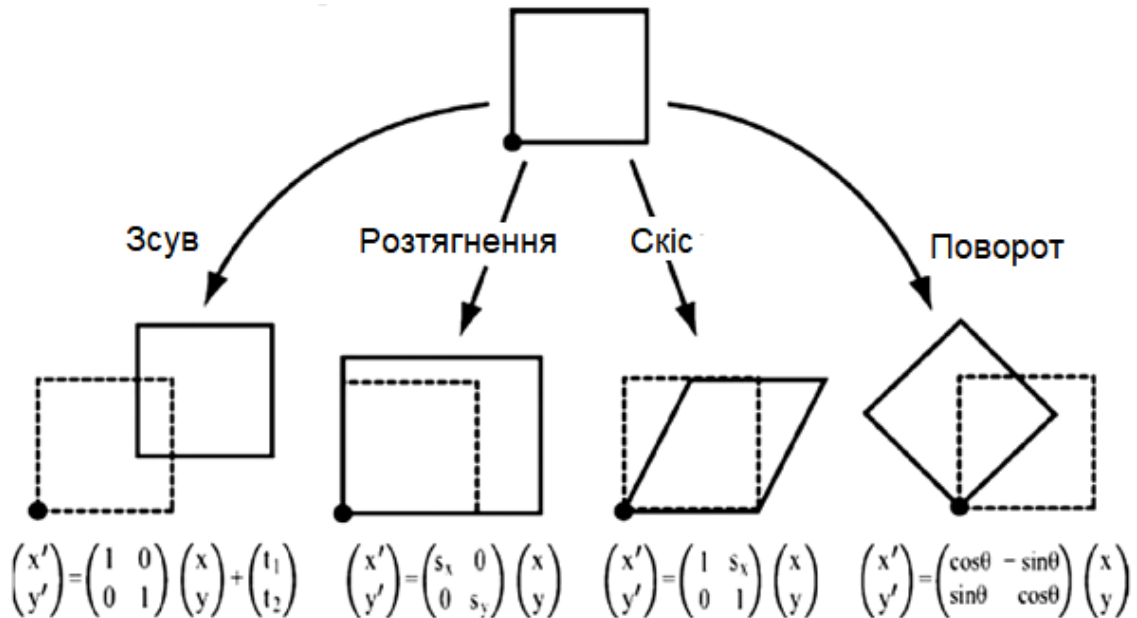


Рисунок 4.2 – Базові перетворення афінного типу

Матриця зсуву має вигляд

$$\mathbf{T}_{cd} = \begin{bmatrix} 1 & 0 & c \\ 0 & 1 & f \\ 0 & 0 & 1 \end{bmatrix}$$

Відбиття зображення щодо осі OY задається формулами

$$x' = x, \quad y' = -y.$$

Матриця відбиття має вигляд

$$\mathbf{T}_{om} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Масштабування (розтягнення) зображення задається у вигляді

$$x' = \alpha x, \quad \alpha > 0, \quad y' = \beta y, \quad \beta > 0.$$

Матриця перетворення в однорідних координатах має вигляд

$$\mathbf{T}_{\alpha\beta} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Необов'язково $\alpha=\beta$. Якщо α і β менше одиниці, зображення зменшується, якщо більше одиниці – збільшується. При збільшенні зображення кількість пікселів зростає, з'являються нові з невизначеним значенням яскравості. Їхні значення обчислюються шляхом інтерполяції, тобто локальними методами, описаними нижче.

Скіс зображення уздовж осі OX описується формулами

$$x' = x + sy, \quad y' = y.$$

Матриця перетворення в однорідних координатах має вигляд

$$\mathbf{T}_{ск} = \begin{bmatrix} 1 & s & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Поворот зображення навколо початку координат на кут φ описується формулами

$$x' = x \cos \varphi - y \sin \varphi, \quad y' = x \sin \varphi + y \cos \varphi.$$

Матриця обертання в однорідних координатах має вигляд

$$\mathbf{T}_{\varphi} = \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Відзначимо, що при повороті на 90° $\cos \varphi = 0$, $\sin \varphi = 1$, тому

$$x' = -y, \quad y' = x.$$

При повороті зображення координати пікселя можуть прийняти нецілі значення, у такому випадку потрібно їх округлити до цілих. Це створює в поверненому зображенні пікселі з невизначеною яскравістю (рис. 4.3).

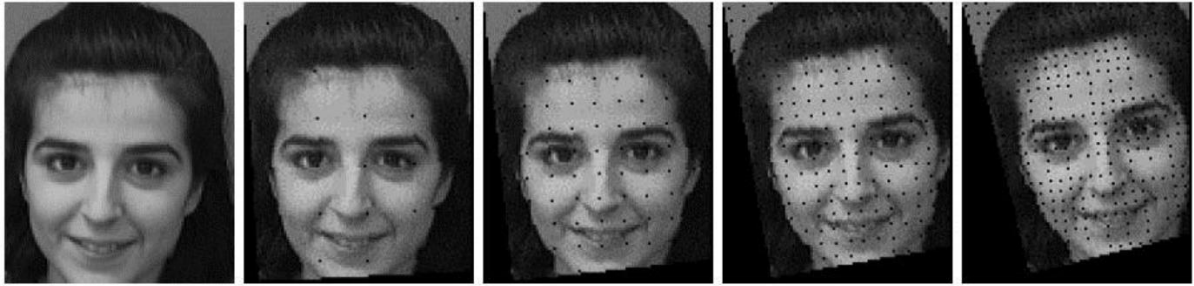


Рисунок 4.3 – Зображення повернене на 3, 6, 10 і 14° навколо верхнього лівого кута (у чорних пікселях яскравість не визначена)

Пікселям з невизначеною яскравістю можна привласнити значення яскравості найближчого пікселя. Альтернативою є поворот за допомогою виконання трьох операцій скосу уздовж осі OX , потім OY і ще раз уздовж OX . Матриця обертання описується у вигляді добутку трьох матриць скосу

$$T(\alpha) = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} = \begin{bmatrix} 1 & -\operatorname{tg} \alpha / 2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ \sin \alpha & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -\operatorname{tg} \alpha / 2 \\ 0 & 1 \end{bmatrix}$$

Після кожної операції скосу виконується округлення координат до цілих значень.

Довільне афінне перетворення можна записати у вигляді композиції послідовно виконуваних базових перетворень. Суперпозиція афінних перетворень також є афінним перетворенням. Перевага суперпозиції: складне перетворення зображення можна описати шістьма параметрами, обчисленими один раз, і застосувати до координат всіх пікселів.

У безперервній геометрії будь-яке афінне перетворення має зворотне перетворення, що також є афінним. Добуток прямого і зворотного дає одиничне перетворення, що залишає всі точки на місці.

Деякі властивості афінних перетворень:

- пряма лінія переходить у пряму;
- паралельні прямі переходять у паралельні;
- зберігаються відношення довжин відрізків, які лежать на одній прямій або паралельних прямих;
- зберігаються відношення площ фігур.

Безліч афінних перетворень утворюють афінну групу. Відзначимо, що Афінна група є підгрупою проєктивної групи перетворень, а евклідова – частковим випадком афінної. Ці групи формують наступну ієрархію:

евклідове перетворення \subset афінне \subset проєктивне перетворення.

Проективні перетворення

Загальними лінійними перетвореннями T у однорідних координатах

описується група проєктивних перетворень

$$\begin{bmatrix} \bar{x}' \\ \bar{y}' \\ 1 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} \\ t_{21} & t_{22} & t_{23} \\ t_{31} & t_{32} & t_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \quad (4.2)$$

При поданні у звичайних координатах співвідношення (4.2) буде мати нелінійний вигляд:

$$x' = \frac{t_{11}x + t_{12}y + t_{13}}{t_{31}x + t_{32}y + t_{33}}, \quad y' = \frac{t_{21}x + t_{22}y + t_{23}}{t_{31}x + t_{32}y + t_{33}}. \quad (4.3)$$

Проєктивні перетворення в загальному випадку не зберігають паралельності ліній. Властивістю, що зберігається при проєктивному перетворенні, є колінеарність точок: три точки, які лежать на одній прямій (колінеарні), після перетворення залишаються на одній прямій.

Проєктивне перетворення пов'язане з відображенням тривимірної візуальної інформації на двовірну площину. З математичної точки зору зручно розглядати світ як тривимірний проєктивний простір P^3 , а площина зображення – включеної в проєктивний простір розмірності два – P^2 . Точки на тривимірній сцені й на зображенні представляються в проєктивних просторах як вектори в однорідних координатах. Проєктивне перетворення з P^3 у P^2 (перспективна проєкція), що відображає евклідову Точку сцени $\mathbf{p}=(X,X,Z)'$ у точку зображення $\mathbf{x}=(x,y)'$ і виражене в однорідних координатах, задається рівняннями (4.2) або (4.3).

На рис. 4.4, б, в показані зображення, отримані проєктивними перетвореннями відповідно:

$$x' = \frac{1,1x + 0,35y}{0,00075x + 0,0005y + 1}, \quad y' = \frac{0,2x + 1,1y}{0,00075x + 0,0005y + 1}$$

і

$$x' = \frac{1,1x + 0,2y}{0,00075x + 0,00005y + 1}, \quad y' = \frac{0,1x + 0,9y}{0,00075x + 0,00005y + 1}$$

Однорідні координати векторів проєктивного простору $(\bar{X}, \bar{Y}, \bar{Z}, W)'$ проєктивної площини $(\bar{x}, \bar{y}, w)'$ співвідносяться з неоднорідними (евклідовими) координатами векторів \mathbf{p} і \mathbf{x} у такий спосіб:

$$\mathbf{p} = \left(\frac{\bar{X}}{W}, \frac{\bar{Y}}{W}, \frac{\bar{Z}}{W} \right)' \quad \text{і} \quad \mathbf{x} = \left(\frac{\bar{x}}{w}, \frac{\bar{y}}{w} \right)'.$$

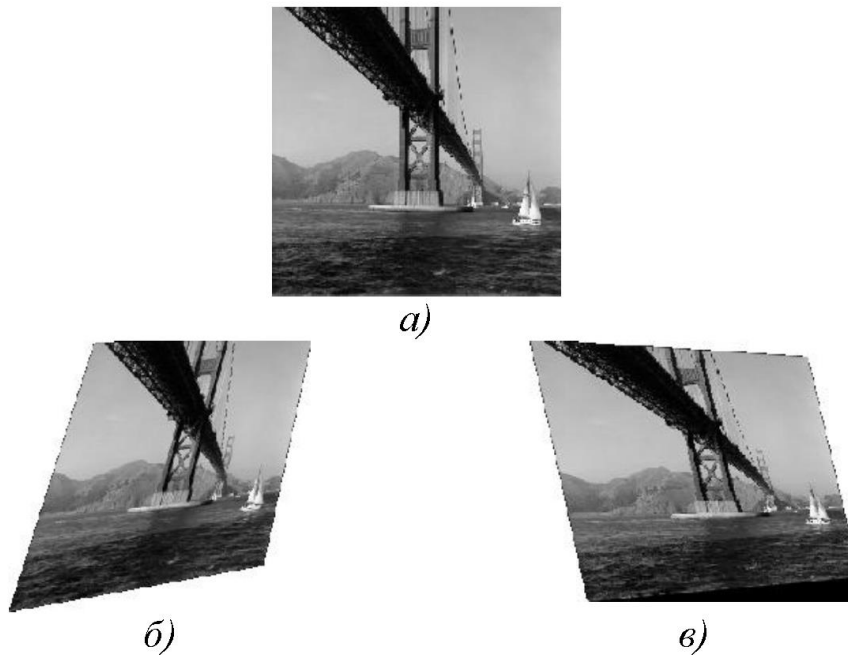


Рисунок 4.4 – Вихідне зображення (а) і два варіанти його проєктивного перетворення (б, в)

Зображення об'єктів на знімку, сформованому реєструвальною камерою, пов'язане із проєктуванням за допомогою пучка прямих, оскільки кожна 2D-точка є проєкцією множини 3D-точок уздовж деякого променя проєктування в площину знімка.

При розгляді геометричних перетворень використовується ідеальна модель камери. У дійсності формування зображень супроводжується різними нелінійними перекручуваннями (типу оптичної дісторсії лінзи). Приведення знімків до реального подання можна виконати нелінійною функцією геометричного перетворення. Для цього звичайно використовують поліноміальну апроксимуючу функцію перетворення:

$$\begin{aligned} x' &= a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 + \dots + a_px^m; \\ y' &= b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + \dots + b_py^m, \end{aligned}$$

де (x, y) – координати точок ідеального зображення; (x', y') – відповідні їм координати на його проєкції.

Поліноміальна апроксимуюча функція перетворення для знімка на рис. 4.5 має вигляд

$$x' = 0,1x + 0,9y + 0,002xy, \quad y' = 0,2x + 1,1y + 0,0022xy. \quad (4.4)$$

Уточнення координат і значень яскравості пікселів після перетворення

Після повороту або збільшення зображення часто виникає завдання уточнення нових координат пікселів (рис. 4.6) у цілочисельній системі

координат. Вона вирішується шляхом округлення координат пікселів до цілих значень і двомірної інтерполяції значень яскравості в невизначених пікселях. Найбільш простою є інтерполяція методом найближчого сусіда, де значення координат округляються до найближчих точок дискретного растра.



Рисунок 4.5 – Приклад білінійного геометричного перетворення знімка

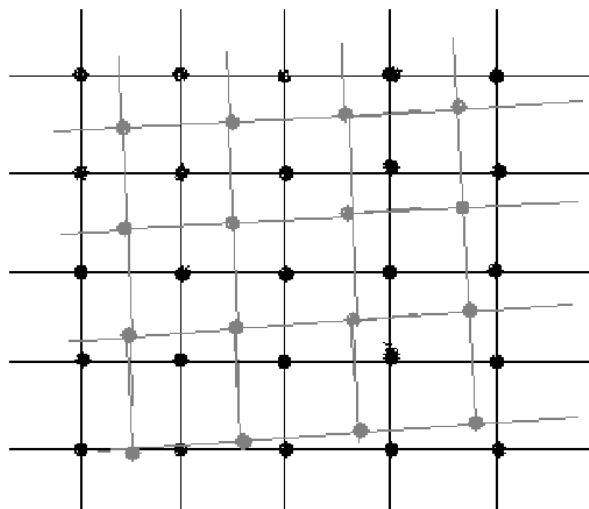


Рисунок 4.6 – При повороті координати пікселів (виділені сірим тоном) не збігаються з дискретною цілочисельною системою координат

Можливі два варіанти визначення яскравості пікселів при геометричних перетвореннях:

- прямий – для кожного пікселя обчислюються нові координати, вони округляються до найближчих цілих значень, потім невизначеним

пікселям нового зображення привласнюється яскравість найближчого після геометричного перетворення пікселя з неокругленими координатами;

- зворотний – для кожного пікселя перетвореного зображення координати піддаються зворотному перетворенню в систему координат вихідного зображення, координати округляються до найближчих цілих значень і береться яскравість пікселя з такими координатами.

Геометрична корекція зображень

Побудова мозаїки із зображень

Як можна використати геометричні перетворення, розглянемо на прикладі побудови мозаїки із двох зображень. Мозаїка, або зшивка, – це об'єднання двох або більше зображень у єдине зображення. Нехай два зображення отримані шляхом сканування вроздріб одного великого малюнка (рис. 4.7). Обов'язкова умова: на обох зображеннях частково присутні ті самі об'єкти, тобто зображення мають області перекриття.

Системи координат двох зображень будуть розрізнятися через різний ракурс зйомки, обертання камери й руху самого фотографованого об'єкта. Для зображень, представлених на рис. 4.7, освітлення, відстань до камери, оптична система однакові, тому основним завданням обробки таких зображень є приведення їх у загальну систему координат. Якщо зображення реєструвалися в різний час, необхідно врахувати зміни яскравості, що виникають через зміну висвітлення, сезонні й добові зміни, використання іншої оптичної системи.



Рисунок 4.7 – Приклад сканування документа вроздріб

Розглянемо завдання знаходження такого просторового перетворення, що дозволяє визначити пікселі обох зображень, представлених на рис. 4.7, у єдиній системі координат таким чином, щоб

точки, що відповідають однаковим об'єктам на двох зображеннях, збіглися. Як загальну систему координат можна використати систему лівого зображення, тоді потрібно знайти перетворення координат всіх пікселів другого зображення (x, y) у загальну систему координат (x', y') . Для спрощення завдання вважаємо, що в процесі реєстрації не відбулося скривлення прямих ліній, а лише афінні трансформації (поворот, зсув). Афінні перетворення є підмножиною поліноміальних перетворень першого порядку й описуються двома рівняннями:

$$\begin{aligned} x' &= a_0 + a_1x + a_2y, \\ y' &= b_0 + b_1x + b_2y. \end{aligned} \quad (4.5)$$

На обох зображеннях (рис. 4.8) можна вибрати інтерактивно або знайти за допомогою спеціальних методів пікселі, які відповідають однаковим об'єктам. Позначимо через (x_i, y_i) координати таких пікселів на правому зображенні в системі координат цього зображення, (x'_i, y'_i) – координати в системі координат лівого зображення. Ці координати відомі (точки визначені), коефіцієнти $a_0...a_2$ і $b_0...b_2$ невідомі. Для рішення завдання обчислення шести коефіцієнтів $a_0...a_2$ і $b_0...b_2$ афінного перетворення мінімальна кількість загальних точок дорівнює $n=3$. Якщо задати більше пікселів, то необхідно мінімізувати помилки їхнього завдання, які неминучі в силу неточного завдання й округлення координат до цілих значень.



Рисунок 4.8 – Приклад вибору пари відповідних пікселів

Таким чином, маємо шість невідомих, для їхнього пошуку потрібно три пари пікселів з відомими координатами до й після трансформації $(x_1,$

y_1); (x_2, y_2) ; (x_3, y_3) і (x_1', y_1') ; (x_2', y_2') ; (x_3', y_3') . Підставимо координати (x, y) у рівняння (4.5), одержимо три пари рівнянь із невідомими a_1 й b_1 :

$$\begin{aligned}x_1' &= a_0 + a_1x_1 + a_2y_1, & y_1' &= b_0 + b_1x_1 + b_2y_1, \\x_2' &= a_0 + a_1x_2 + a_2y_2, & y_2' &= b_0 + b_1x_2 + b_2y_2, \\x_3' &= a_0 + a_1x_3 + a_2y_3, & y_3' &= b_0 + b_1x_3 + b_2y_3,\end{aligned}$$

У матричній формі це виглядає в такий спосіб:

$$\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}, \quad \begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}.$$

Для обчислення коефіцієнтів a_i і b_i кожна частина матричного рівняння повинна бути помножена на зворотну матрицю. Наприклад, для a_i :

$$\begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}.$$

У матричній формі коефіцієнти a_i і b_i обчислюються за формулами

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}, \quad \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{bmatrix}^{-1} \begin{bmatrix} y_1' \\ y_2' \\ y_3' \end{bmatrix}.$$

Отриманий набір коефіцієнтів підставляється у вихідні рівняння, і нові координати обчислюються для кожного пікселя другого зображення. Залишилося визначити, яку частину другого зображення необхідно приєднати до першого.

До отриманого зображення можна приєднувати праворуч і знизу інші зображення. Результуюче зображення представлено на рис. 4.9.

При зйомці більших об'єктів можна їх фотографувати вроздріб, а потім виконати зшивку отриманих зображень. Приклад обробки панорамних фотографій наведений на рис. 4.10. У такий же спосіб була створена панорама Праги з 2600 знімків. У результаті отримане зображення розміром 260 000×130 000 пікселів. Якщо його роздрукувати з відрізненням 150 dpi, вийде фотографія довжиною близько 40 м.



Рисунок 4.9 – Приклад мозаїки, складеної із двох зображень



Рисунок 4.10 – Приклад зшивки фотографій з урахуванням геометричних перекручувань

Корекція проєктивних перекручувань
Проективне перетворення описується рівняннями

$$x' = \frac{a_{01}x + a_{10}y + a_{00}}{c_{01}x + c_{10}y + 1}, y' = \frac{b_{01}x + b_{10}y + b_{00}}{c_{01}x + c_{10}y + 1}$$

де x и y – точки ідеального зображення; x' й y' – точки проективно перекрученого зображення.

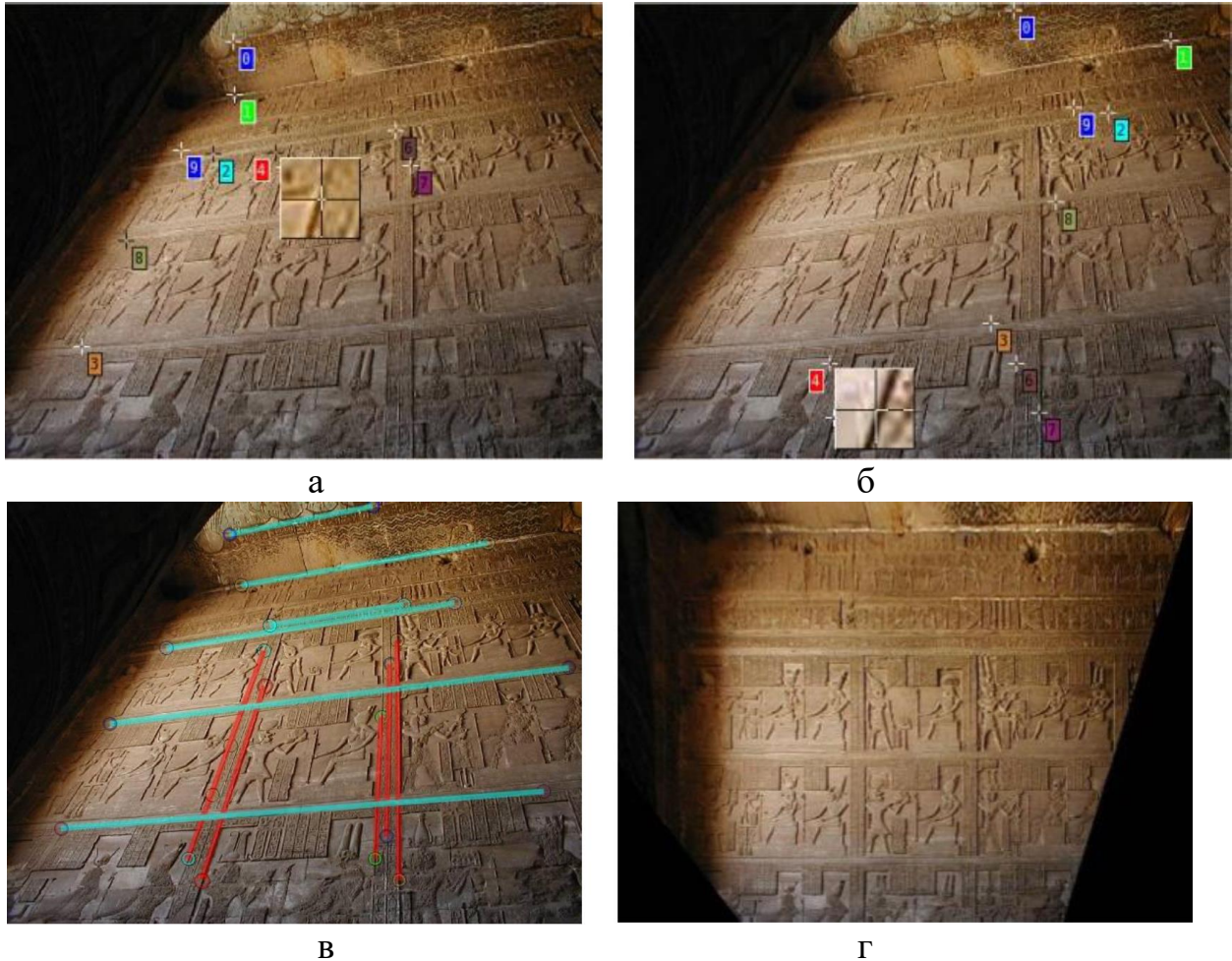


Рисунок 4.11 – Приклад перспективної корекції зображення: а) задані ліві кінці горизонтальних відрізків і верхні вертикальних; б) задані праві кінці горизонтальних відрізків і нижні вертикальних (кінці одного відрізка позначені однаковими цифрами); в) зазначені відрізки відображені; г) зображення відкоректоване

При корекції перспективних перекручувань вихідне зображення одне (рис. 4.11). Інтерактивно задаються кінці відрізків, які повинні бути вертикальними й горизонтальними на відкоректованому зображенні. Для визначення восьми невідомих коефіцієнтів мінімально необхідно задати вісім точок. За допомогою рішення системи лінійних рівнянь обчислюються невідомі параметри перетворення a_{ij} , b_{ij} й c_{ij} . Всі пікселі вихідного зображення трансформуються, і відзначені відрізки стають

вертикальними й горизонтальними.

Нелінійна геометрична корекція

Не завжди завдання геометричної корекції можна вирішити за допомогою лінійних перетворень. Нелінійні перетворення (другого й більше порядку) допомагають коректувати нелінійні перекручування зображення, які виникли в процесі його реєстрації.

Перетворення другого порядку можуть бути використані для перетворення даних з географічної системи координат (широта, довгота) у прямокутну, більших областей (для врахування кривизни Землі), для точної прив'язки перекручених по тій або іншій причині (через перекручування лінз камери, погано відсканованих матеріалів) даних і т.д. (рис. 4.12)

Спосіб рішення завдання трансформації аналогічний попередній і зводиться до знаходження коефіцієнтів системи рівнянь. Докладніше рішення цього завдання представлено в лабораторній роботі, присвяченій топографічній прив'язці супутникового знімка.



Рисунок 4.12 – Приклади нелінійних перекручувань

Корекція дісторсії

При фотографуванні на зображенні може виникнути перекручування, яке називається дісторсією. Дісторсія – це оптичне перекручування, що виражається в скривленні прямих ліній. Світлові промені проходять через центр лінзи й сходяться в точці, розташованій далі від лінзи, чим промені, які проходять через її краї (рис. 4.13).

Дісторсія не порушує різкість й яскравість зображення, але вносить перекручування в його форму. Прямі лінії зображуються кривими, крім тих, які лежать в одній площині з оптичною віссю. Наприклад, зображення квадрата, центр якого перетинає оптична вісь, має вигляд подушки (подушкоподібна дісторсія) при позитивній дісторсії в ширококутних об'єктивів при зйомці на максимальній фокусній відстані й вид бочки (бочкоподібна дісторсія) при негативній дісторсії при зйомці телеоб'єктивом на мінімальній фокусній відстані (рис. 4.14). Це найпоширеніші варіанти дісторсії оптичних систем фотоапаратів з осью

симетрією.

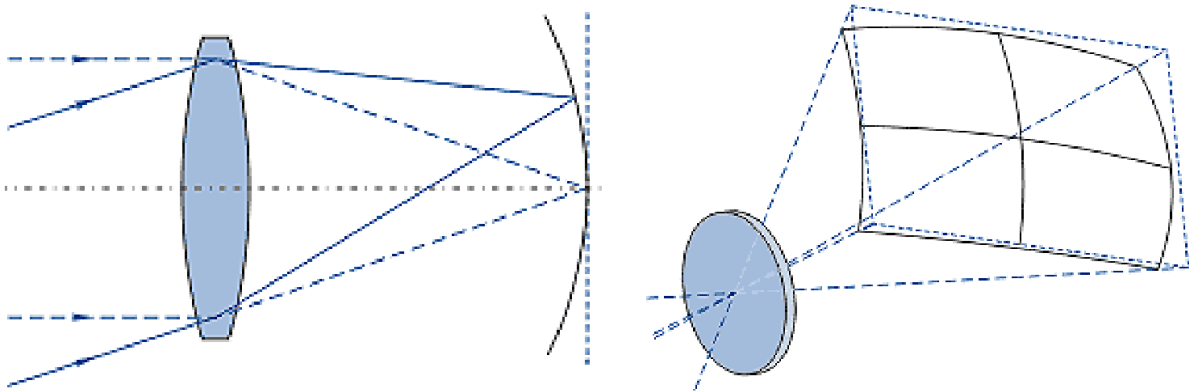


Рисунок 4.13 – Оптичні перекручування при реєстрації зображення породжують дісторсію

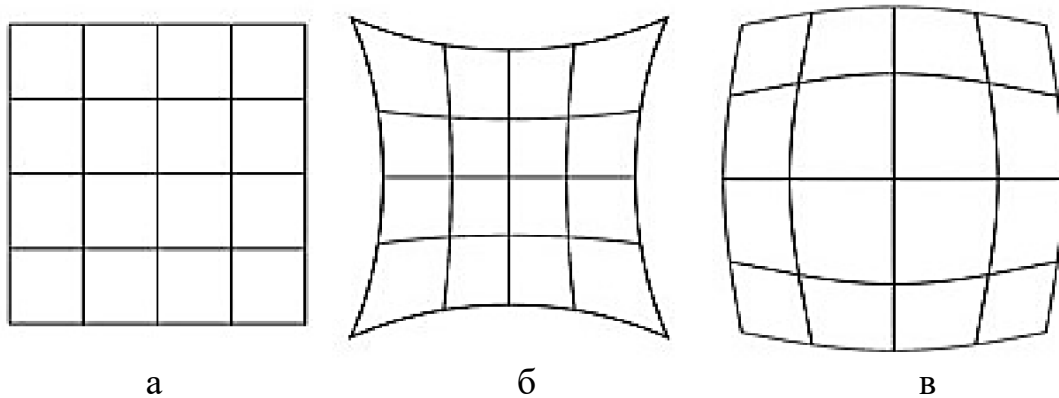


Рисунок 4.14 – Основні варіанти дісторсії фотооб'єктивів: а) дісторсія відсутня; б) подушкоподібна дісторсія; в) бочкоподібна дісторсія

Дісторсія об'єктива може бути скоректована оптично за рахунок особливої конструкції об'єктива, використання асферичних лінз або програмним шляхом. Як правило, така дісторсія має осьову симетрію, це значить, що горизонтальна й вертикальна осі, які проходять через центр зображення, не перекручені.

Якщо скривлення присутнє в цифровому зображенні, його можна виправити програмним методом, описаним нижче.

Будемо вважати, що центр системи координат перебуває в центрі зображення. Для осесиметричних оптичних систем дісторсія може бути тільки непарних порядків: третього, п'ятого, сьомого й т.д. Нехай $r = (x, y)$ – вектор, що задає дві координати в площині, розташованій перпендикулярно оптичній осі (координата Z уздовж осі сюди не входить). Для ідеального зображення всі промені, що вийшли із точки з координатами r і минули через оптичну систему, потраплять у точку

зображення з координатами \mathbf{R} , які визначаються за формулою $\mathbf{R}=b_0\mathbf{r}$, де b_0 – коефіцієнт лінійного збільшення. Для однієї й тієї ж оптичної системи дісторсія залежить від відстані до об'єкта, а, отже, і від коефіцієнта b_0 . У довгофокусних об'єктивів дісторсія менше, ніж у нормальних, у ширококутних – більше. Дісторсія також незначно залежить від довжини відбитої хвилі.

Якщо присутня дісторсія третього порядку, тоді промені, що вийшли із точки з координатами r , потраплять у точку зображення з координатами $\mathbf{R}=b_0\mathbf{r}+F_3r^2\cdot\mathbf{r}$, где r – довжина вектора \mathbf{r} ; F_3 – коефіцієнт дісторсії третього порядку, який вносить найбільший вклад у перекручування форми зображення. Якщо F_3 має той же знак, що й b_0 , виникає подушкоподібне перекручування, у протилежному випадку – бочкоподібна.

Для обчислення параметрів коригувального перетворення використовують зображення регулярної сітки і її скривлені зображення (рис. 4.15).

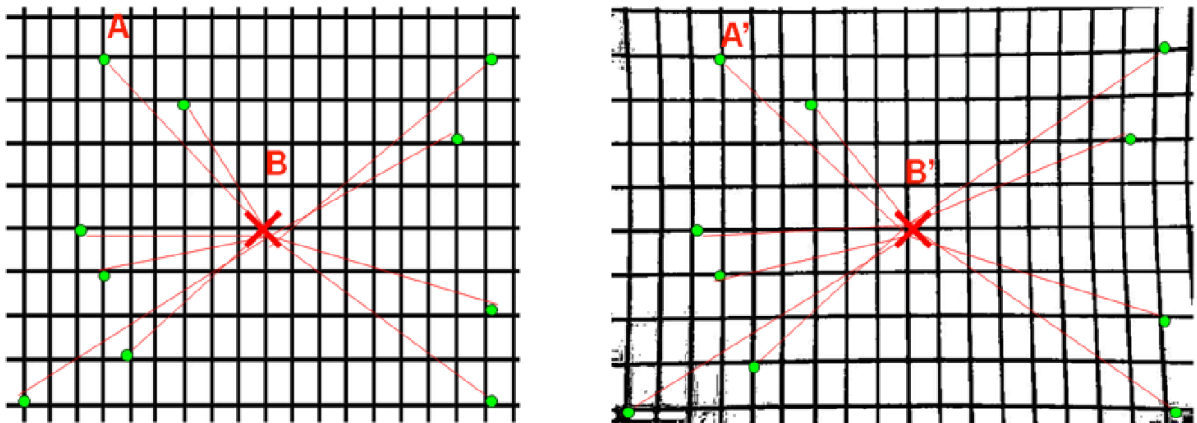


Рисунок 4.15 – Зображення регулярної сітки (ліворуч) і її знімок (праворуч), зроблений мобільним телефоном з парами відповідних точок А і В

На них вибираються пари відповідних точок; обчислюються вектори, що з'єднують ці точки з початком координат; отримані параметри підставляються в рівняння дісторсії й вирішується система лінійних рівнянь для визначення невідомих коефіцієнтів b_0 і F_3 .

Приклади виправлення дісторсія реальних зображень наведені на рис. 4.16.



Рисунок 4.16 – Приклади двох видів дісторсії й виправлене зображення

Контрольні запитання

1. Що таке цифрова обробка зображення ?
2. Яка мета обробки зображень ?
3. Що таке покращення зображення ?
4. Що таке відновлення зображення ?
5. Чим обумовлене виправлення зображень ?
6. В чому полягає структурне редагування зображення ?
7. Які існують рівні обробки зображень ?
8. Які існують класи методів обробки зображень в залежності від обчислювальної складності ?
9. В чому полягають геометричні перетворення зображень ?
10. Які існують види афінних перетворень ?
11. На які класи поділяються зображення ?

12. З чим пов'язане проєктивне перетворення зображень ?
13. Яким чином уточнюються координати пікселів після перетворення зображення ?
14. Яким чином уточняється яскравість пікселів після перетворення зображення ?
15. Як і для чого виконується мозаїка з зображень ?
16. Як здійснюється корекція проєктивних перекручувань ?
17. Що таке дісторсія зображення ?
18. Як здійснюється корекція дісторсії зображення ?

5 ПЕРЕТВОРЕННЯ ЗОБРАЖЕНЬ ТОЧКОВОГО ТИПУ

5.1 Логічні операції над зображеннями

Яскравість бінарних зображень задається числами 0 й 1. Ці зображення можна інтерпретувати як множини чорних пікселів на білому тлі. До них можна застосовувати операції теорії множин AND, OR, NOT, XOR (рис. 5.1). Нехай $L1$ й $L2$ – два бінарних зображення рівних розмірів. Операції застосовуються до пікселів з однаковими координатами. У табл. 5.1 показано, як змінюються значення яскравості при застосуванні цих логічних операцій. Операція NOT створює негативне зображення, її можна описати функцією $L=1 - L1$.

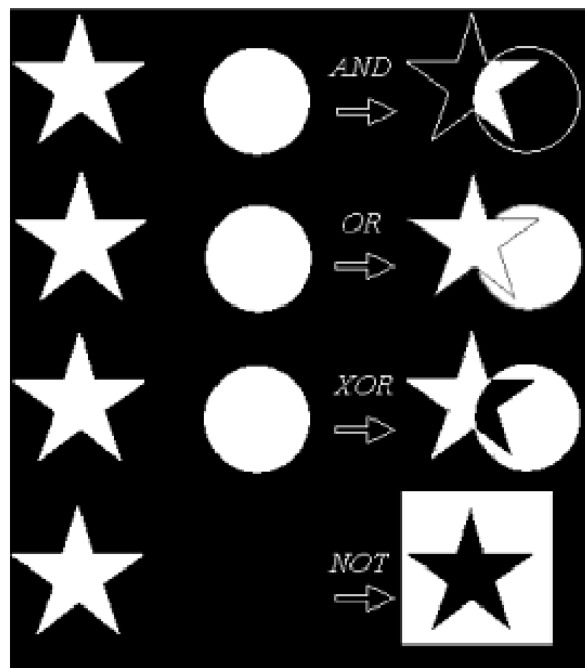


Рисунок 5.1 – Приклади логічних операцій над бінарними зображеннями

Таблиця 5.1 – Результати логічних операцій

L ₁	L ₂	L ₁ AND L ₂	L ₁ OR L ₂	NOT L ₁	L ₁ XORL ₂
0	0	0	0	1	0
0	1	0	1	1	1
1	0	0	1	0	1
1	1	1	1	0	0

5.2 Лінійні перетворення яскравості напівтонових зображень

Яскравісними перетвореннями зображення $I(i,j)$ називаються перетворення двовірних функцій яскравості, описувані формулою $I'(i, j) = f(I(i, j))$. Перетворення яскравості зображень відносять до точкового типу, якщо значення яскравості пікселя після перетворення залежить від яскравості однієї точки (пікселя) вихідного зображення й не залежить від її місця розташування. Нехай $x = I(i, j)$ – функція яскравості вихідного зображення, $v = J(i, j)$ – функція яскравості перетвореного зображення. Поелементна обробка означає, що зміну яскравості можна описати функцією $v = f(x)$ незалежно від координат пікселя.

Препарування зображень

За допомогою яскравісного зрізу зображення можна виділити області зображення з яскравістю з певного інтервалу. При цьому іншим областям можна привласнити чорні кольори й одержати бінарне зображення (рис. 5.2, а) або залишити незмінними (рис. 5.2, б). Переміщаючи виділений інтервал по шкалі яскравості й змінюючи його ширину, можна детально досліджувати зміст зображення. При цьому, пікселі, які мають яскравість, що не потрапила в цей інтервал, відображаються чорними, з яскравістю 0 (рис. 5.2, а), або лінійно контрастуються (рис. 5.2, б). Такі перетворення використовуються, якщо інтервали яскравості можна визначити (обчислити або задати інтерактивно).

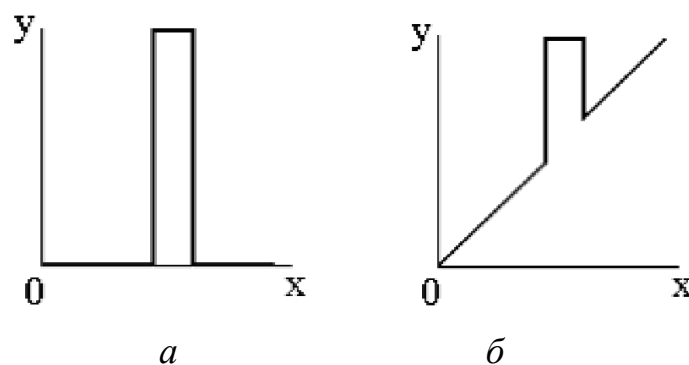


Рисунок 5.2 – Два варіанти графічного завдання функції яскравісного зрізу

Аналітично функція зрізу для першого варіанта (рис. 5.2, а) задається в такий спосіб (замість значення яскравості 1 використовується 255 для відображення на екрані):

$$y=255, \text{ якщо } T_1 < x < T_2, \text{ інакше } y=0.$$

Пилкоподібне контрастування – це лінійне контрастування, застосовуване до одного або декількох діапазонів яскравості (рис. 5.3). Воно також дозволяє підвищити детальність зображення в обраному інтервалі яскравості.

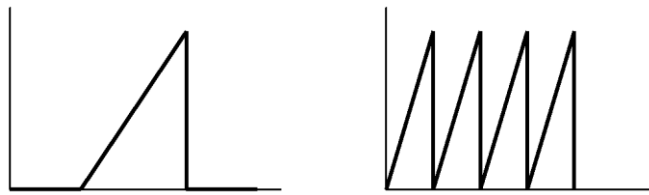


Рисунок 5.3 – Пилкоподібні варіанти яскравісного зрізу

Бінаризація зображень

Найпростішим методом препарування зображень є бінаризація. Вона перетворює напівтонове зображення в бінарне (чорно-біле). Перетворення має єдиний параметр – поріг, щодо якого яскравість пікселів міняється на чорну або білу. Функція бінаризації із глобальним (тобто єдиним для всіх пікселів) порогом має вигляд

$$y=255, \text{ якщо } x > T, \text{ інакше } y=0,$$

а її графічне подання показане на рис. 5.4. Це перетворення використовують при виділенні темних об'єктів на світлому тлі або навпаки (рис. 5.5).

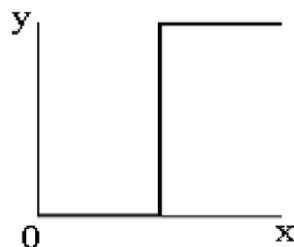


Рисунок 5.4 – Функція бінаризації із глобальним порогом



Рисунок 5.5 – Приклад бінаризації зображення

Опишемо *три варіанти* обчислення порога.

Найбільш простий з них полягає в знаходженні мінімального g_{min} і максимального g_{max} значень яскравості зображення й обчислення середнього значення між ними:

$$t = (g_{max} - g_{min}) / 2$$

Метод Отсу використовує гістограму розподілу значень яскравості зображення й шукає оптимальний поріг t , який розділяє два класи пікселів (темні й світлі) і зменшує внутрішньокласову дисперсію, що визначається як зважена сума дисперсій двох класів:

$$\sigma_w^2(t) = \omega_1(t)\sigma_1^2 + \omega_2(t)\sigma_2^2$$

де ваги ω_i – це ймовірності двох класів, розділених порогом t ; σ_i^2 – дисперсія цих класів. Мінімізація дисперсії усередині класу рівносильна максимізації дисперсії між класами, що виражається ймовірністю ω_i середнього арифметичного i -го класу μ_i :

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = \omega_1(t) \omega_2(t) [\mu_1(t) - \mu_2(t)]^2$$

де σ^2 – дисперсія всього зображення.

Алгоритм обчислення порога бінаризації методом Отсу.

Крок 1. Обчислити гістограму яскравостей й ймовірність для кожного рівня інтенсивності $p_i = n_i / N$, де N – число пікселів у зображенні; n_i – число пікселів з рівнем яскравості i . Діапазон яскравостей ділиться на два класи порогом k , де k – ціле значення від 0 до L (звичайно $L = 255$).

Крок 2. Обчислити початкові значення для $\omega_i(0)$ і $\mu_i(0)$:

$$\omega_0(k) = \sum_{i=1}^k p_i \qquad \omega_1(k) = \sum_{i=k+1}^L p_i = 1 - \omega_0(k)$$

$$\mu_0(k) = \sum_{i=1}^k \frac{ip_i}{\omega_0} \quad \mu_1(k) = \sum_{i=k+1}^L \frac{ip_i}{\omega_1}$$

Крок 3. Для кожного значення порога від $t = 1$ до максимальної яскравості (рівної 255):

- обновити ω_i й μ_i ;
- обчислити $\sigma_b^2(t)$;
- якщо $\sigma_b^2(t)$ більше, ніж наявне, запам'ятати σ_b і значення порога t .

Крок 4. Оптимальний поріг t відповідає максимуму $\sigma_b^2(t)$.

Алгоритм обчислення порога на основі градієнта яскравості.

Крок 1. Обчислити модуль градієнта яскравості для кожного пікселя (m, n) зображення f :

$$G(m,n) = \max\{|G_m(m,n)|, |G_n(m,n)|\},$$

де $G_m(m, n) = f(m+1, n) - f(m-1, n)$, $G_n(m, n) = f(m, n+1) - f(m, n-1)$.

Крок 2. Обчислити значення порога t :

$$t = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) G(m, n)}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} G(m, n)}$$

Лінійне розтягання контрасту зображень

Призначено для поліпшення узгодження динамічного діапазону зображення й екрана, на якому виконується візуалізація. Якщо для цифрового подання яскравості кожного пікселя зображення приділяється 1 байт (8 біт), то вхідне або вихідне зображення може приймати одне з 256 значень у діапазоні від 0 до 255. Нехай мінімальна й максимальна яскравості вихідного зображення рівні x_{min} й x_{max} відповідно. Якщо $x_{min} \gg 0$ й $x_{max} \ll 255$, тобто динамічний діапазон вузький, то зображення виглядає сірим і малокоонтрастним.

При лінійному розтяганні контрасту використовується перетворення яскравості типу

$$y = ax + b,$$

де a й b визначаються бажаними значеннями мініимальної й максимальної яскравості результуючого зображення (звичайно $y_{min} = 0$ для x_{min} й $y_{max} = 255$ для x_{max}). З урахуванням цього перетворення яскравості приймає вид

$$y = 255 \cdot (x - x_{min}) / (x_{max} - x_{min}) \quad (5.1)$$

Графічне подання цієї функції й приклад лінійного розтягання

контрасту представлені на рис. 5.6 й 5.7 відповідно.

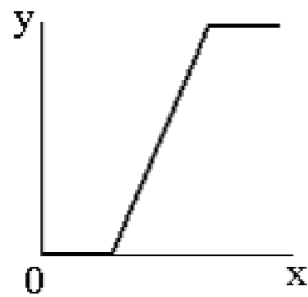


Рисунок 5.6 – Графічне подання функції лінійного контрастування зображення



Рисунок 5.7 – Світле зображення (а) з діапазоном яскравості $x_{min} = 180$, $x_{max} = 240$ перетвориться за допомогою лінійного розтягання контрасту (б)

Частковим випадком цієї лінійної зміни яскравості є перетворення зображення в негативне подання: $v = 255 - x$.

Контрастування можна виконати кусочно-лінійним способом. Нехай задані $x_1 < x_2 < \dots < x_n$ – границі діапазонів зміни яскравості (рис. 5.8).

Тоді кусочно-лінійна функція зміни яскравості має вигляд

$$y = \begin{cases} a_0x + b_0, & \text{при } x < x_1 \\ a_1x + b_1, & \text{при } x_1 < x < x_2 \\ \dots & \dots \\ a_nx + b_n, & \text{при } x_n < x \end{cases}$$

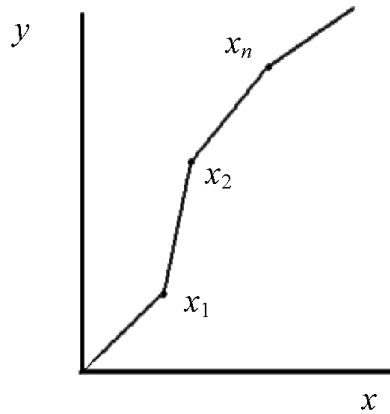


Рисунок 5.8 – Графічне подання функції кусочно-лінійного контрастування зображення

5.3 Арифметичні операції над зображеннями

Арифметичні перетворення зображень відносять до точкового типу.
Множення яскравості зображення на константу:

$$v(i, j) = ax(i, j),$$

де $a > 0$ – константа. Якщо $v(i, j) > 255$, замінити на $v = 255$.

Збільшення яскравості на константу c :

$$z(i, j) = x(i, j) + c.$$

Додавання або віднімання рівних за розміром зображень (рис. 5.9):

$$z(i, j) = ax(i, j) \pm (1 - a) v(i, j),$$

де x, y, z – цифрові зображення одного розміру, $0 < a < 1$ – константа.

Просторове маскуваня зображення (рис. 5.10):

$$y(i, j) = a(i, j) x(i, j)$$

де $a(i, j)$ – матриця (маска) з нулів й одиниць того ж розміру, що зображення x . Операція являє собою поелементний добуток двох матриць.



а)



б)



в)

Рисунок 5.9 – Приклад додавання двох зображень:
а), б) – вихідні зображення; в) результат додавання



Рисунок 5.10 – Приклад маскування

5.4 Нелінійна корекція яскравості зображень

Гама-корекція – це операція корекції яскравості зображення I , використовується для однакового подання зображень на моніторах різних конструкцій з нелінійною яскравісною характеристикою. Звичайно використовується ступенева функція виду

$$I = c \cdot I^\gamma.$$

Історично введення γ -корекції було обумовлено тим, що в електронно-променевої трубки залежність між кількістю випущених фотонів і напругою на катоді близька до ступеневої формули. У результаті це ввійшло в стандарт і для сучасних рідкокристалічних моніторів, де залежність між напругою і яскравістю більше складна. Стандартне значення параметра γ для відеозображень NTSC дорівнює 2,2. Для більшості моніторів значення становить від 1,8 до 2,4.

При обробці цифрових зображень із діапазоном яскравості від 0 до 255 це перетворення виконується за формулою

$$I = 255 \cdot (I/255)^\gamma.$$

Якщо зображення виглядає темним або бляклим, γ -корекція використовується для корекції цих недоліків.

Аналогічно виконується *логарифмічна зміна яскравості* (c – задана константа, $b \leq 1$):

$$I = c \log(I+b).$$

Перетворення, назване *соляризацією зображення*, змінює яскравість за допомогою параболічної функції:

$$y = k \cdot x(x_{max} - x),$$

де x_{max} – максимальне значення яскравості вихідного зображення (звичайно 255), а k – константа, яка управляє динамічним діапазоном перетвореного зображення. Функція перетворення є квадратичною параболою, її графік при $k = 1/64$ наведений на рис. 5.11.

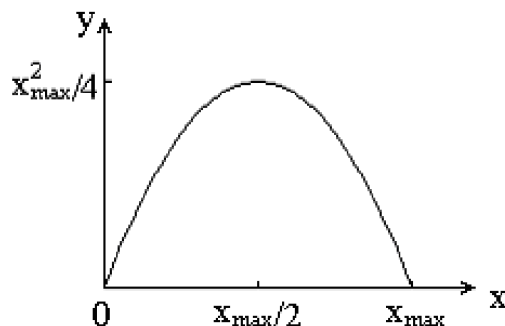


Рисунок 5.11 – Функція зміни яскравості при соляризації

Зміст перетворення полягає в тому, що ділянки вихідного

зображення з яскравістю, близькою до білих кольорів, після обробки приймають значення яскравості ближче до чорного. При цьому темні ділянки залишаються темними. Значення, близькі до білого, здобувають ділянки, що споконвічно мали середні рівні яскравості (рис. 5.12).



Рисунок 5.12 – Вихідне зображення (а) і результат його соляризації (б)

Як видно на рис. 5.12, перетворення підвищує чіткість темних і світлих областей: поліпшені зображення очей, підвищений контраст на переході «обличчя-волосся».

Алгоритм корекції яскравості шляхом приведення середньої яскравості до заданого значення.

Нехай L_{ij} – яскравість пікселя з координатами i, j .

Крок 1. Задати необхідне значення середньої яскравості K . $L = L/255$, $K = K/255$.

Крок 2. Обчислити мінімальне L_{min} , максимальне L_{max} і середнє арифметичне значення яскравості вихідного зображення (Av):

$$Av = \text{mean}(L) = \frac{1}{N} \sum_{i,j} L_{ij} .$$

Крок 3. Визначити параметри перетворення значень яскравості так, щоб значення середньої яскравості стало рівним K .

$$K = \frac{1}{N} \sum_{i,j} F(L_{ij}) .$$

Для цього змінити значення яскравості всіх пікселів зображення, наприклад, нелінійно

$$F(L_{ij}) = \left(\frac{L_{ij} - L_{min}}{L_{max} - L_{min}} \right)^\alpha \quad (5.2)$$

одержимо

$$K = \frac{1}{N} \sum_{i,j} F(L_{ij}) = \frac{1}{N} \sum_{i,j} \left(\frac{L_{ij} - L_{min}}{L_{max} - L_{min}} \right)^\alpha .$$

Підібрати значення α так, щоб середньоарифметичне значення яскравості пікселів зображення стало дорівнює K .

Крок 4. Виконати перетворення яскравості всіх пікселів вихідного зображення по формулі (5.2), додатково помноживши на 255.

Приклади перетворення зображення до середніх значень яскравості наведені на рис. 5.13.

Табличний метод зміни яскравості. У перетвореннях точкового типу яскравість пікселя змінюється відповідно до деякої формули, однакової для всіх пікселів зображення. Нехай розмір зображення дорівнює $N \times N$, де $N \gg 256$. При нелінійних змінах яскравості значень буде потрібно N^2 разів обчислити нові значення пікселів.

Від цих недоліків вільний табличний метод заміни яскравості (LookUpTable – LUT). Його суть полягає в тому, що попередньо обчислюються значення функції зміни яскравості $f(x)$ для всіх x від 0 до 255 і створюється таблиця нових значень $y = f(x)$ (табл. 5.2). Потім яскравість змінюється відповідно до цієї таблиці, при цьому вихідна яскравість x використовується як індекс в отриманій таблиці, а y – як нове значення яскравості.

Таблиця 5.2 – Табличне завдання функції $y = f(x)$

x	x_1	x_2	...	x_{n-1}	x_n
$y = f(x)$	y_1	y_2	...	y_{n-1}	y_n

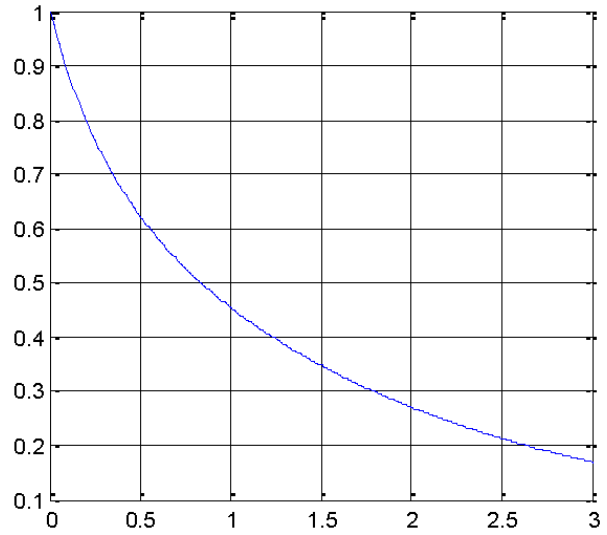
Якщо вихідне зображення має діапазон яскравості $[0, 255]$, то таблиця LUT складається з 256 елементів. Наприклад, для перетворення соляризації по формулі $y = 4x(255-x)/255$ таблиця описується одновимірним масивом:

0,4,8,12,16,20,23,27, ..., 198,199,201,203, ..., 254,255,255,255, ..., 12,8,4,0.

Нові значення яскравості при такому підході обчислюються 256 разів при формуванні таблиці. Якщо обчислювати її по формулі для кожного пікселя, то буде потрібно N^2 обчислень, де N може дорівнювати декільком тисячам. Тому таблична зміна яскравості є дуже швидкою.



а



б



в



г

Рисунок 5.13 – Перетворення середньої яскравості до заданого значення: а) вихідне зображення із середньою яскравістю 119; б) графік функції $K = f(a)$ при $a = 1,85$; в) перетворене зображення із середньою яскравістю 177; г) перетворене зображення із середньою яскравістю 77

Інтерактивна зміна функції яскравості зображення. Розглянемо метод перетворення яскравості, аналогічний реалізованому в графічному редакторі Photoshop (функція «криві»). Споконвічно графік зміни яскравості описується тотожною функцією $y=x$, де x – яскравість вихідного, а y – результуючого зображення (рис. 5.14). Значення x й y змінюються в діапазоні від 0 до 255.

Інтерактивно змінюємо положення деякої точки цього графіка, наприклад (x_1, y_1) , де $x_1 = y_1$, на нове з координатами (x_{11}, y_{11}) , при цьому $x_{11} \neq y_{11}$ (рис. 5.15).

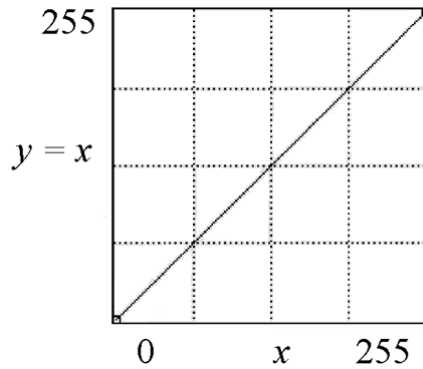


Рисунок 5.14 – Вихідний графік функції зміни яскравості

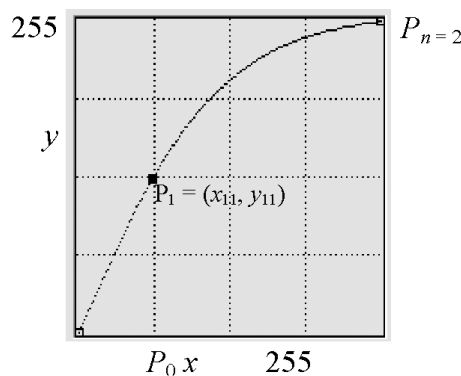


Рисунок 5.15 – Параболічна функція зміни яскравості

Крива, що задає функцію зміни яскравості, проходить через точки P_0 , P_1 , P_n (рис. 5.15), де $n = 2$, $P_0 = (0,0)$, $P_1 = (x_{11}, y_{11})$, $P_n = (255, 255)$. При $n=2$ це крива другого порядку (у загальному випадку парабола).

Можна змінити форму графіка, задавши ще одну точку $P_2=(x_{22}, y_{22})$ між точками P_0 і P_n (рис. 5.16). У цьому випадку крива повинна проходити через кінцеві $P_0 = (0,0)$, $P_n = (255, 255)$ і задані точки P_1, P_2 , при цьому $n=3$. Щоб обчислити параметри такої функції, крива описується поліномом ступеня на одиницю більше числа заданих (проміжних) точок. У даному прикладі – ступеня $n = 3$:

$$a_0 x^n + a_1 x^{n-1} + \dots + a_{n-1} x + a_n = y. \quad (5.3)$$

Оскільки $P_0 = (0,0)$, тобто при $x = 0$ з рівняння (3) маємо $v = 0$, то $a_n=0$. Інші a_i невідомі, їхня кількість дорівнює n . Щоб їх знайти, у рівняння (5.3) потрібно підставити координати точок $P_i = (x_i, y_i)$. Так, $P_3 = (255, 255)$ означає, що при $x=255$ маємо $v=255$. Одержимо систему з n лінійних рівнянь із невідомими a_i .

$$\begin{cases} a_0x_1^n + a_1x_1^{n-1} + \dots + a_{n-1}x_1 = y_1 \\ a_0x_2^n + a_1x_2^{n-1} + \dots + a_{n-1}x_2 = y_2 \\ \dots \\ a_0x_{n-1}^n + a_1x_{n-1}^{n-1} + \dots + a_{n-1}x_{n-1} = y_{n-1} \\ a_0255^n + a_1255^{n-1} + \dots + a_{n-1}255 = 255 \end{cases} \quad (5.4)$$

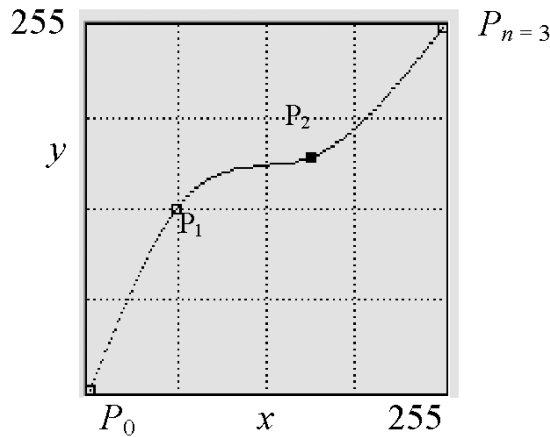


Рисунок 5.16 – Функція зміни яскравості третього порядку

У матричному виді система (5.4) записується як

$$X \cdot A = Y,$$

де $Y = (y_1, y_2, \dots, y_{n-1}, 255)$ – вектор-стовпець констант; $A = (a_0, a_1, a_2, \dots, a_{n-1})$ – вектор-стовпець невідомих; X – квадратна матриця констант,

$$X = \begin{bmatrix} x_1^n & x_1^{n-1} & \dots & x_1 \\ x_2^n & x_2^{n-1} & \dots & x_2 \\ \dots & \dots & \dots & \dots \\ x_{n-1}^n & x_{n-1}^{n-1} & \dots & x_{n-1} \\ 255^n & 255^{n-1} & \dots & 255 \end{bmatrix}$$

Система має рішення, якщо її визначник не дорівнює нулю. Ця умова виконується, якщо всі точки P_i мають різні координати. Тоді невідомі a_i знаходимо, обчисливши зворотну до X матрицю X^{-1} :

$$A = X^{-1} \cdot Y.$$

Обчисливши коефіцієнти a_i , підставляємо їх у рівняння (5.3) і будуємо таблицю LUT для всіх $0 \leq x \leq 255$, одержуємо нові значення яскравості v . Якщо $y < 0$, замінюємо на 0, якщо $y > 255$, замінюємо на 255.

Таблиця LUT готова, її графічне подання показано на рис. 5.16. Відповідно до отриманої таблиці заміняємо всі значення яскравості у вихідному зображенні в циклі за рядками і стовпцями.

5.5 Перетворення яскравості на базі гистограми зображення

Гістограма – це функція розподілу яскравостей напівтонового зображення. Нехай L – максимально можливе значення яскравості (звичайно 255). Обчислити гістограму означає побудувати масив H розміром в 256 елементів, значення яких дорівнюють кількості пікселів з яскравістю, рівної номеру елемента $H(i)$, тобто i . Обчислення виконуються проходом по всім пікселям зображення й додаванням одиниці до значення відповідного елемента масиву H .

У гістограмі типового зображення яскравість пікселів розподілена нерівномірно (рис. 5.17). У результаті часто на темних ділянках зображень деталі виявляються погано помітними.

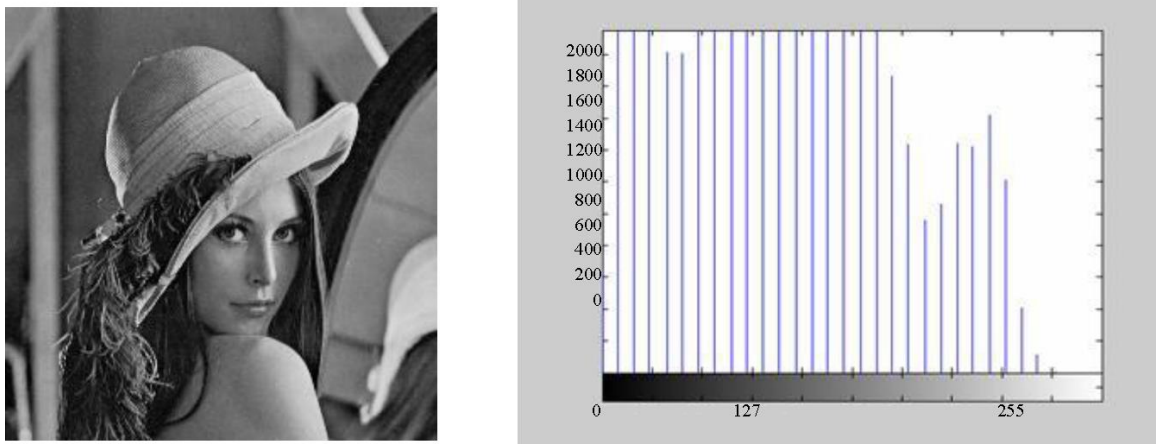


Рисунок 5.17 – Напівтонове зображення і його гістограма

Розроблено ряд методів поліпшення зображень на базі зміни гістограми. Найбільш простим з них є еквалізація (лінійне вирівнювання) гістограми. Це перетворення прагне змінити розподіл яскравостей зображення, роблячи графік гістограми максимально близьким до горизонтальної прямої (рис. 5.18).

Алгоритм еквалізації (лінійного вирівнювання) гістограми.

Крок 1. Обчислити гістограму вихідного зображення $f(x, y)$, записати її в масив H . Обчислити число пікселів N .

Крок 2. Масив H нормалізувати так, щоб сума всіх елементів дорівнювала 255: $H(j) = 255 \cdot H(j) / N$.

Крок 3. Обчислити кумулятивну гістограму, яка підсумовує розподіл яскравості від 0 до i : $Sum(i) = \sum_0^i H(i)$.

Крок 4. Обчислити нові значення яскравості кожного пікселя з координатами (x, y) за формулою: $g(x, y) = \text{Sum}(f(x, y))$.

На рис. 5.18 показаний результат вирівнювання гистограми зображення. Якщо зображення має багато темних ділянок і мале число ділянок з високою яскравістю або навпаки, вирівнювання гистограми коректує яскравість зображення, вирівнюючи інтегральні площі ділянок з різними яскравостями.

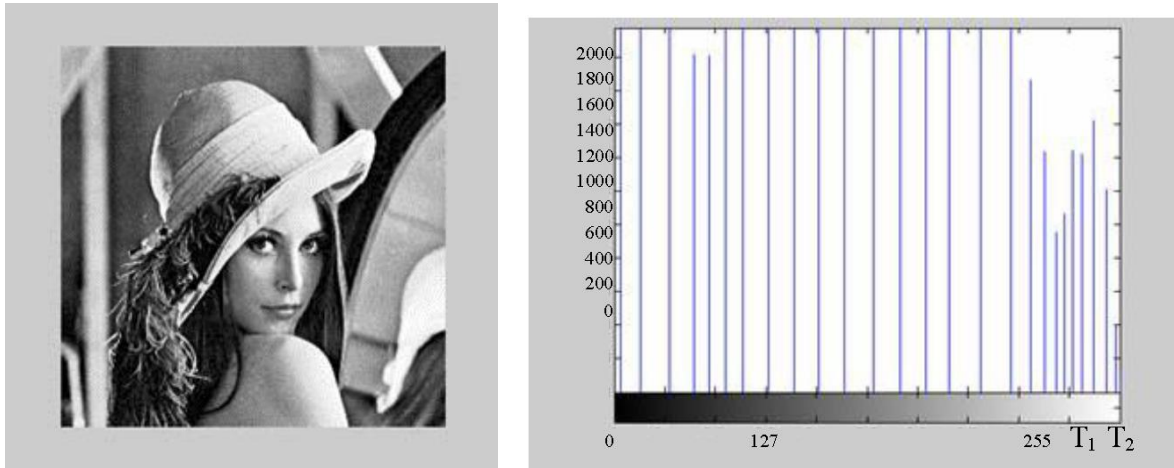


Рисунок 5.18 – Приклад вирівнювання гистограми зображення, представленого на рис. 5.17: перетворене зображення і його гистограма

Перетворення гистограми можуть бути й нелінійними, наприклад, наближені до кривої будь-якої форми (рис. 5.19), отримані за наведеним нижче алгоритмом.

Алгоритм нелінійного перетворення гистограми.

Крок 1. Для зображення X обчислити гистограму h_x вихідного зображення і його кумулятивну гистограму H_x , тоді функція вирівнювання цієї гистограми має вигляд

$$H_x[j] = \sum_{i=0}^j h_x[i]$$

Крок 2. Визначити бажану гистограму h_z й її кумулятивну гистограму H_z , тоді функція вирівнювання цієї гистограми має вигляд

$$H_z[j] = \sum_{i=0}^j h_z[i]$$

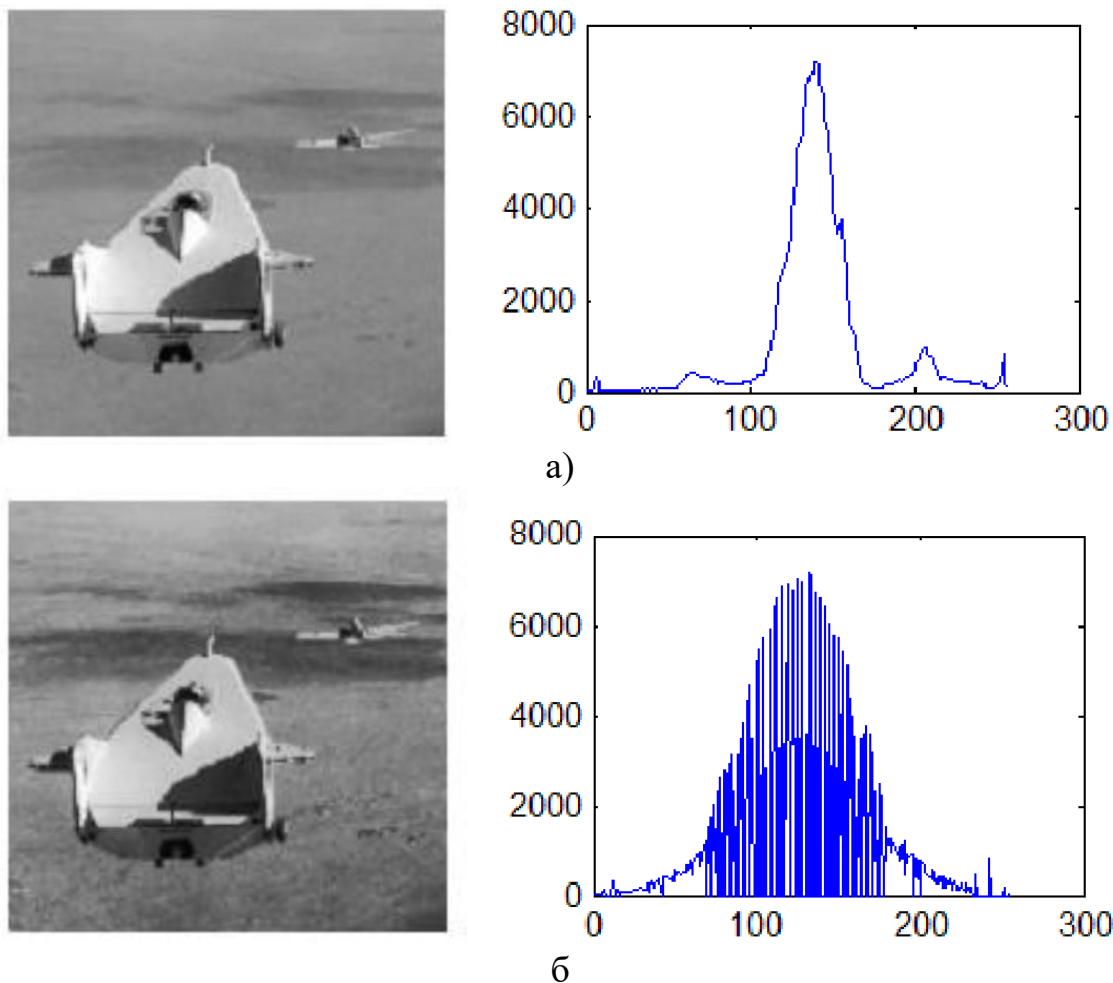


Рисунок 5.19 – Приведення гістограми до розподілу яскравостей за нормальним законом (функція Гауса): а) вихідне зображення і його гістограма; б) результуюче зображення і його гістограма

Крок 3. Побудувати таблицю LUT-перетворення яскравостей. У циклі по i для всіх $0 \leq i \leq 255$ виконати

$$j = i$$

якщо $H_x[i] \leq H_z[j]$, тоді $LUT[i] = j$;

інакше, поки $H_x[i] > H_z[j]$, збільшити $j = j + 1$;

якщо $H_x[i] - H_z[j] > H_x[i] - H_z[j-1]$, зменшити $j = j - 1$;

інакше, $LUT[i] = j$.

Визначити бажану гістограму можна різними способами, наприклад обчисливши гістограму зображення-зразка або задавши h_z у вигляді функції. Якщо функція має вигляд $h_z = 1/256$ для всіх i , результатом перетворення буде лінійне вирівнювання гістограми.

5.6 Локально-адаптивна обробка зображень

Середню яскравість зображення можна скорегувати глобально: збільшивши її, затемнені ділянки зображення можна зробити більше

світлими, а зменшивши – світлі ділянки зображення затемнити. Однак якщо на зображенні присутні й темні, і світлі області, глобальна корекція погіршить візуальне сприйняття зображення. Статистичні характеристики, необхідні для роботи алгоритмів обробки зображень, оцінюються по самих зображеннях. Якщо характеристики не змінюються по всьому зображенню, можна застосовувати алгоритми обробки з фіксованими параметрами, тобто глобальну обробку. Але в багатьох випадках яскравість, контраст, освітлення, тло не є однорідними. Тому глобальні методи обробки не дають бажаний результат. У таких випадках використовуються адаптивні, тобто перетворення, які підстроюють під локальні характеристики зображення.

Найпростіший підхід до побудови адаптивних алгоритмів полягає в розбивці зображення на невеликі непересічні фрагменти, у них обчислюються характеристики й кожен фрагмент обробляється як окреме зображення. Перевагою такого підходу є простота, недоліком – погане стикування частин об'єктів на границях оброблених фрагментів, скачки яскравості.

Для зниження небажаних ефектів на краях оцінки локальних характеристик обчислюють у пересічних фрагментах. У граничному випадку ці оцінки обчислюються й використовуються для корекції єдиного пікселя в центрі фрагмента. Це обробка методом «ковзного вікна», центр якого проходить всі пікселі вихідного зображення. Вона ставиться до перетворень локального типу.

Розглянемо особливості адаптивних алгоритмів стосовно до перетворень точкового типу.

Адаптивне підвищення контрасту.

Воно аналогічно лінійному контрастуванню, але коефіцієнти міняються залежно від значень пікселів, що потрапили в ковзне вікно:

$$g(i, j) = a(i, j) f(i, j) + \mu - \mu_f(i, j) a(i, j),$$

де $a(i, j) = \sigma/\sigma_f(i, j)$, $\mu_f(i, j)$, $\sigma_f(i, j)$ – локальне середнє й дисперсія; μ й σ – глобальне середнє й дисперсія.

Якщо у вікні зміна яскравості мала, $\sigma_f(i, j)$ теж мало, тому діапазон значень яскравості більше розтягується усередині вікна. По цій же причині розтягується діапазон значень у темних ділянках зображення (рис. 5.20).

Ще один варіант адаптивного контрастування вихідного зображення L_{in} при обчисленні інтенсивності нового зображення L_{out} здійснюється наступною операцією для кожного пікселя (i, j) :

$$L_{out}(i, j) = k \frac{\bar{L}}{\sigma(i, j)} [L_{in}(i, j) - \bar{L}_{loc}(i, j)] + \bar{L}_{loc}, \quad (5.5)$$

де \bar{L} – середнє значення яскравості всього зображення L_{in} ;
 $\sigma(i, j)$ – середньоквадратичне відхилення яскравості в локальній околиці зображення в пікселєві з координатами (i, j) ;
 $\bar{L}_{loc}(i, j)$ – середнє значення яскравості в околиці із центром у пікселєві (i, j) ;
 k – деяка константа, $0 < k < 1$.

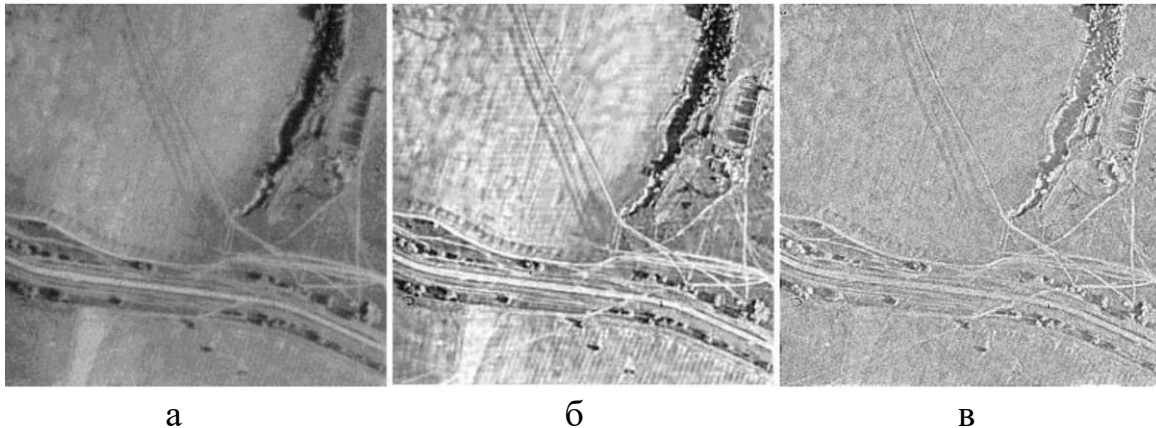
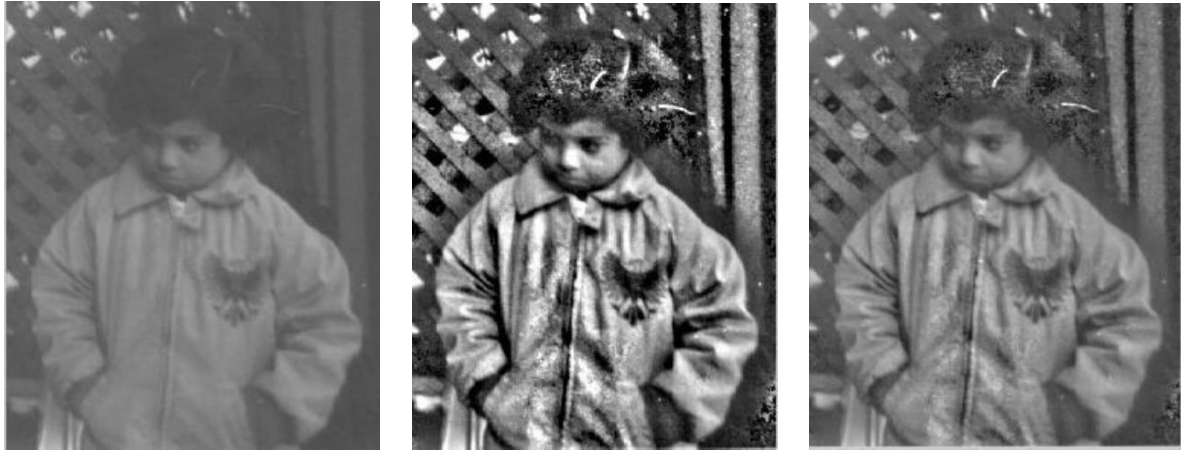


Рисунок 5.20 – Вихідне зображення (а), результати локального (б) і глобального (в) контрастування вихідного зображення

Відзначимо, що значення параметрів σ і L_{loc} – залежать від розмірів і форми околиці пікселя, що робить цей метод адаптивним. Локальні зміни збільшуються за рахунок множення різниці між значенням яскравості пікселя вихідного зображення $L_{in}(i, j)$ і локальним середнім $\bar{L}_{loc}(i, j)$ на $k \frac{\bar{L}}{\sigma(i, j)}$.

Середньоквадратичне відхилення буде менше в малоконтрастних околицях і більше на околицях з більш високим контрастом, тому ділянки з низьким контрастом будуть мати більше посилення, чим ділянки з більшим контрастом. Доцільно обмежувати діапазон значень множника $k \frac{\bar{L}}{\sigma(i, j)}$ щоб уникнути більших відхилень інтенсивностей на окремих ділянках.

На рис. 5.21 представлені результати адаптивного підвищення контрасту за формулою (5.5). Зменшення розміру локального вікна m приводить до збільшення детальності обробки. Зменшення коефіцієнта k приводить до усунення різких перепадів на зображенні й зниженню його контрастності. Таким чином, використовуючи різні значення параметрів m й k , можна управляти рівнем контрастності й детальності обробки зображень.



а

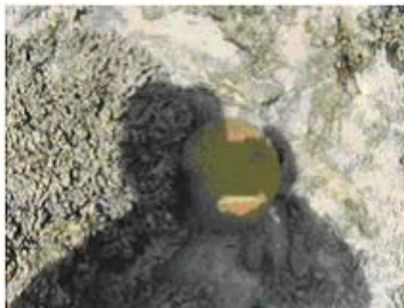
б

в

Рисунок 5.21 – Результати обробки зображення, отримані з використанням формули (5.5): а) вихідне зображення; б) результуюче зображення при $k = 0,7$, $m = 35$; в) при $k = 0,3$, $m = 35$

Адаптивне перетворення гістограми

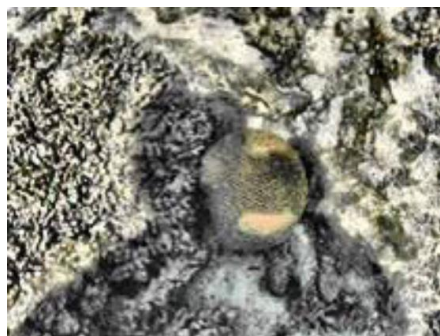
Цю операцію можна виконати аналогічно адаптивному підвищенню контрасту. Вона виконується так само, як і стандартне (глобальне) перетворення гістограми, але застосовується локально, тобто до фрагментів зображення. Наприклад, вирівнювання гістограми можна виконати в ковзному вікні (рис. 5.22).



а



б



в

Рисунок 5.22 – Вихідне зображення (а), після глобального вирівнювання гістограми (б), після адаптивного вирівнювання гістограми компоненти L у просторі Lab (в), розмір вікна 8×8

Основний недолік адаптивних методів обробки – збільшення обсягу обчислень. Для зменшення обчислень використовують рекурсивні процедури корекції статистичних характеристик поточного фрагмента при обчисленні характеристик сусіднього фрагмента. Наприклад, якщо фрагмент має форму прямокутника й у ньому обчислена гистограма яскравостей, при зсуві вправо на один піксель локальна гистограма коректується шляхом віднімання внеску пікселів, що утворюють лівий стовпець поточного фрагмента, і додавання пікселів, що утворюють правий стовпець сусіднього фрагмента.

Контрольні запитання

1. Які логічні операції проводять з зображеннями ?
2. В чому полягають яскравісні перетворення зображення ?
3. В чому полягає препарування зображень ?
4. Що таке бінаризація зображення ?
5. Які існують методи обчислення порогу бінаризації ?
6. Для чого проводять лінійне розтягання контрасту зображення ?
7. Які арифметичні операції проводять з зображеннями ?
8. Для чого проводять нелінійну корекцію яскравості зображення ?
9. Що таке соляризація зображення ?
10. Що таке гистограма зображення ?
11. В чому полягає еквалізація гистограми зображення ?
12. В чому полягає нелінійне перетворення гистограми зображення ?
13. В чому полягає локально-адаптивна обробка зображення ?
14. Як здійснюється адаптивне підвищення контрасту зображення ?
15. Як здійснюється адаптивне перетворення гистограми зображення ?

6 ПЕРЕТВОРЕННЯ ЛОКАЛЬНОГО ТИПУ

Цифрові зображення, сформовані різними оптико-електронними системами й зареєстровані різними приймачами, спотворюються під дією перешкод різного характеру. Так, перекручування зображення вносяться всіма компонентами пристрою відображення, починаючи з освітлювальної системи (наприклад, нерівномірність освітленості предмета). Оптична система вносить перекручування, які називаються абераціями, а електронні приймачі випромінювання, наприклад ПЗЗ-матриці, – електронний шум. Всі перешкоди, внесені в цифрове зображення, прийнято називати шумом. Шум утрудняє як візуальний аналіз зображень людиною, так й їхню автоматичну обробку.

Існує багато методів поліпшення зображень, які зменшують вплив

шуму, внесеного в процесі формування цифрового зображення.

6.1 Типи шумів

Шум, що з'являється на зображеннях, може бути класифікований різними способами. Він може бути корельований або некорельований, залежний або незалежний від самого реєстрованого сигналу й т.д. Знання щодо системи одержання зображень, візуальний аналіз зашумленого зображення допомагають у побудові моделі шуму й оцінці його статистичних характеристик. Математично певна модель шуму допомагає побудувати оптимальну для даного додатка систему поліпшення зображень.

Найчастіше використовуються наступні моделі шумів на зображеннях.

Під *імпульсним шумом* розуміється перекручування сигналу імпульсами, тобто викидами з дуже більшими позитивними або негативними значеннями й малою тривалістю. При обробці зображень імпульсний шум виникає, наприклад, внаслідок помилок декодування, які приводять до появи чорних і білих крапок на зображенні. Тому його часто називають крапковим шумом. Викиди шуму особливо помітні на дуже темних або дуже світлих ділянках зображень.

Імпульсний шум моделюється в такий спосіб. Поява шумового викиду в кожному пікселі (i, j) зображення має ймовірність p і не залежить ні від наявності шуму в інших точках зображення, ні від якості вихідного зображення. Значення вихідної яскравості пікселя замінюється новим значенням d (від 0 до 255). Нехай $\{x_{i,j}\}$ – перекручене зображення. Тоді

$$x_{i,j} = \begin{cases} d \text{ з вигогідністю } p \\ s_{i,j} \text{ з вигогідністю } (1 - p) \end{cases}$$

де $s_{i,j}$ – значення яскравості вихідного зображення. Якщо нове значення $d = 0$, то вносяться чорні значення яскравості (шум типу «перець»), якщо $d = 255$ – білі значення яскравості (шум типу «сіль»).

Параметр d може також приймати випадкові значення. Передбачається, що вони є незалежними випадковими величинами з гаусовим розподілом на інтервалі $[0, 255]$. Видалення такого шуму представляє більше складне завдання.

Адитивний шум описується співвідношенням

$$g(x, y) = f(x, y) + \eta(x, y),$$

де $f(x, y)$ – вихідне зображення; $g(x, y)$ – зашумлене зображення; $\eta(x, y)$ – адитивний і не залежний від сигналу шум з гаусовим або іншим

розподілом функції щільності ймовірності.

Гаусів шум (називаний також нормальним шумом) виникає на зображенні в результаті прояву таких факторів, як шум в електричних колах, шум сенсорів (через недолік висвітлення й/або високу температуру). Модель цього шуму широко застосовується при фільтрації зображень і сигналів, оскільки припускає незалежність і нульове математичне очікування перешкод.

Функція щільності розподілу ймовірностей гаусової випадкової величини z задається наступним вираженням (рис. 6.1):

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2 / 2\sigma^2},$$

де z визначає рівень яскравості із всіх можливих значень для зображення (наприклад, для восьмибітового зображення $0 < z < 255$); μ – середнє значення (математичне очікування) випадкової величини z ; σ – середньоквадратичне відхилення. Величина, яка дорівнює квадрату середньоквадратичного відхилення, називається дисперсією σ^2 і визначає потужність внесеного шуму.

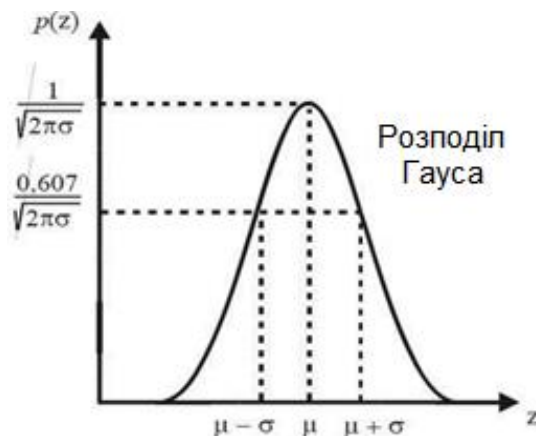


Рисунок 6.1 – Функція щільності розподілу ймовірностей гаусової випадкової величини z у формі дзвона

Таким чином, $p(x)$ визначає ймовірність того, що значення поточного відліку шуму буде мати яскравість, рівну z . Якщо випадкова величина z описується функцією Гауса, то приблизно 67 % її значень зосереджено в діапазоні $[(\mu - \sigma), (\mu + \sigma)]$ і близько 96 % у діапазоні $[(\mu - 2\sigma), (\mu + 2\sigma)]$.

Мультиплікативний шум описується співвідношенням

$$g(x, y) = f(x, y) \cdot \eta(x, y),$$

тобто шумова складова множить прийнятий сигнал. Прикладами такого шуму є: зернистість фотоплівки; зображення, які одержані за допомогою ультразвуку; радіолокаційні зображення.

Шум квантування залежить від сигналу й обраного кроку квантування значень сигналу. Наприклад, у результаті бінаризації яскравість зображення описується двома значеннями, що відповідають чорному й білому кольорам. Такий шум може робити небажані артефакти, наприклад помилкові контури навколо об'єктів, або забирати низько контрастні деталі на зображенні. Він не усувається.

Спекл-шум, як і у всіх сканувальних системах з когерентним формуванням зображення, викликається енергетичними перешкодами через безладно розподілені відбивачі сигналу, занадто дрібних для того, щоб їх могла відобразити система. Це клас шумів, які залежать від самого сигналу зображення. Для об'єктів з низьким відрізненням звичайно є мультиплікативним. Спекл-шум з'являється при одержанні цифрових зображень за допомогою ультразвукових медичних сканерів і радарів. У результаті зображення виглядає «зернистим», а на радарних зображеннях з'являються перекручування у вигляді світлих хрестів.

6.2 Фільтрація зображень

Ослаблення перешкод досягається фільтрацією. Для цього зображення описується двомірною функцією, значення якої змінюються повільніше, ніж значення двомірної функції, що описує шум. Різноманіття методів фільтрації зображень пов'язане з розмаїтістю математичних моделей сигналів, перешкод і критеріїв оптимальності фільтрації. Фільтрація проводиться в просторовій або частотній областях. Так, у частотній області зображення повинне бути перетворене в частотне подання, наприклад, за допомогою перетворення Фур'є, що відноситься до перетворень глобального типу.

Розглянемо основні методи поліпшення зображень у просторовій області на прикладі напівтонових зображень.

Перетворення, які для обчислення значення яскравості одного пікселя враховують значення пікселів у деякій околиці, називають *локальними*, а околиця – *вікном*. Форма й розміри вікна можуть бути довільними, положення центра вікна також може бути будь-яким, але для зручності обчислення частіше використовують прямокутні вікна.

Часто в операціях використовують значення деякої матриці, яка має ті ж розміри, що й околиця. Таку матрицю називають *маскою*, *фільтром*, *ядром фільтра*, а значення елементів матриці – коефіцієнтами. При лінійній фільтрації зображень форма й розмір використовуваної околиці визначають маску фільтра. *Маска фільтра* – це матриця вагових

коефіцієнтів. Центр маски сполучається з аналізованим пікселем, її коефіцієнти множаться на значення яскравості пікселів, накритих маскою. На практиці в основному використовують маски симетричної форми, центри яких збігаються із центром околиці, тобто маски. Розповсюдженим видом околиці є квадрат розміром 3×3 , 5×5 і т.п.

Фільтрація зображення $f(x, y)$, яке має розміри $M \times N$, за допомогою маски розмірами $m \times n$ описується формулою

$$g(x, y) = \sum_s \sum_t w(s, t) f(x + s, y + t)$$

де s й t – координати елементів маски відносно до її центра (у центрі $s=t=0$). Такого роду *перетворення називаються лінійними* і їх можна виконати за допомогою *операції згортки* (схематично ця операція представлена на рис. 6.2).

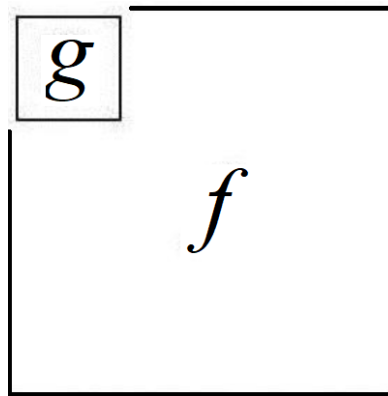


Рисунок 6.2 – Операція згортки зображення f з маскою g

Нехай f – функція яскравості зображення, а g – маска фільтра. Операція згортки зображення f за допомогою g позначається як $f * g$ й описується в такий спосіб:

$$(f * g)(m, n) = \sum_{k,l} f(m - k, n - l) g(k, l) \quad (6.1)$$

Якщо маска не симетрична, то її треба перетворити симетрично відносно її ж центра, а потім виконати множення за формулою (6.1).

Ідея фільтрації зображення ґрунтується на інтегральному використанні яскравості пікселів, що потрапили в задане вікно, з урахуванням коефіцієнтів маски фільтра (рис. 6.3).

У цьому складається головна відмінність фільтрації від розглянутих вище операцій обробки зображень точкового типу: фільтрація є процедурою обробки зображень локального типу. Після обчислення нового значення яскравості пікселя вікно, у якому описана маска фільтра, зсувається й обчислюється яскравість наступного пікселя. Тому подібне

перетворення називають *фільтрацією в ковзному вікні*.

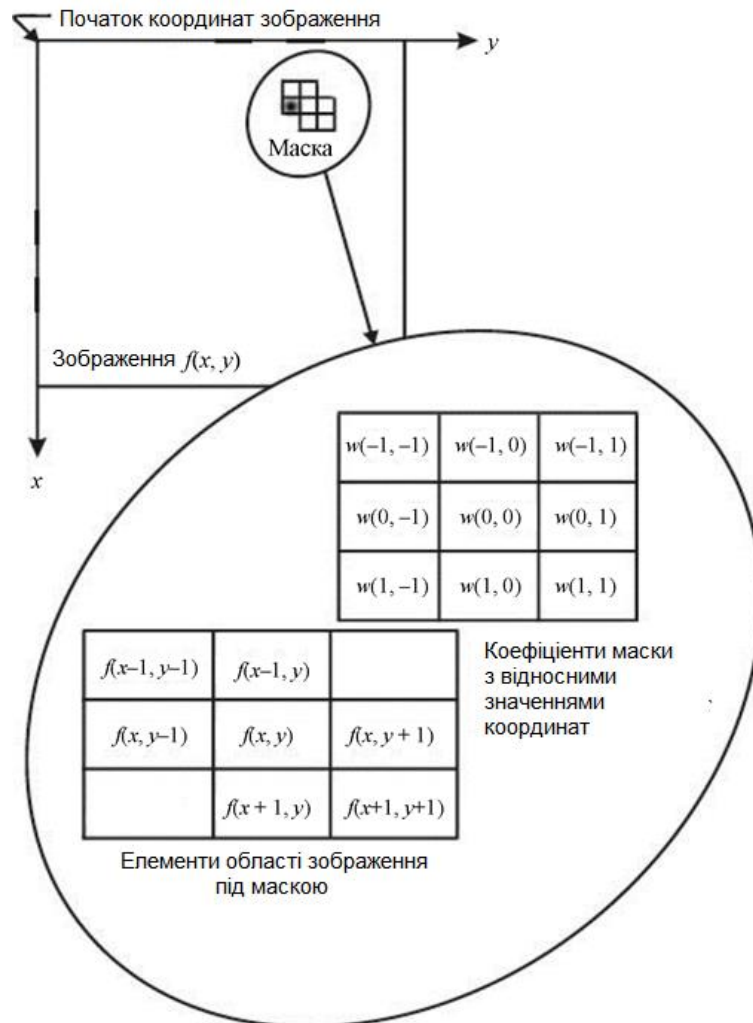


Рисунок 6.3 – Зображення, вікно фільтра, накладене на піксель (показаний темними кольорами), коефіцієнти ядра фільтра (причому $w(1, -1) = w(-1, 1) = 0$), значення яскравості пікселів, які потрапили у вікно

При реалізації операцій просторової фільтрації у вікні виникає невизначена ситуація, коли центр вікна близький до границі зображення й кілька елементів маски перебувають поза зображенням. Існує три способи, що враховують цю обставину:

- 1) обмеження переміщення центра вікна по зображенню (тобто не виходити вікном за краї зображення);
- 2) розширення зображення шляхом додавання рядків і стовпців з нульовими значеннями;
- 3) розширення зображення, повторюючи рядки й стовпці зображення симетрично границі.

Після обробки додані рядки й стовпці видаляються. Це дозволяє одержати оброблене зображення розміром з вихідне. Варто пам'ятати, що

значення пікселів, які використалися для розширення зображення, впливають на результат фільтрації пікселя усередині граничної смуги, ширина якої залежить від розмірів околиці, використовуваної при фільтрації. Якщо околиця має розміри 5×5 пікселів, то з кожної сторони додаються по два рядка і стовпця.

6.3 Низькочастотні фільтри

Низькочастотні просторові фільтри залишають низькочастотні компоненти зображення (області із плавними змінами яскравості) недоторканими й послабляють високочастотні компоненти. Такі фільтри використовуються для зниження шуму, що втримується в зображенні, а також для видалення високочастотних компонентів із зображення для того, щоб можна було ретельніше досліджувати зміст низькочастотних компонентів. Результатом низькочастотної фільтрації є розмиття зображення. Відмітні ознаки низькочастотних фільтрів: не негативні коефіцієнти маски; сума всіх коефіцієнтів дорівнює одиниці.

Приклади ядр низькочастотних фільтрів:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Розглянемо основні варіанти низькочастотних фільтрів. Вони реалізуються лінійними операціями локального типу.

Більшу групу низькочастотних фільтрів становлять *усереднювальні (або що згладжувальні) фільтри*. В усереднювальних фільтрах застосовують різні способи знаходження середнього значення яскравості в околиці. Розрізняють арифметичні, геометричні, гармонійні й контргармонійні усереднювальні фільтри.

Арифметичний усереднювальний фільтр або усереднювальний *бох-фільтр* усереднюють значення яскравості пікселя по околиці шляхом використання маски з однаковими коефіцієнтами, наприклад, для маски розміром 3×3 коефіцієнти рівні $1/9$, при 5×5 – $1/25$. Таке нормування необхідне для того, щоб привести значення результату фільтрації до діапазону яскравості вихідного зображення. Графічне подання двомірної функції, що описує маску цього фільтра, схоже на коробку або паралелепіпед, звідси й англійська назва – *бох-фільтр*.

На рис. 6.4 наведений приклад графічного подання функції

$$f(x, y) = \frac{1}{MN} \sum_{i=0}^M \sum_{j=0}^N g(i, j) \quad (6.2)$$

де $f(x, y)$ – значення пікселя вихідного зображення; $g(i, j)$ – поточне

значення пікселя вихідного зображення, що відповідає центру маски; M й N – ширина й висота маски відповідно.

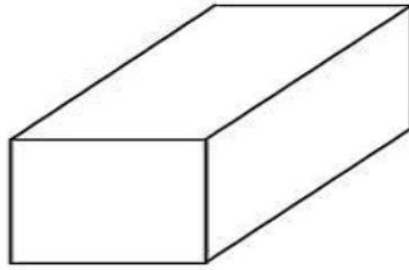


Рисунок 6.4 – Графічне подання функції усереднювального фільтра (бокс-фільтра)

Ця операція аналогічна операції лінійної фільтрації з коефіцієнтами маски, рівними $1/(MN)$. З формули (6.2) витікає, що арифметичний метод усереднення є часткою випадку алгоритму лінійної згортки й може застосовуватися для усунення високочастотних перешкод і згладжування зображень. Алгоритм арифметичного усереднення ефективний лише для слабо зашумлених зображень.

Геометричний усереднювальний фільтр. Геометричне усереднення може бути записане у вигляді наступної формули:

$$f(x, y) = \left[\prod_{i=0}^M \prod_{j=0}^N g(i, j) \right]^{\frac{1}{MN}}.$$

При такому усередненні відбувається згладжування зображення, аналогічне арифметичному усередненню, спостерігається погіршення різкості, властиве всім фільтрам із цього класу, але окремі об'єкти вихідного зображення спотворюються менше. Цей фільтр так само, як і усереднювальний арифметичний, може використовуватися для придушення високочастотного адитивного шуму, маючи при цьому кращі статистичні характеристики, хоча нові значення яскравості довше обчислюються.

Гармонійний усереднювальний фільтр. Фільтр на базі гармонійного усереднення ґрунтується на виразі

$$f(x, y) = \frac{MN}{\sum_{i=0}^M \sum_{j=0}^N \frac{1}{g(i, j)}}.$$

Даний фільтр добре працює з імпульсними шумами типу «сіль» (білими крапками) і не працює із шумами типу «перець».

Контргармонічний усереднювальний фільтр являє собою

узагальнення арифметичного й гармонійного усереднень. Цей фільтр міняє яскравість пікселя відповідно до функції

$$f(x, y) = \frac{\sum_{i=0}^M \sum_{j=0}^N g(i, j)^{Q+1}}{M \cdot N \sum_{i=0}^M \sum_{j=0}^N g(i, j)^Q},$$

де Q – порядок фільтра.

Контргармонічний фільтр ідеально підходить для зменшення або повного усунення імпульсних шумів типу «соль-і-перець», причому при додатних значеннях Q придушуються шуми «перець», а при від’ємних – «сіль». Однак одночасне видалення біполярного імпульсного шуму (білих і чорних крапок) неможливо. При $Q = 0$ контргармонічний фільтр перетворюється в арифметичний, а при $Q = -1$ – у гармонічний, тому даний фільтр можна вважати узагальненим фільтром із класу усереднювальних.

Фільтр Гауса. При завданні фільтрів можна використовувати маски з різними ваговими коефіцієнтами (рис. 6.3). Логічно припустити, що пікселі, розташовані ближче до аналізованого пікселя, впливають на яскравість обчислювану в процесі фільтрації. До такого типу фільтрів відносять фільтри, значення ваг яких описуються дзвоноподібною функцією Гауса

$$f(x) = \frac{1}{\sigma^2 \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

де μ – координата центральної крапки; σ – дійсна константа, яка визначає ширину дзвона.

При фільтрації зображень використовується двовимірний фільтр Гауса

$$G_\sigma = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}.$$

Інші варіанти подання цієї функції наведені на рис. 6.5.

Чим більше σ , тим більше розмивається зображення при застосуванні фільтра. Значення даного параметра варто вибирати залежно від того, наскільки сильно необхідно розмити зображення. Звичайно радіус фільтра r вибирається рівним 3σ . При цьому розмір маски дорівнює $2r+1$, тобто вона описується матрицею розміром $(6\sigma+1) \times (6\sigma+1)$. За межами цієї околиці значення функції Гауса будуть дорівнюють нулю або нехтувано малі.

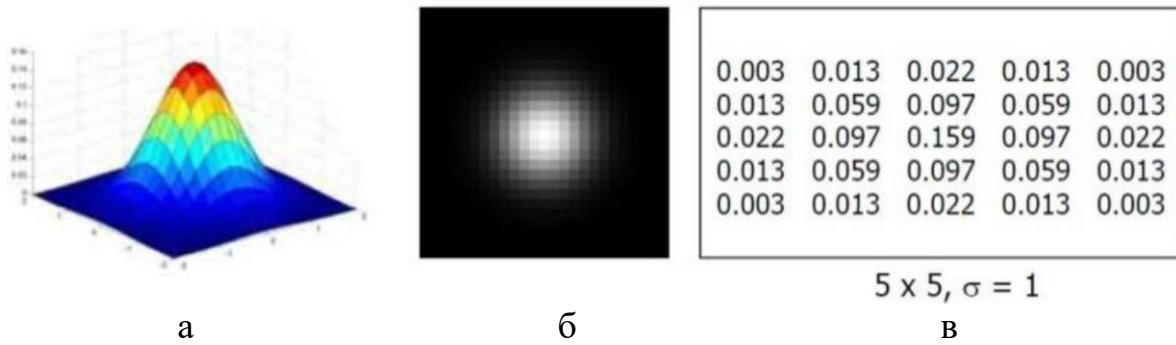


Рисунок 6.5 – Різні подання функції Гауса у вигляді:
 а) поверхні; б) напівтонового зображення; в) маски з коефіцієнтами

Фільтр Гауса має важливі властивості:

1) є сепарабельним (роздільним). Це значить, що двовимірна функція Гауса може бути представлена у вигляді добутку двох одновимірних функцій:

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) = \left(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right)\right) \cdot \left(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{y^2}{2\sigma^2}\right)\right)$$

що дозволяє знизити кількість обчислень із $(2r+1)^2$ до $2(2r+1)$ на кожен піксель, тобто приблизно в r раз, де $(2r+1)$ – розмір маски;
 2) згортка, виконана два рази з маскою фільтром радіуса r , дає той же результат, що й з маскою радіуса $r\sqrt{2}$

6.4 Високочастотні фільтри

Високочастотні фільтри використовуються для виділення перепадів яскравості. На їхній базі будуються алгоритми підкреслення границь і виділення невеликих об'єктів. Вони підсилюють різкі перепади на зображенні, у той же час як плавні придушуються. Після обробки високочастотним фільтром зображення стає більше різким.

Основний клас високочастотних фільтрів виділяє краї. Крайовим називається піксель, у якому різко змінюється локальна яскравість зображення. Методи виділення таких пікселів називають детекторами краю. Їхня зв'язна сукупність утворить границі об'єктів, представлених на зображенні.

На зображенні можна виділити два основних типи країв: у вигляді сходинки (або ската) і у вигляді даху (рис. 6.6).



Рисунок 6.6 – Основні типи країв

Як правило, різку зміну яскравості можна визначити, аналізуючи першу похідну функції яскравості (рис. 6.7). Похідна характеризує величину градієнта. Границя об'єкта на зображенні звичайно перпендикулярна вектору градієнта.

Для обчислення похідної застосовують високочастотні фільтри, які апроксимують обчислення похідних за напрямом. При цьому Δx дорівнює одиниці або двом. У масках високочастотних фільтрів частина коефіцієнтів має негативні значення, але сума всіх коефіцієнтів маски дорівнює нулю, що означає обчислення різниці значень яскравості сусідніх пікселів. Для прискорення обчислень всі коефіцієнти, як правило, цілочисельні.

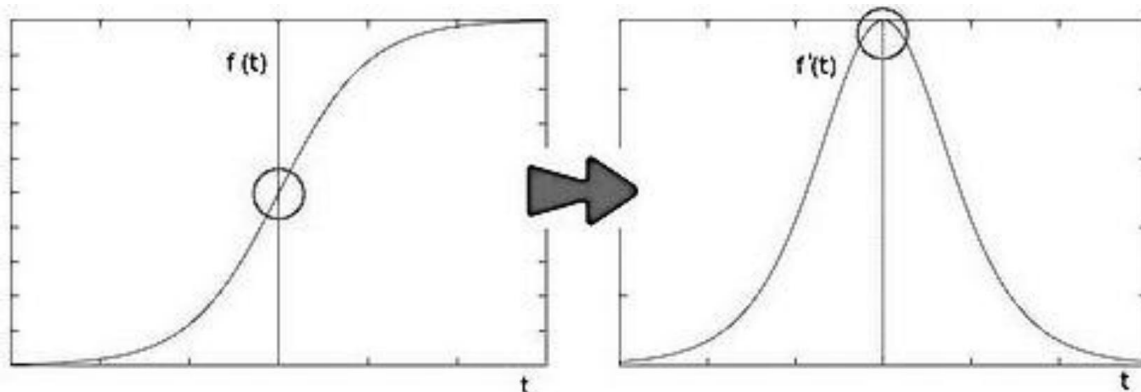


Рисунок 6.7 – Функція яскравості і її перша похідна, максимум похідної відповідає краю

Фільтр Робертса. Метод Робертса є найпростішим, швидким і досить ефективним. Він працює з найменшими можливими масками розміром 2×2 (рис. 6.8).

У результаті згортки зображення з масками фільтра Робертса одержуємо оцінку градієнта (перепаду яскравості) за напрямом G_x , G_y . Модуль градієнта (максимум перепаду) всіх детекторів краю обчислюється за формулою $G = \sqrt{G_x^2 + G_y^2}$ або спрощено $G = |G_x| + |G_y|$. Напрямок градієнта, тобто максимального перепаду яскравості, обчислюється за формулою $\arctg(G_y/G_x)$.

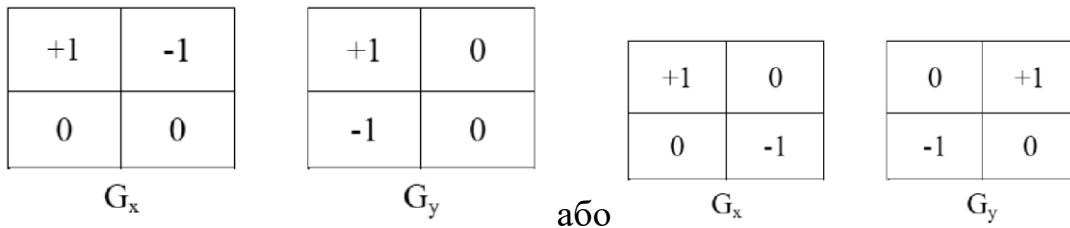


Рисунок 6.8 – Маски фільтра Робертса для обчислення похідних по X і по Y

Фільтр Превіта. Метод Превіта пропонує дві ортогональні матриці (маски Превіта) розмірністю 3×3 для більш точного обчислення похідних по X і по Y :

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix}, \quad G_y = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ +1 & +1 & +1 \end{bmatrix}.$$

Фільтр Собела використовує різні ваги в масках. Класичним варіантом цього фільтра є маски

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}, \quad G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Фільтр Лапласа. Попередні фільтри використовували апроксимацію першої похідної по осі X і по осі Y . У фільтрі Лапласа застосовується апроксимація других похідних по цих осях (рис. 6.9).

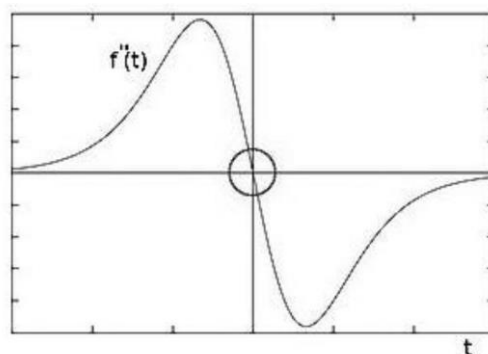


Рисунок 6.9 – Друга похідна функції яскравості міняє знак (проходить через нуль у точці, яка відповідає краю)

У цьому випадку градієнт обчислюється незалежно від напрямку й границі виділяються точніше, ніж у попередніх методах. Лапласіан від функції $f(x, y)$ – це сума других похідних

$$L(f(x, y)) = \frac{d^2 f}{dx^2} + \frac{d^2 f}{dy^2}.$$

Для дискретних функцій другі похідні можуть бути апроксимовані в такий спосіб:

$$L(f(x, y)) = [(f(x, y) - f(x - 1, y)) - (f(x + 1, y) - f(x, y))] + [(f(x, y) - f(x, y - 1)) - (f(x, y + 1) - f(x, y))] = -f(x, y - 1) - f(x - 1, y) - f(x, y + 1) - f(x, y - 1) + 4f(x, y).$$

Даний вираз дозволяє реалізувати фільтр Лапласа за допомогою згортки маски $w(s, t)$ із зображенням $f(x, y)$:

$$w(s, t) = \begin{vmatrix} 0 & -1 & 0 \\ -1 & +4 & -1 \\ 0 & -1 & 0 \end{vmatrix}.$$

Після згортки близькі до нуля значення відповідають крайовим пікселям. Маска фільтра Лапласа може бути описана й іншими способами із цілочисельними вагами, наприклад:

$$\begin{vmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{vmatrix}, \quad \begin{vmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{vmatrix}.$$

Виділення країв *методом Кенні*. Ціль алгоритму Кенні – більш точне виділення не крайових пікселів, а зв'язних граничних ліній на зображенні (рис. 6.10). Він заснований на використанні двох порогів при аналізі матриці градієнтів. Пороги виділяють сильні й слабкі границі, причому слабкі відзначаються тільки в тих випадках, якщо вони з'єднані із сильними границями.

Алгоритм Кенні.

Крок 1. Згладити зображення. Задати параметри радіуса й сигми. Виконати згладжування зображення фільтром Гауса.

Крок 2. Обчислити градієнти всіх пікселів. Застосувати фільтр Собела для обчислення значень модуля градієнта у всіх пікселях зображення. Напрямок градієнта округляється до 0, 45, 90 й 135.

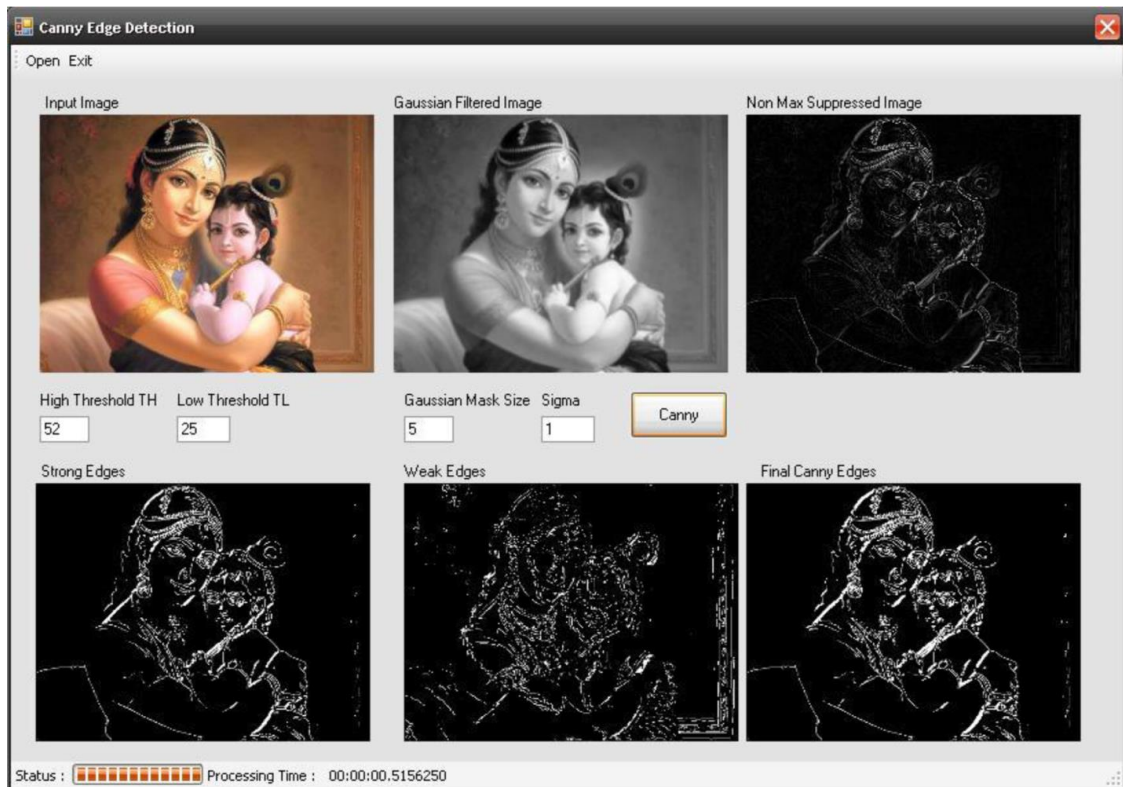


Рисунок 6.10 – Приклад послідовної реалізації алгоритму виділення границь методом Кенні

Крок 3. Придушити немаксимуми модуля градієнта. Крайовим пікселям відповідають максимальні значення модуля градієнта. Порівнюючи значення модулів градієнтів, відібрати пікселі ортогональним обчисленням напрямкам градієнта. Якщо напрямок градієнта в пікселі дорівнює 0, проаналізувати значення модулів градієнта ортогонально розташованих сусідніх пікселів, тобто вертикальних сусідів. Піксель вважається крайовим, якщо його градієнт більше, ніж у сусідів, інакше він – не максимум. Аналогічно аналізуються значення модуля градієнта, якщо він має три інших напрямки (45, 90 й 135). Після придушення немаксимумів краї стають більше точними, тонкими й зв'язковими. Простий вибір локальних максимумів виділить мало крайових пікселів.

Крок 4. Виконати подвійну граничну фільтрацію. Задати два граничних значення T_1 і T_2 . Для пікселів, відібраних на попередньому кроці, уточнити, чи є вони крайовими, зрівнявши значення їхніх модулів градієнтів з порогом T_1 . Чим нижче значення порога T_1 , тим більше пікселів буде відібране й більше зашумленим буде результат. Високе значення порога T_1 , навпаки, може відкинути слабкі крайові пікселі й одержати лінію краю з безліччю розривів. Для виділення країв використовуються два значення порога: якщо значення модуля градієнта в пікселі вище верхнього порога T_2 , то він приймає максимальне значення (наявність краю в ньому вважається достовірною); якщо менше нижнього

порога T_1 – піксель не вважається крайовим. Пікселі зі значеннями модуля градієнта, що попадають у діапазон між порогами T_1 і T_2 , вважаються неоднозначними, їм привласнюють фіксоване середнє значення, а їхній статус уточняється на наступному кроці.

Крок 5. Уточнити краї шляхом трасування області неоднозначності. Підсумкові краї одержують шляхом придушення всіх крайових пікселів, не пов'язаних з раніше певними достовірними крайовими пікселями. Виділяються групи пікселів, що одержали на попередньому етапі середнє значення. Якщо вони з'єднані з одним з достовірних країв, то вони відносяться до краю, а в протилежному випадку відкидаються. Піксель додається до групи, коли він сусідить із нею по одному з восьми напрямків (8-зв'язне сусідство). Крайовим пікселям привласнюються білі кольори, іншим – чорні, або навпаки (рис. 10).

6.5 Нелінійна фільтрація

Низькочастотні методи фільтрації приводять до згладжування різких перепадів яскравості на зображенні. Вони лінійні й оптимальні при гаусовому розподілі перешкод у спостережуваних даних. На реальних зображеннях у районі границь різних об'єктів розподіл яскравості має інший вид. У таких випадках для фільтрації використовують нелінійні методи. Розглянемо їх докладніше.

Медіанна фільтрація. Перешкоди у вигляді білих або чорних крапок на зображенні – це шуми імпульсного типу. Лінійні фільтри не усувають їх повністю, а лише локально усереднюють їхнє значення. Шуми такого типу видаляються за допомогою нелінійних фільтрів, наприклад медіанних.

При класичній медіанній фільтрації використовується поняття околиці і її центра, але не задаються вагові коефіцієнти. Околиця може мати довільну форму й розмір, а центр розташовуватися довільно щодо околиці. При сполученні центра околиці з аналізованим пікселем околиця стає вікном, у яке попадає ряд сусідніх із центром пікселів. Значення яскравості пікселів, що потрапили у вікно, сортуються за зростанням, значення середнього в ряді (медіанного) елемента після сортування й буде результатом медіанної фільтрації в даному вікні. Потім вікно зсувається, і процедура повторюється для всіх пікселів вихідного зображення.

На практиці вікно часто має прямокутну форму з непарною кількістю елементів, а його центр розташовується в геометричному центрі. Наприклад, нехай у вікно розміром 3×3 потрапили дев'ять пікселів, після сортування їхніх значень у результаті медіанної фільтрації центральний піксель приймає значення 6 (рис. 6.11).



Рисунок 6.11 – Медіанна фільтрація у вікні 3×3

Оскільки вікно може мати довільну форму, для його опису використовують прямокутну маску, елементи якої містять значення 0 й 1. Безліч одиниць описує вікно потрібної форми. У процесі фільтрації в сортуванні беруть участь тільки пікселі, які відповідають ненульовим елементам такої маски.

Існують алгоритми *зваженої медіанної фільтрації*. У масці, що відповідає вікну такого фільтра, крім 0 й 1 використовуються цілі числа 2, 3 й ін. Вони означають, скільки разів повторити яскравість відповідного пікселя перед сортуванням. Як правило, більші ваги розташовуються ближче до центра вікна для посилення впливу яскравості центральних пікселів на результат. Число значень яскравості N , які беруть участь у сортуванні, дорівнює сумі ваг маски. Номер медіанного елемента після сортування буде дорівнювати $(N+1)/2$.

Медіанний фільтр має наступні властивості:

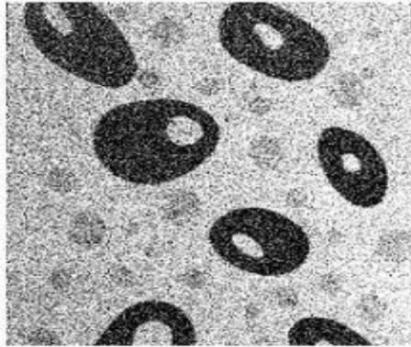
- несепарабельний;
- нелінійний;
- на напівтонових зображеннях не вносить нових значень яскравості, які відсутні у вихідному зображенні;
- якісно видаляє шуми імпульсного типу.

Медіанні фільтри успішно поліпшують відскановані старі паперові фотографії з білими смугами на місці згинів.

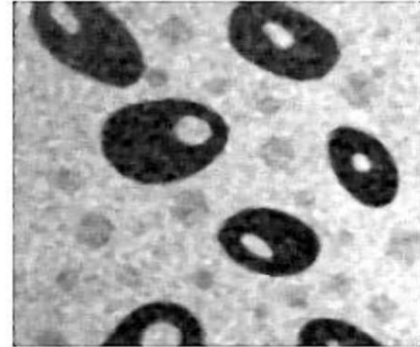
На рис. 6.12 приводяться приклади медіанної фільтрації напівтонового, зашумленого гаусовим адитивним шумом з різними розмірами вікна.

Медіанний фільтр добре справляється зі слабким і середнім ступенем зашумлення, але для придушення більш інтенсивних шумів необхідно використати медіанний фільтр із більшими розмірами вікна фільтрації. Відзначимо, що зі збільшенням розміру вікна росте здатність медіанного фільтра придушувати шумовий компонент. Однак збільшення розмірів вікна фільтра може привести до перекручування обрисів об'єктів. Крім того, малі за розміром об'єкти можуть бути цілком вилучені із зображення.

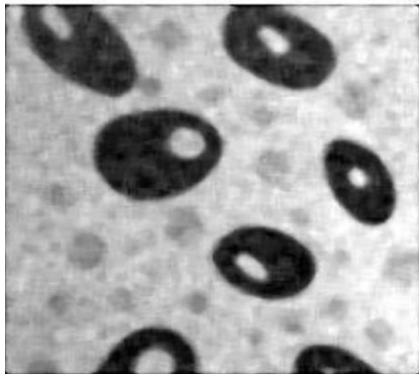
Тому в кожному конкретному випадку параметри фільтра необхідно налаштувати залежно від ступеня перекручувань і характерних розмірів спостережуваних об'єктів. Іноді краще застосувати два рази фільтр із вікном 3×3 , чим один раз із вікном 5×5 (рис. 6.13).



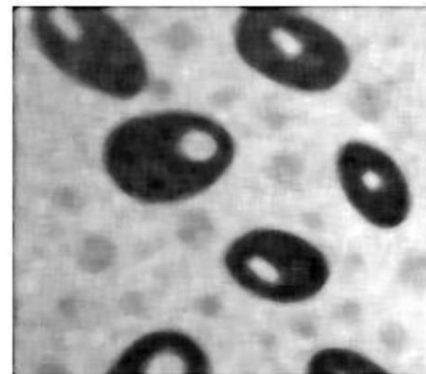
а)



б)

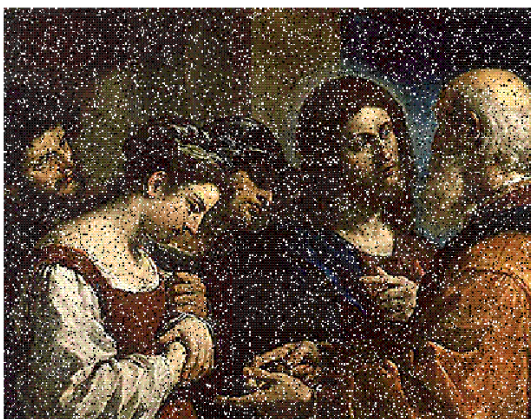


в)

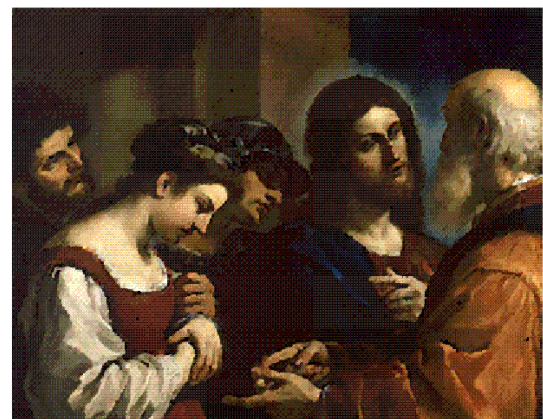


г)

Рисунок 6.12 – Приклади медіанної фільтрації: а) вихідне зображення з гаусовим шумом (сильний ступінь зашумлення); б) результат у вікні 5×5 ; в) 7×7 ; г) 9×9



а)



б)

Рисунок 6.13 – Приклад трикратного застосування медіанного фільтра: а) шум займає 15 % площі зображення; б) відфільтроване зображення

Результат медіанної фільтрації не чутливий до щільності розподілу шуму. Це пов'язане з тим, що медіанний фільтр із площею вікна $2M+1$ ефективно придушує шумові пікселі у вікні сумарною площею менш M . У той же час, при фільтрації контрастних великорозмірних об'єктів медіанний фільтр не розмиває й не зміщає їхнього краю (точки перепаду яскравості).

Варіант *адаптивного медіанного фільтра* використовує прямокутне вікно розміром $S \times S$. Адаптація складається в збільшенні розміру вікна S у процесі фільтрації залежно від його локальних статистичних характеристик.

Введемо позначення: Z_{min} , Z_{max} , Z_{med} – мінімальне, максимальне, медіанне значення у вікні; $Z_{i,j}$ – значення пікселя з координатами (i, j) ; S_{max} – максимально припустимий розмір вікна.

Алгоритм адаптивної медіанної фільтрації.

Крок 1. Завантажити зображення. Задати параметр S_{max} – максимально припустимий розмір вікна. Організувати цикл по рядках і стовпцям.

Крок 2. Задати вихідні розміри вікна фільтрації S .

Крок 3. Для пікселя (i, j) з яскравістю $Z_{i,j}$ обчислити Z_{min} , Z_{max} , Z_{med} , $A_1 = Z_{med} - Z_{min}$, $A_2 = Z_{med} - Z_{max}$ в заданому вікні. Якщо $A_1 > 0$ й $A_2 < 0$, перейти на крок 4. У протилежному випадку варто збільшити розмір вікна. Якщо поточний розмір вікна $S \leq S_{max}$, то повторити крок 3. У протилежному випадку результат фільтрації дорівнює величині $Z_{i,j}$.

Крок 4. Обчислити $B_1 = Z_{i,j} - Z_{min}$, $B_2 = Z_{i,j} - Z_{max}$. Якщо $B_1 > 0$ й $B_2 < 0$, результат фільтрації дорівнює $Z_{i,j}$. У протилежному випадку результат фільтрації дорівнює Z_{med} .

Крок 5. Змінити координати (i, j) . Якщо не вийшли за межі зображення, перейти на крок 2. У протилежному випадку фільтрація кінчена.

Ціль кроку 3 – визначити, чи є результат медіанної фільтрації в околиці пікселя Z_{med} імпульсним шумом (чорним або білим), чи ні. Якщо умова $Z_{min} < Z_{med} < Z_{max}$ виконується, то Z_{med} не може бути імпульсом. У цьому випадку здійснюється перехід на крок 4 і перевіряється, чи є піксель $Z_{i,j}$ у центрі вікна імпульсом. Якщо умови $B_1 > 0$ й $B_2 < 0$ вірні, то $Z_{min} < Z_{med} < Z_{max}$ й $Z_{i,j}$ не може бути імпульсом по тій же причині, що й Z_{med} . Тоді алгоритм повертає вихідне значення яскравості $Z_{i,j}$. Тому що подібні пікселі із проміжними значеннями не змінюються (на відміну від медіанного фільтра, який обробляє всі пікселі зображення), сумарний рівень перекручувань у відновленому зображенні менше (рис. 6.14). Якщо умова $B_1 > 0$ й $B_2 < 0$ невірна, тоді $Z_{i,j} = Z_{min}$ або $Z_{i,j} = Z_{max}$. У тім й іншому випадку значення пікселя є критичним й алгоритм виводить медіану Z_{med} , що, як треба із кроку 3, не є імпульсним шумом.

Припустимо, що на кроці 3 виявлений імпульсний шум. Тоді збільшується розмір вікна й повторюється крок 3. Повтор кроку

відбувається доти, поки алгоритм не знаходить медіанне значення, яке не є імпульсним шумом (і виконує перехід до кроку 4), або поки не досягнутий максимальний розмір вікна. Якщо це так, то алгоритм повертає величину Z_{ij} . Помітимо, немає гарантії, що це значення не є шумовим імпульсом. Чим менше ймовірність шуму p , тим вище ймовірність визначення підходящого значення яскравості. Для фільтрації шумів з високою щільністю потрібен більший розмір вікна. Останній етап – обробка аналогічно алгоритму стандартного медіанного фільтра.

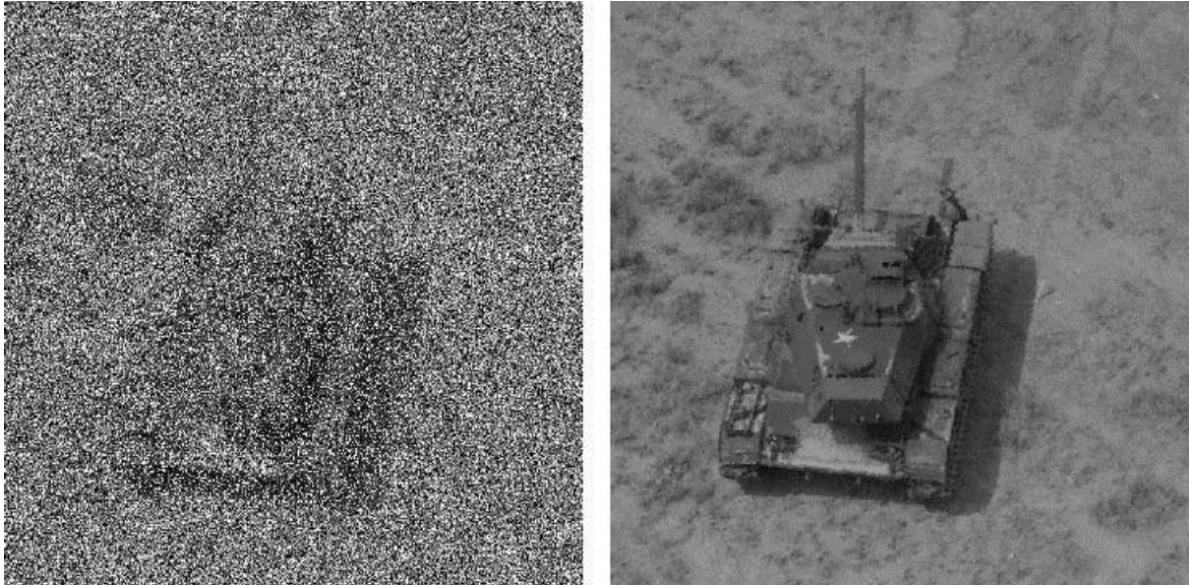


Рисунок 6.14 – Приклад фільтрації адаптивним медіанним фільтром: зображення з імпульсним шумом (ліворуч), результат фільтрації описаним алгоритмом (праворуч)

Таким чином, розглянута модифікація медіанного фільтра переслідує наступні цілі:

- оптимальне видалення імпульсних шумів;
- згладжування інших типів шумів;
- зменшення перекручувань у вигляді втрати дрібних деталей.

Основний недолік адаптивних методів обробки – збільшення обсягу обчислень. Для зменшення обчислень використовують рекурсивні процедури корекції статистичних характеристик поточного фрагмента при обчисленні характеристик сусіднього фрагмента. Наприклад, якщо фрагмент має форму прямокутника й у ньому обчислена гістограма яскравості й, при зсуві вправо на один піксель локальна гістограма коректується шляхом віднімання внеску пікселів, що утворюють лівий стовпець поточного фрагмента, і додавання пікселів, що утворюють правий стовпець сусіднього фрагмента.

Рангова фільтрація. Медіанний фільтр є частковим випадком класу

фільтрів, названих ранговими, або порядковими. Ранговий фільтр порядку r ($1 < r < N$, де N – число елементів в околиці) вибирає з отриманого ряду елемент із номером r і привласнює його значення як результат фільтрації пікселя вихідного зображення.

Якщо число N непарне й $r = (N + 1)/2$, фільтр стає медіанним. Якщо $r = 1$, фільтр вибирає мінімальне значення яскравості у вікні й називається *min-фільтром*. Якщо $r = N$, фільтр вибирає максимальне значення яскравості у вікні й називається *max-фільтром*.

Іноді ранг задається у відсотках. Тоді вибір мінімального значення відповідає 0 %, медіанного – 50 %, а максимального – 100 %.

6.6 Внесення ефектів у зображення

Одним зі способів обробки зображень є їхня зміна шляхом внесення різних ефектів. Опишемо деякі з них.

Підвищення різкості зображення. Якщо до центрального елемента маски Лапласа додати одиницю й виконати згортку зображення з такою маскою, фактично одержимо додавання до вихідного зображення градієнтної інформації. У результаті підвищиться різкість вихідного зображення. Для збільшення контрасту між центральним пікселем і сусідами використовуються від'ємні вагові коефіцієнти.

Приклад маски для збільшення різкості зображення:

$$G = \begin{vmatrix} -0,1 & -0,1 & -0,1 \\ -0,1 & 1,8 & -0,1 \\ -0,1 & -0,1 & -0,1 \end{vmatrix}$$

Сума коефіцієнтів маски дорівнює +1, але може й незначно відрізнятись за рахунок зміни значення центрального елемента маски.

Створення ефекту тиснення. Ефект тиснення можна додати зображенню за допомогою фільтрів

$$T_4 = \begin{vmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & -1 & -2 \end{vmatrix} \quad T_5 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{vmatrix} \quad T_6 = \begin{vmatrix} -2 & -1 & 0 \\ -1 & 0 & +1 \\ 0 & +1 & 2 \end{vmatrix}$$

Коефіцієнти в масці антисиметричні щодо центра або осі, що проходить через центр, а їхня сума дорівнює нулю.

При зміні знаків у масці фільтра на протилежні змінюється напрямок підсвічування.

Створення ефекту акварелізації. У результаті застосування даного

фільтра зображення буде виглядати так, начебто воно намальовано аквареллю.

На першому етапі згладжується яскравість редагованого зображення. Можна використати медіанний фільтр або низькочастотний, наприклад, з маскою

$$M = \frac{1}{16} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix}$$

На наступному етапі підвищується різкість переходів отриманого зображення для завершення створення ефекту акварелі, наприклад, високочастотним фільтром з маскою (рис. 6.15)

$$R = \begin{vmatrix} -0,5 & -0,5 & -0,5 \\ -0,5 & 5 & -0,5 \\ -0,5 & -0,5 & -0,5 \end{vmatrix}$$



Рисунок 6.15 – До вихідного зображення (ліворуч) застосували медіанний фільтр у вікні 5×5 і згортку з маскою R ; результуюче зображення праворуч

Перспективні алгоритми. Вищезгадані алгоритми не враховують локальних особливостей зображень, і якщо зображення містить дуже темні й дуже світлі ділянки, то вони не можуть забезпечити якісного поліпшення зображення в цих ділянках. Для таких зображень застосовують так називані просунуті алгоритми, наприклад Retinex і його модифікації. Слово Retinex – композиція від слів retina (сітківка) і cortex (кора головного мозку). Едвін Н. Ленд в 1971 р. сформулював теорію ретінекса, що пояснює особливість колірної постійності людського зору, що полягає в тім, що людина сприймає кольори об'єкта приблизно однаково при зміні

кольорів освітлення (апельсин завжди виглядає жовтогарячим). Він же запропонував однойменний алгоритм Retinex і запатентував його (патент належить NASA).

Алгоритм Single-Scale Retinex (SSR) вирівнює освітленість зображення, зберігаючи локальний контраст у погано освітлених і яскраво освітлених областях (рис. 6.16). Корекція зображення I виконується за формулою

$$R(x, y, \sigma) = \log [I(x,y)] - \log [I(x,y)*G(x,y,\sigma)], \quad (6.3)$$

де G – гаусіан; σ – параметр розмиття; $*$ – оператор згортки.

Після застосування формули (6.3) більшість значень R перебуває в діапазоні $[-1; 1]$. Для візуалізації R його значення необхідно перетворити в діапазон $[0..255]$ і округлити до цілих значень за формулою $I = \text{Int}(127,5*(R+1))$.



Рисунок 6.16 – Приклад застосування алгоритму Single-Scale Retinex при $\sigma = 18$: вихідне зображення (ліворуч), результат (праворуч)

Алгоритм SSR можна застосувати до вихідного зображення кілька разів з різними параметрами фільтра Гауса і обчислити зважену суму яскравості кожного пікселя, тоді одержимо алгоритм Multi-Scale Retinex (MSR) (рис. 6.17):

$$MSR = w_1 \cdot SSR_1 + w_2 \cdot SSR_2 + \dots + w_n \cdot SSR_n, \quad w_1 + w_2 + \dots + w_n = 1.$$

Звичайно n невелике, три або п'ять.

Порівняння фільтрів. Вибір конкретного фільтра і його параметрів залежить від завдання й умов її реалізації, наприклад швидкості обчислень. Арифметичний усереднювальний фільтр крім придушення шуму спотворює різкі границі й розмиває дрібні деталі зображення, але він найшвидший.

Фільтр Гауса менше розмиває дрібні деталі й краще вбирає шум,

досить швидкий, якщо коефіцієнти обчислюються один раз.

Медіанний фільтр різких границь не спотворює, забирає дрібні деталі, зображення стає більше однорідним. Якщо маска велика, сортування сповільнює обчислення.

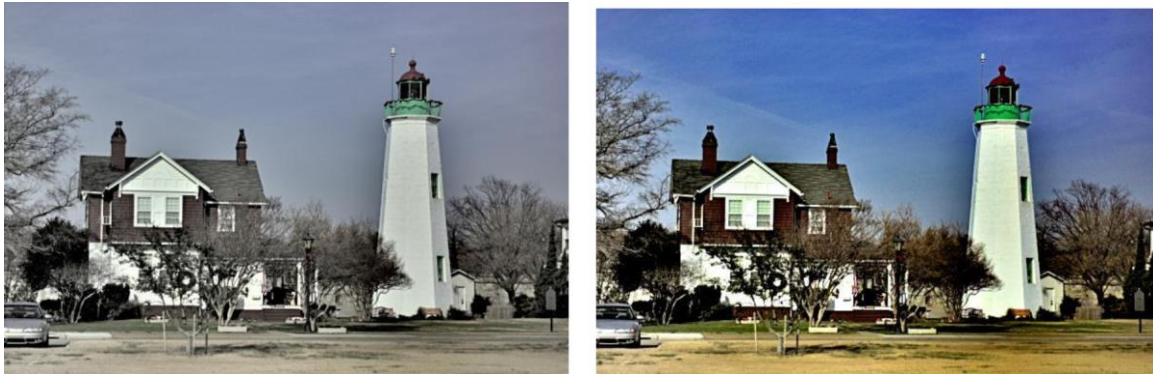


Рисунок 6.17 – Приклади застосування алгоритму Multi-Scale Retinex для значень $Q=7, 480, 4800$ (ліворуч) плюс лінійне розтягання контрасту (праворуч)

Адаптивний фільтр менше спотворює деталі, але залежить від більшого числа параметрів, які вимагають додаткових обчислень.

Перспективні фільтри краще зберігають деталі й менше розмивають контури на зображенні, але вони часто складні в реалізації й працюють повільніше.

Контрольні запитання

1. Що таке імпульсний шум ?
2. Що таке шум типу «сіть» ?
3. Що таке шум типу «перець» ?
4. Що таке адитивний шум ?
5. Що таке гаусів шум ?
6. Що таке мультиплікативний шум ?
7. Що таке шум квантування ?
8. Що таке спекл-шум шум ?
9. Що таке глобальне перетворення зображення ?
10. Що таке локальне перетворення зображення ?
11. Що таке фільтр зображення ?
12. В чому полягає фільтрація зображення в ковзному вікні ?
13. Яку функцію виконує низькочастотний фільтр ?
14. Які існують види низькочастотних фільтрів ?
15. Яку функцію виконує високочастотний фільтр ?
16. Які існують види високочастотних фільтрів ?

17. Для чого призначається нелінійна фільтрація ?
18. Що таке медіанний фільтр зображення ?
19. В чому полягає адаптивна медіанна фільтрація ?
20. Які ефекти можуть вноситися у зображення ?

7 ОПЕРАЦІЇ МАТЕМАТИЧНОЇ МОРФОЛОГІЇ ІЗ ЗОБРАЖЕННЯМИ

7.1 Математична морфологія бінарних зображень

Одним з порівняно нових напрямків в обробці зображень є математична морфологія. Вона базується на теорії множин, поданні зображень у вигляді геометричних форм й їхньому аналізу. Нагадаємо основні поняття з теорії множин.

Якщо $X = \{x\}$ й $Y = \{y\}$ – множини, то *об'єднанням множин* X й Y називається множина $Z = \{z : z \in X \text{ або } z \in Y\} = X \cup Y$ (множина, що складається з таких елементів z , які належать X або Y), а *перетинанням множин* X й Y називається $Z = X \cap Y = \{z : z \in X, z \in Y\}$. Множина $Z = X^c = \{z : z \notin X\}$ називається *доповненням множини* X . *Різницею множин* X й Y називається множина $Z = \{z : z \in X, z \notin Y\} = X \setminus Y$. Множина Z називається *порожньою*, якщо вона не містить жодного елемента, позначається $Z = \emptyset$. Справедливі наступні співвідношення:

$$\begin{aligned}(X \cup Y)^c &= X^c \cap Y^c; \\ (X \cap Y)^c &= X^c \cup Y^c; \\ X \setminus Y &= X \cap Y^c;\end{aligned}$$

Бінарні зображення легко інтерпретувати у вигляді множин. Якщо яскравість пікселя дорівнює одиниці (білі кольори), він належить деякій множині, якщо нулю (чорний) – не належить.

Бінарне зображення можна розглядати як індикаторну функцію набору множин, тобто плям (наприклад, об'єднання множин на рис. 7.1). До такого подання бінарних зображень застосуємо математичний апарат теорії множин.

Для будь-якої множини X можна задати деякий елемент (не обов'язково приналежній цій множині), називаний центром (або початком координат) цієї множини. Позначимо через X_a множини X центр якого поміщений у точку a .

Для бінарних зображень легко визначити міру множини X – це його площа або кількість вхідних у множини пікселів.

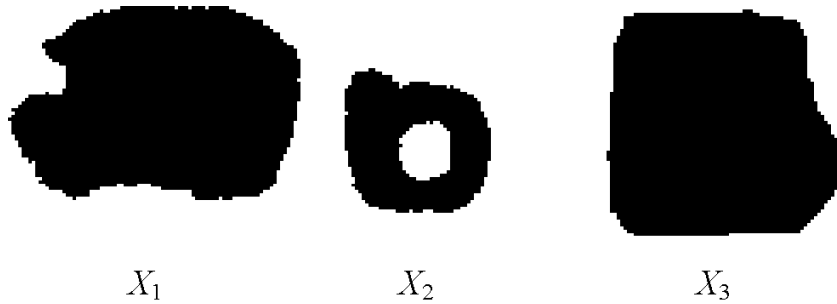


Рисунок 7.1 – Приклади бінарних зображень

Одним з основних понять математичної морфології є поняття структурного елемента. *Структурний елемент* B – це множина точок, для якого визначений його центр (не обов'язково є центром цієї множини). Важливо, що структурний елемент визначає деяку геометричну форму, за допомогою якої виконуються морфологічні операції. Він може мати довільний розмір і форму, але на практиці частіше використовуються симетричні елементи (прямокутник, коло), а центром є їхні геометричні центри.

Розглянемо основні операції математичної морфології, використовувані для обробки бінарних зображень.

Ерозія та дилатація

Нехай заданий структурний елемент B . *Ерозією* множини X називається множина Y , що складається з тих елементів вихідної множини X , для яких виконується умова (рис. 7.2):

$$Y = X \ominus B = \{x : B_x \leq X\}.$$

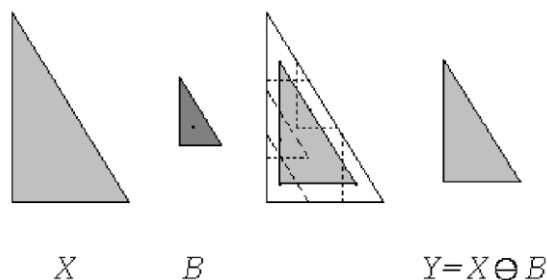


Рисунок 7.2 – Приклад виконання операції ерозії множини X з елементом B

Інакше кажучи, якщо $x \in X$, а $b \in B$, то в множину Y включаються такі елементи x , для яких при всіх b виконується умова

$$x + b \in X. \quad (7.1)$$

Ерозію можна трактувати в такий спосіб: поміщаємо центр структурного елемента в точку $x \in X$, якщо елемент повністю належить X , тоді точка $x \in Y$, де Y – множина подібних точок. Очевидно, що після цієї операції множина Y менше X .

З іншого боку, якщо точка b пробігає всі можливі положення в множині B , умова (7.1) виконується тоді й тільки тоді, коли x належить зміщеній множині X_{-b} (рис. 7.3). Тому операцію ерозії можна записати як

$$Y = X \ominus B = \bigcap_{b \in B} X_{-b} = \bigcap_{-b \in B} X_b = \bigcap_{b \in \check{B}} X_b.$$

де $\check{B} = \{b : -b \in B\}$ – множина, симетрична елементу Y відносно його центра.

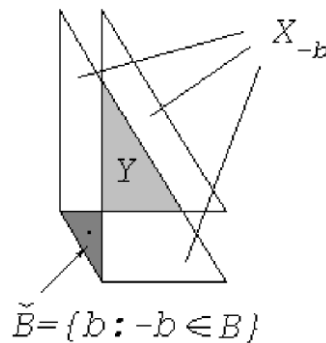


Рисунок 7.3 – Ерозія як перетинання зміщених множин

Операцією, двоїстою до ерозії, є дилатація (або розширення) (рис. 7.4), що визначається в такий спосіб:

$$Y = X \oplus B = \{x : B_x \cap X \neq \{\}\}.$$

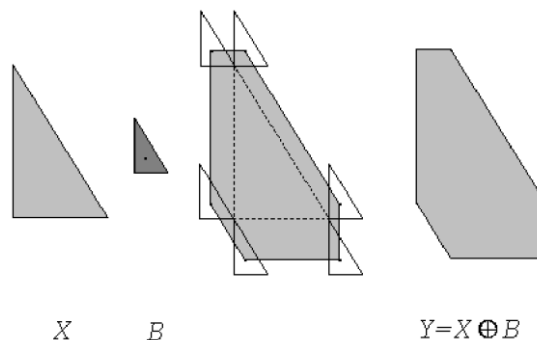


Рисунок 7.4 – Приклад виконання операції дилатації множини X з елементом B

Цю операцію можна записати по-іншому (рис. 7.5):

$$Y = X \oplus B = \bigcup_{-b \in B} X_b.$$

Дилатацію можна трактувати в такий спосіб: поміщаємо центр структурного елемента в усі точки $x \in X$ і додаємо до X всі точки зміщеного структурного елемента. Об'єднання отриманих точок визначає множину Y . Очевидно, що після цієї операції множина X збільшується.

Якщо розглядати множину X як об'єкт, а його доповнення X^c як тло на зображенні, то дилатацію об'єкта можна інтерпретувати як ерозію тла:

$$X \oplus B = (X^c \ominus B)^c. \quad (7.2)$$

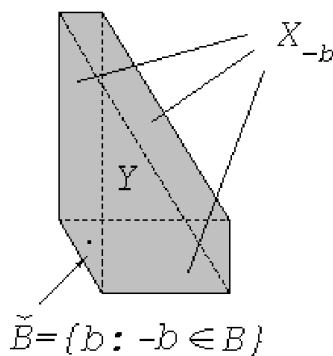


Рисунок 7.5 – Дилатація як об'єднання зміщених множин

Операції *ерозії* й *дилатації* є базовими в математичній морфології. Приведемо їхні основні властивості.

1. Дилатація й ерозія дистрибутивні щодо об'єднання множин

$$\left(\bigcup_{i=1}^N X_i \right) \oplus B = \bigcup_{i=1}^N (X_i \oplus B), \quad \left(\bigcup_{i=1}^N X_i \ominus B \right) = \bigcup_{i=1}^N (X_i \ominus B)$$

і щодо їхнього перетинання

$$\left(\bigcap_{i=1}^N X_i \right) \oplus B = \bigcap_{i=1}^N (X_i \oplus B), \quad \left(\bigcap_{i=1}^N X_i \ominus B \right) = \bigcap_{i=1}^N (X_i \ominus B)$$

Властивість дистрибутивності з врахуванням (7.2) дозволяє виконувати операції над X по фрагментах, комбінуючи потім результати за допомогою об'єднання або перетинання.

2. Дилатація асоціативна

$$(X \oplus A) \oplus B = X \oplus (A \oplus B),$$

а ерозія – ні, але вірно наступне:

$$(X \ominus A) \ominus B = X \ominus (A \oplus B),$$

Дані властивості дозволяють спростити виконання операцій за допомогою розкладання складних структурних елементів у композицію більше простих (рис. 7.6). При цьому операції зі складними структурними елементами виконуються швидше як послідовність операцій з більш простими елементами, які містять меншу кількість точок. Так, ерозію за допомогою структурного елемента H , наведеного праворуч на рис. 7.6, можна замінити чотирма послідовними операціями ерозії зі структурними елементами A, B, C, D , що значно зменшує обсяг обчислень.

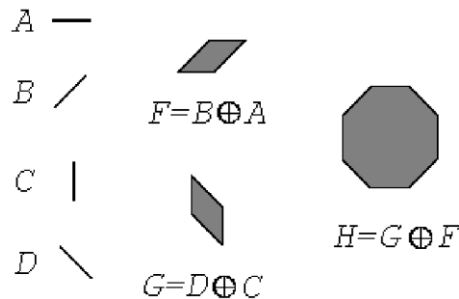


Рисунок 7.6 – Декомпозиція структурних елементів F, G й H

3. Інваріантність до зміни масштабу:

$$\lambda X \oplus \lambda B = \lambda(X \oplus B);$$

$$\lambda X \ominus \lambda B = \lambda(X \ominus B);$$

де $\lambda B, \lambda A$ – множини, промасштабовані щодо своїх центрів (рис. 7.7):

$$B' = \lambda B = \{b' = \lambda b : b \in B\},$$

$$X' = \lambda X = \{x' = \lambda x : x \in X\}.$$

Якщо центр структурного елемента B перебуває всередині нього, тоді для будь-яких множин A й C справедливі наступні відношення:

$$\text{якщо } A \subseteq C, \\ \text{то } A \ominus B \subseteq C \ominus B, \quad A \oplus B \subseteq C \oplus B.$$

Ерозія й дилатація – операції, призначені в першу чергу для виділення на бінарному зображенні об'єктів (зв'язних множин точок) з

певною формою шляхом використання структурних елементів різної форми й розмірів. Наприклад, ерозія за допомогою кола з радіусом r дозволяє знайти в зображенні об'єкти, мінімальний діаметр яких перевищує $2r$. Якщо ж в якості структурного елемента взяти дві точки, що утворюють вектор B , ерозія дозволяє виділити об'єкти, які мають сусідів у напрямку y на відстані задану цим вектором (рис. 7.8). Під об'єктами маються на увазі однозв'язні множини пікселів.

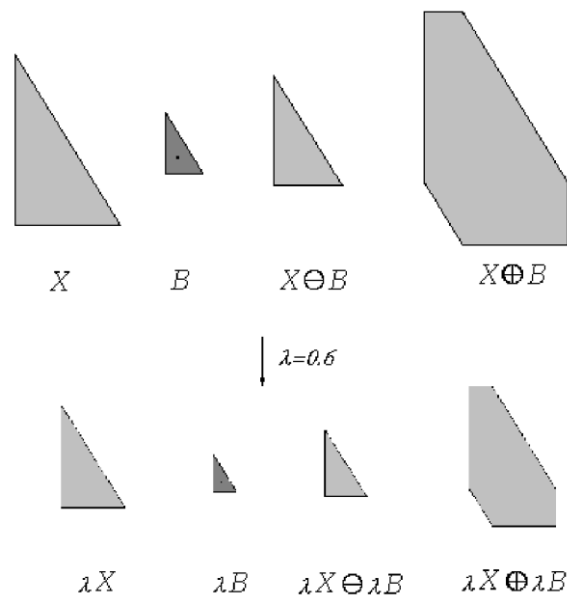


Рисунок 7.7 – Операції ерозії й дилатації інваріантні до масштабування з коефіцієнтом λ .

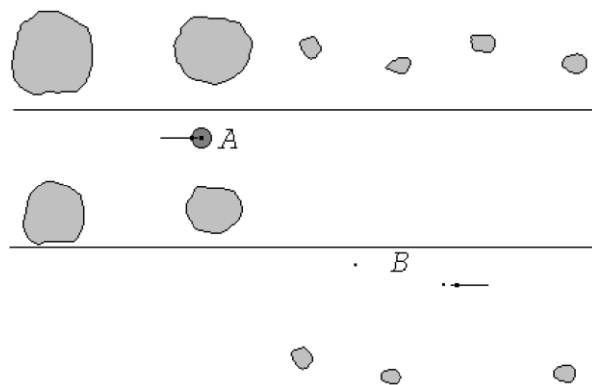


Рисунок 7.8 – Верхній ряд: вихідна множина X ; середній ряд: структурний елемент A і результат операції $X \ominus A$; нижній ряд: структурний елемент B і результат операції $X \ominus B$ (стрілками відзначені центри структурних елементів A й B)

Одне з найпростіших застосувань дилатації – це усунення розривів

ліній шляхом їхнього з'єднання. На рис. 7.9, а показане зображення з розірваними символами. Нехай відомо, що максимальна довжина розривів становить два пікселя. На рис. 7.9, б показаний найменший структурний елемент, який дозволяє усунути розриви. Результат дилатації вихідного зображення з таким примітивом показаний на рис. 7.9, в. У місцях розривів утворилися «містки». Одна з переваг морфологічного підходу в порівнянні з методом усунення розривів шляхом низькочастотної фільтрації полягає в тому, що морфологічний метод відразу дає бінарне зображення. При низькочастотній фільтрації з вихідного бінарного зображення спочатку виходить напівтонове, котре потім потрібно перетворити знову у бінарне зображення.

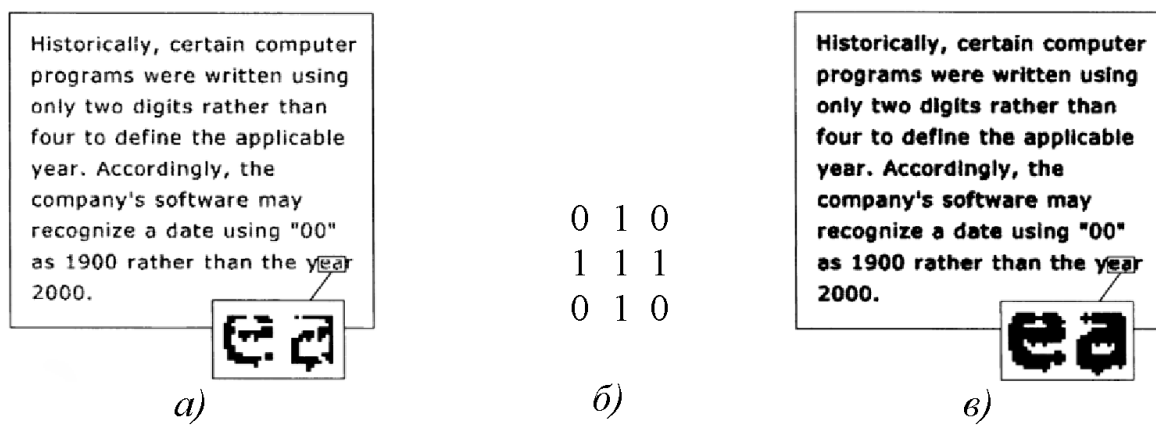


Рисунок 7.9 – Приклад дилатації тексту з недостатнім відірзненням

Розглянемо спосіб морфологічного виділення границь (рис. 7.10). Визначимо морфологічну границю множини X (тобто всіх темних плям на зображенні) як різниця дилатації й ерозії цієї множини із круглим структурним елементом B : $\delta X = X - (X \ominus B)$.



Рисунок 7.10 – Приклад морфологічного виділення границі бінарного об'єкта

Відімкнення й замикання

У загальному випадку неможливо повністю відновити вихідну множину X після виконання ерозії $X \ominus B$ шляхом застосування операції

дилатації з тим же структурним елементом B , але послідовне виконання ерозії й дилатації породжує нові операції математичної морфології.

Операція відімкнення множини X за допомогою структурного елемента B визначається формулою

$$A \circ B = (A \ominus B) \oplus B. \quad (7.3)$$

Операція замикання множини X за допомогою структурного елемента B визначається застосуванням ерозії й дилатації у зворотному порядку

$$A \bullet B = (A \oplus B) \ominus B. \quad (7.4)$$

Розглянемо основні властивості операцій (7.3) і (7.4). Вони двоїсті одна одній:

$$A \bullet B = (A^C \circ B^S)^C. \quad (7.5)$$

Стосовно до бінарних зображень співвідношення (7.5) означає, що операція замикання еквівалентна доповненню результату відімкнення доповнення A з елементом B^S , симетричним Y відносно центра. Очевидно, якщо B симетричний, то $B = B^S$.

Важливо наступна властивість операцій (7.3)...(7.5), їх ідемпотентність:

$$(A \circ B) \circ B = A \circ B \quad \text{и} \quad (A \bullet B) \bullet B = A \bullet B.$$

Дана властивість означає незмінність результату при повторенні операцій відімкнення й замикання.

Якщо центр структурного елемента B перебуває всередині нього, тоді для будь-якої множини A справедливо

$$A \ominus B \subseteq A \circ B \subseteq A \subseteq A \bullet B \subseteq A \oplus B.$$

Відімкнення й замикання використовуються для фільтрації бінарних зображень: згладжування форми об'єктів шляхом видалення дрібних і вузьких частин, видалення дрібних порожнин і вузьких «каналів» відповідно. Ступінь згладжування й розміри артефактів, які видаляються, залежать від розмірів і форми структурного елемента.

Розглянемо спосіб морфологічного виділення кутів і гострих піків. Детектор гострих кутів можна визначити як різниця між відімкненням множини X і замиканням X із круглим структурним елементом B :

$$gX_m = X \circ B - X \cdot B.$$

Метод нормалізації тла зображення використовує два структуруючих елементи B_1 й B_2 . Елемент B_1 вибирається дуже малим, щоб зберегти навіть найменші деталі об'єкта на зображенні, але відігнати імпульсний шум. Елемент B_2 має більші розміри, щоб придушити будь-який, навіть найбільший за геометричними розмірами, об'єкт (пляма) на зображенні. Операція нормалізації тла – це різниця операцій відімкнення й замикання:

$$h = (X \circ B_1) \cdot B_1 - (((X \circ B_1) \cdot B_1) \circ B_2) \cdot B_2.$$

Комбінуючи більше складні послідовності базових операцій математичної морфології, можна сконструювати інші операції, орієнтовані на рішення конкретних завдань.

Операція тоншання

Ціль тоншання – одержати зв'язну множину пікселів товщиною не більше одного. Ця операція виконується при обробці зображень, які містять тонкі лінії: креслення, текст, відбитки пальців і т.п.

Операція тоншання визначається в такий спосіб:

$$Y = XOT = X \setminus (X \otimes T),$$

де $T = (T_1, T_2)$ – структурний елемент, що складається із двох непересічних підмножин T_1 й T_2 .

Відзначимо, якщо початок структурного елемента належить T_1 , то $(X \otimes T) \subseteq X$, якщо ж початок належить T_2 , то $X \otimes T \subseteq X^C$. Тому $XOT = X$ при будь-якому T_1 . Щоб уникнути цього при виконанні операції тоншання, початок структурного елемента не повинний належати T_1 .

Модифікований алгоритм Зонга-Суня. Т. Зонг й У. Сунь в 1984 р. запропонували швидкий алгоритм тоншання об'єктів на бінарних зображеннях. В 1988 р. Ю. Зонг і П. Вонг модифікували цей алгоритм, спростивши й прискоривши його в 1,5 рази.

Вихідне зображення – бінарне зі значенням яскравості 0 й 1. Будемо використовувати наступну нумерацію сусідів аналізованого пікселя P_1 :

$$\begin{bmatrix} P_{10} & P_{11} & P_{12} & P_{13} \\ P_9 & P_2 & P_3 & P_{14} \\ P_8 & P_1 & P_4 & P_{15} \\ P_7 & P_6 & P_5 & P_{16} \end{bmatrix}.$$

Уведемо позначення: $B(P_1) = \sum_{i=2}^9 P_i$, $A(P_1)$ – число переходів

яскравості з 0 на 1 у послідовності сусідніх пікселів $P_9, P_2, P_3, P_4, P_5, P_6, P_7, P_8, P_9$,

Крок 1. У циклі по i й j для кожного пікселя $P(i, j)$ виконати: $P_1 = P(i, j)$.

Пікселю P_1 привласнюється 0, якщо

$P_1 = 1$,

$2 \leq B(P_i) \leq 6$,

$A(P_i) = 1$,

$(P_2 \times P_4 \times P_8 = 0$ або $P_{11} = 1)$,

$(P_2 \times P_4 \times P_6 = 0$ або $P_{15} = 1)$.

Крок 2. Якщо на попередньому кроці зображення змінилося, повторити крок 1.

Інакше – закінчити роботу.

Приклад виконання операції тоншання представлений на рис. 7.11.

Скелетон множини точок, заданого на дискретних ґратах, тільки приблизно нагадує скелетон безперервної множини. Більше того, для тієї ж самої множини результат побудови скелетона за допомогою послідовних стоншень може бути різним залежно від порядку застосування структурних елементів у послідовності (такі топологічні властивості скелетона, як кількість зв'язних компонентів, точок розгалуження, галузей, кінцевих точок й інші при цьому зберігаються).



Рисунок 7.11 – Вихідне бінарне зображення (ліворуч), результат тоншання (праворуч)

7.2 Математична морфологія напівтонових зображень

Розглянемо найбільш простий варіант узагальнення операцій математичної морфології для обробки напівтонових зображень. Напівтонове зображення описується функцією яскравості $f(x)$ і обробляється плоским структурним елементом B . При цьому можна вважати, що елемент також описується функцією нульової яскравості,

тобто $g(z) = 0$, якщо $z \in B$. У цьому випадку

$$(f \oplus B)(x) = \max\{f(x-z), z \in B\};$$

$$(f \ominus B)(x) = \min\{f(x+z), z \in B\}.$$

Результати виконання операцій ерозії й дилатації напівтонового зображення представлені на рис. 7.12. Перша розширює множину темних пікселів, друга – світлих.



Рисунок 7.12 – Вихідне зображення із прикладами ерозії (у центрі) і дилатації (праворуч) плоским структурним елементом квадратної форми

Ерозію напівтонового зображення плоским структурним елементом можна пояснити геометрично. Представимо, що функція яскравості задає якусь поверхню. Рухаємо плоский структурний елемент усередині поверхні по її краю так, щоб елемент повністю помістився усередині. Поверхня, обкреслена центром елемента усередині вихідної поверхні, буде необхідним результатом.

Дилатацію напівтонового зображення плоским структурним елементом можна трактувати подібним чином. Функція яскравості задає якусь поверхню. При дилатації, якщо структурний елемент не симетричний щодо свого центра, перетворимо його, сформувавши новий, симетричний щодо центра елемент. Рухаємо центр перетвореного структурного елемента по поверхні, розширюючи її в сторони відповідно до форми елемента. Поверхня, обкреслена краями елемента зовні вихідної поверхні, буде необхідним результатом.

Розглянемо чисельний приклад виконання цих операцій. Нехай плоский структурний елемент несиметричний і його форма задана одиницями $[0 \ 1 \ 1]$, центр елемента – це його геометричний центр (виділений жирним шрифтом), а рядок яскравостей оброблюваного зображення має значення

0 0 8 8 9 0 0.

Застосовуючи дилатацію, знаходимо максимум в околиці кожного пікселя, з яким сполучаємо центр елемента, симетричного заданому, його околиця відзначена одиницями $[1 \ 1 \ 0]$. У результаті наведені вище яскравості $0 \ 0 \ 8 \ 8 \ 9 \ 0 \ 0$ приймуть значення

$0 \ 0 \ 8 \ 8 \ 9 \ 9 \ 0$.

Застосовуючи ерозію, знаходимо мінімум в околиці кожного пікселя, з яким сполучаємо центр елемента, околиця відзначена одиницями $[0 \ 1 \ 1]$. У результаті яскравості $0 \ 0 \ 8 \ 8 \ 9 \ 0 \ 0$ приймуть значення

$0 \ 0 \ 0 \ 8 \ 8 \ 0 \ 0$.

Контрольні запитання

1. На чому базуються операції математичної морфології з зображеннями ?
2. В чому полягає операція ерозії ?
3. В чому полягає операція дилатації ?
4. В чому полягає операція відімкнення ?
5. В чому полягає операція замикання ?
6. В чому полягає операція тоншання ?

8 ОБРОБКА КОЛЬОРОВИХ ЗОБРАЖЕНЬ

Існують два основних типи обробки кольорових зображень: колірні компоненти обробляються окремо або кольори розглядається як єдиний вектор у тривимірному просторі.

У методах першого типу фактично виконується обробка R, G, B кольорів як окремих напівтонових зображень. Однак у результаті такої незалежної обробки (наприклад, розтягання контрасту) трьох кольорів можуть виникнути неприродні для вихідного зображення відтінки.

У методах другого типу зміни в колірні складові вносяться пропорційно щодо їх «внеску» у кожен піксель. Розглянемо деякі методи поліпшення кольорових зображень.

8.1 Корекція кольорів на зображеннях

Зміна колірного балансу зображення. Мета цього методу – компенсація невірної настроювання колірної камери або

кольорового освітлення сцени.

Передбачається, що користувач може інтерактивно вказати пікселі (або області) з потрібними опорними кольорами, наприклад білі кольори – кольори хмар на фотографії, або джерелом «правильних» кольорів може бути фотографія цієї ж (або схожої) сцени з добре переданими кольорами (рис. 8.1).



Рисунок 8.1 – Корекція зображення з опорними кольорами: а) зображення з поганим колірним балансом; б) відкоректоване зображення

Перетворити кольори по кожному з каналів можна за формулами

$$R = R \frac{R_{real}}{R_{orig}} \quad G = G \frac{G_{real}}{G_{orig}} \quad B = B \frac{B_{real}}{B_{orig}}$$

де індекс *orig* означає коригований колір пікселя на вихідному зображенні; *real* – бажаний колір.

Формування ідеальних відблисків. Передбачається, що найбільш яскраві області зображення є відблисками на поверхнях. У випадку ідеальних відбивачів (дзеркало) кольори відблиску повинні дорівнювати кольорам освітлення, тобто білому. Якщо значення кольорів у відблисках менше припустимого максимуму (звичайно 255), то всі кольори на зображенні можна скорегувати за наступним алгоритмом.

Алгоритм корекції кольору при ідеальних відблисках.

Крок 1. Знайти максимуми значень у відблисках по кожному з каналів R_{max} , G_{max} , B_{max} .

Крок 2. Змінити значення кольорів для кожного пікселя по формулах

$$R = R \frac{255}{R_{max}}, \quad G = G \frac{255}{G_{max}}, \quad B = B \frac{255}{B_{max}} .$$

Перетворення «Сірий світ». В основі лежить припущення, що сума всіх кольорів на зображенні природної сцени має сірі кольори. Приклад

корекції зображення представлений на рис. 8.2.



а б

Рисунок 8.2 – Зображення з поганим колірним балансом (а);
відкоректоване перетворенням «Сірий світ» (б)

Алгоритм корекції кольору перетворенням «Сірий світ».

Крок 1. Обчислити середні значення всіх кольорів за формулами

$$\bar{R} = \frac{1}{N} \sum R(x, y) \quad \bar{G} = \frac{1}{N} \sum G(x, y) \quad \bar{B} = \frac{1}{N} \sum B(x, y) .$$

Крок 2. Скорегувати значення кольорів відповідно до коефіцієнтів:

$$R' = R \frac{Avg}{\bar{R}} , \quad G' = G \frac{Avg}{\bar{G}} , \quad B' = B \frac{Avg}{\bar{B}} .$$

де R' , G' , B' – нові значення кольорів по кожному з каналів.

Лінійне розтягання контрасту кольорового зображення. Суть цього методу – лінійно розтягти інтенсивності по кожному колірному каналу на весь динамічний діапазон. Цей метод використовується в пакеті Photoshop у функції autolevels.

Алгоритм лінійного розтягання контрасту.

Крок 1. Знайти мінімум і максимум зображення по кожному з каналів

$$(R_{min}, R_{max}, G_{min}, G_{max}, B_{min}, B_{max}).$$

Крок 2. Перетворити відтінки кольорів кожного пікселя:

$$R' = (R - R_{min}) \frac{255 - 0}{R_{max} - R_{min}} ,$$

$$G' = (G - G_{min}) \frac{255-0}{G_{max}-G_{min}},$$

$$B' = (B - B_{min}) \frac{255-0}{B_{max}-B_{min}}.$$

8.2 Автоматичне обчислення порога бінаризації кольорового зображення

У деяких завданнях потрібно перетворити кольорове зображення в бінарне зображення (наприклад, подібно факсимільному апарату).

Евристичний вибір порога бінаризації T , наприклад, на підставі візуального вивчення гістограми, вимагає інтерактивного втручання користувача, що неможливо при автоматичній обробці безлічі зображень. На сьогоднішній день завдання автоматичного обчислення порога бінаризації T , при якому можна одержати високу якість і швидкість бінаризації зображень різних типів при будь-яких умовах зйомки, залишається невирішеною.

Існує кілька способів автоматичного підстроювання значення порога T . Для схематичних і картографічних зображень використовується метод Бернсена, де константа повинна підбиратися інтерактивно.

Метод Ейквеля часто застосовується для обробки чітких і контрастних зображень. Відповідно до даного методу зображення обробляється за допомогою двох концентричних вікон. При обробці тонких пересічних ліній можуть виникати розриви, тому даний метод рекомендується застосовувати винятково для товстих ліній і великих об'єктів.

Методом Яновиця і Брукштейна найкраще обробляти зображення, яким властива яскравісна зональна нерівномірність, коли ті самі об'єкти зображення в різних частинах мають значні розходження яскравості. До недоліків даного методу варто віднести вкрай повільну швидкість роботи.

Ефективним методом автоматичного вибору порога бінаризації є метод Отсу, що використовує гістограму розподілу значень яскравості растрового зображення. Цей метод дає гарний результат поділу об'єктів й однорідного тла на напівтонових зображеннях. До недоліків можна віднести розмивання й втрату тонких ліній, «злипання» об'єктів у місцях їхніх перетинань, а також ітераційний процес обчислення значення порога T .

При повній автоматизації й адаптації до різних умов одержання зображень при обчисленні порога T рекомендується метод, який називається *алгоритмом бінаризації k -середніх*. Опишемо його суть.

Нехай задана деяка початкова величина порога T , вона може бути будь-якою, тоді

$$\forall x \in 1 \dots M-2, \forall y \in 1 \dots N-2, \exists B(x, y) = \begin{cases} 1, & \nabla F \geq T, \\ 0, & \nabla F < T, \end{cases} \quad (8.1)$$

де M, N – відповідні значення ширини й висоти зображення; $B(x, y)$ – бінарне зображення (функція яскравості, що приймає значення 0 або 1); ∇F – модуль градієнта зображення;

$$\begin{aligned} \forall B(x, y) = 1, & H_i = R(x, y) + G(x, y) + B(x, y), \quad i \in 1 \dots \alpha, \\ \forall B(x, y) \neq 1, & L_j = R(x, y) + G(x, y) + B(x, y), \quad j \in 1 \dots \beta, \end{aligned} \quad (8.2)$$

де $R(x, y), G(x, y), B(x, y)$ – відповідні значення інтенсивності червоної, зеленої й синьої компоненти на вихідному кольоровому зображенні; α, β – відповідні значення кількості всіх точок перепаду, які залишились.

Отже, значення порога T дорівнює

$$T = \frac{1}{2} \cdot \left(\frac{\sum_{i=1}^{\alpha} H_i + \sum_{j=1}^{\beta} L_j}{(M-2) \cdot (N-2)} \right). \quad (8.3)$$

З виразів (8.1)...(8.3) витікає, що спочатку проводиться бінаризація зображення з використанням будь-якої оцінки порога T . У результаті утворюються дві групи пікселів: H – з пікселів зі значенням ∇F , більшим або рівним T ; L – з пікселів зі значенням ∇F , меншим T . Далі знаходиться половина середнього значення яскравості пікселів повнокольорового зображення по областях H й L відповідно. Отриманий результат і буде необхідним глобальним порогом T для всього зображення при використанні якого відбувається бінаризація зображень у масштабі реального часу.

В оригінальному методі бінаризації k -середніх наявність двох невідомих факторів (початкової оцінки значення порога для обчислення якої пропонується ряд методів, і значення наперед заданого параметра N_0 , що служить орієнтиром для кількості ітерацій даного алгоритму) утрудняє його застосування на практиці, з огляду на виникаючу невизначеність із часом виконання алгоритму. Зміст й основна частина оригінального методу не мінялися. У модифікованому методі зміни (точніше, ліквідації) підлягали два вищенаведених фактори. Також важливою особливістю модифікованого методу, описаного вище, є перехід від обробки вихідного монохромного зображення до мультиспектрального. Експериментально з'ясувалося, що яскравості червоної, зеленої й синьої компоненти

вихідного зображення не слід множити на коефіцієнти, що враховують сприйняття останніх у видимому діапазоні довжин хвиль. У протилежному випадку наведений метод не буде забезпечувати потрібну якість бінаризації зображення.

На рис. 8.3 представлений приклад виконання бінаризації зображення із застосуванням фільтра Собела для обчислення градієнта й автоматичним обчисленням порога бінаризації T .



Рисунок 8.3 – Приклад виконання бінаризації: а) вихідне кольорове зображення; б) бінарне зображення, отримане з вихідного вищеописаним методом

8.3 Підвищення чіткості кольорового портрета

Підвищити чіткість кольорового портрета можна різними способами. Один з них – перетворити RGB-зображення в простір Lab, що дозволяє уникнути появи кольорових ореолів, які виникають при значному збільшенні чіткості кольорового зображення. Простір Lab точніше інших колірних просторів моделює сприйняття кольорів людиною. Основна відмінність колірної моделі Lab полягає в повному поділі кольорів і контрасту. У моделях RGB й CMYK всі канали відповідають за кольори й контраст. В Lab весь контраст утримується в каналі L , інформація про кольори (точніше, про кольоровість) – у каналах a й b . Канал L можна представити як чорно-білу версію кольорового зображення. Канали a й b є чисто колірними. Якщо зображення не кольорове, канали a й b повинні бути сірими (точніше, на 50% сірими). Чим далі від цього рівня сірого й ближче до білого або чорного, тим кольоровіше стає зображення. Канал a розглядаємо як зображення з розподілом тонів від 0 до 100%. Величина 50% – це ні пурпурний і ні зелений колір. Більш світлі області означають наближення до пурпурного, а більш темні – до зеленого. У каналі b схема кольорів інша: ясно-сірі області представляють жовтий, а темно-сірі –

синій. Рівень 50% означає нейтральний сірий.

При обробці кольорових зображень Lab часто є проміжним поданням кольорів при корекції. Зображення перетворюються в цей простір для обробки, а по її завершенні конвертуються в простір відображення (не обов'язково RGB). Цим уникають появи кольорових ореолів, оскільки канали а й b, тобто канали, де перебуває інформація про кольори, не задіяні.

Алгоритм підвищення чіткості портрета.

Крок 1. Відкрити зображення й перетворити його з RGB в Lab.

Крок 2. До напівтонового зображення каналу L застосувати фільтр типу нерізкого маскування, одержимо L_{new} .

Крок 3. Перетворити зображення, представлене в $L_{new}ab$, в $R_{new}G_{new}B_{new}$ (рис. 8.4).

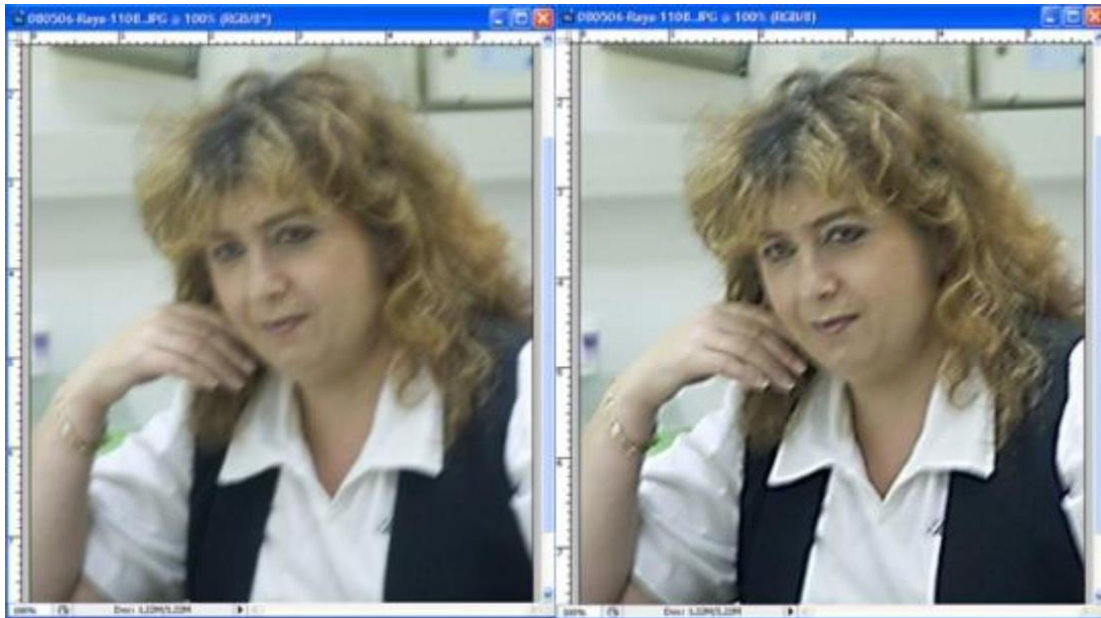


Рисунок 8.4 – Приклад підвищення чіткості в просторі Lab

Умови застосування параметрів максимального збільшення чіткості:

- портрет не у фокусі, тому необхідно максимально збільшити його чіткість, щоб спробувати повернути об'єкту знімка первісну різкість;
- на знімку багато об'єктів з яскраво вираженими границями (наприклад, архітектурні спорудження, монети, автомобілі, які-небудь механізми й ін.).

Для збільшення різкості в пакеті Photoshop використовується фільтр UnsharpMask з параметрами: Amount (величина) – 65 %, Radius (радіус) – 4 й Threshold (поріг) – 3. Дані параметри дають максимальне збільшення чіткості для сильно розфокусованих зображень. Причому фільтр

UnsharpMask з тими самими параметрами застосовується двічі.

Описаний алгоритм часто застосовується при обробці зроблених цифровою камерою в несприятливих умовах й відсканованих знімків. При застосуванні цього алгоритму для збільшення чіткості портрета, зробленого великим планом, часто з'являються на особі чітко видимі пори, зморшки й ін. Якщо збільшити чіткість всіх компонентів зображення RGB або ж чіткість тільки каналу L, використовуючи UnsharpMask, то текстура шкіри буде виглядати грубіше, ніж на вихідному зображенні. У цьому випадку можна виконати обробку в просторі RGB, застосувавши до каналу Red (червоний) фільтр UnsharpMask, тобто змінити чіткість тільки одного каналу Red. Інші використати без змін. Червоний канал, як правило, містить найменшу кількість дрібних деталей зображення, тому застосування фільтра UnsharpMask до одного цього каналу дозволяє підвищити чіткість тільки потрібних областей особи (око, губ і т.д.), а чіткість текстури шкіри залишається практично незмінної й гладкої (рис. 8.5). Якщо замість Red використати канал Blue, то збільшиться інтенсивність текстури шкіри.



Рисунок 8.5 – Приклад підвищення чіткості шляхом зміни тільки каналу R

Незначне збільшення чіткості використовується для обробки м'яких знімків, тобто зроблених на природі (наприклад, квітів, кошеняти, веселки, людей і т.п.). У цьому випадку варто використати параметри фільтра UnsharpMask (Amount – 150 %, Radius – 1 й Threshold – 10).

Для збільшення чіткості веб-зображень використовують значення Amount – 400 %, Radius – 0,3 й Threshold – 0. Зображення виходить м'яким і декілька розмитим. Перед розміщенням його в мережі Інтернет відрізнення високоякісного знімка зменшується з 300 до 72 dpi.

Якщо інтенсивність ефекту занадто велика, можна зменшити

значення параметра Amount до 200 %. Значення параметра Amount 400 % використовується для сильно розфокусованих знімків.

Контрольні запитання

1. Які існують типи обробки кольорових зображень ?
2. Як змінюють колірний баланс зображення ?
3. Як здійснюють корекцію кольору зображення ?
4. Як здійснюють лінійне розтягання контрасту кольорового зображення ?
5. Як обчислюють поріг бінаризації кольорового зображення ?
6. Як підвищують чіткість кольорового портрету ?

9 СЕГМЕНТАЦІЯ ЗОБРАЖЕНЬ

Часто зображення представляється для аналізу у вигляді сукупності однорідних ділянок. Кількість різних типів (або класів) ділянок звичайно невелике. Все зображення може бути розділене на непересічні області, кожна з яких представлена кольорами або текстурою одного типу (наприклад, море, пісок і рослинність на фотографії). При аналізі таких зображень важливим є виділення однорідних областей і визначення їхніх геометричних характеристик. Процес поділу зображення на однорідні області називається *сегментацією*. Відзначимо, що схожі за властивостями області відносяться до одного класу, але можуть розташовуватися в різних ділянках зображення.

Мета сегментації – розбивка всього зображення на однорідні в певному сенсі області. Позначимо через R область, займану всім зображенням. Сегментацію можна розглядати як процес розбивки R на n підобластей: R_1, R_2, \dots, R_n , при цьому повинні виконуватися умови:

- 1) $\bigcup_i^N R_i = R$;
- 2) R_i є зв'язною областю, $i = 1, 2, \dots, n$;
- 3) $R_i \cap R_j = \emptyset$ для всіх $i, j = 1, 2, \dots, n; i \neq j$;
- 4) $P(R_i) = \text{ІСТИНА}$ для $i = 1, 2, \dots, n$;
- 5) $P(R_i \cup R_j) = \text{НЕПРАВДА}$ для $i \neq j$.

Згідно з даними умовами: 1) сегментована повинна бути вся область зображення; 2) пікселі однієї області зв'язні; 3) області не повинні перетинатися; 4) пікселі одного класу однорідності; 5) різні класи відрізняються.

Кількість ознак, якими можуть відрізнятися одна від одної області різних типів, може бути різним. Велика розмаїтість у завдання сегментації вносить обробка кольорових зображень. Одною з найважливіших ознак

для сегментації в цьому випадку служить колір, що доповнює сукупність характеристик, застосовуваних при обробці напівтонових зображень.

Єдиного, або універсального, підходу до рішення завдання сегментації немає. Існує безліч алгоритмів сегментації. Кожний з них призначений для роботи з певним класом зображень і рішення конкретних завдань. Головна ідея більшої частини алгоритмів полягає в знаходженні відповідностей або розходжень у характеристиках пікселів. Характеристиками можуть бути їхнє взаємне розташування, яскравість, кольори і текстура.

У більшості випадків використовуються методи сегментації на основі граничної обробки, нарощування областей, поділи й злиття областей, морфологічного вододілу, аналізу текстур.

9.1 Сегментація зображень за яскравістю на основі граничної обробки

Обчислення порогів за гістограмою

Гранична обробка – найпростіший метод сегментації, орієнтований на обробку зображень, окремі однорідні ділянки яких розрізняються середньою яскравістю. Найпростішим варіантом сегментації є бінаризація зображення, тобто сегментація на два типи однорідних ділянок. При цьому перетворення кожного пікселя вихідного зображення $x(i_1, i_2)$ у вихідне $y(i_1, i_2)$ виконується за правилом

$$y(i_1, i_2) = \begin{cases} y_0, & \text{при } x(i_1, i_2) \leq T \\ y_1, & \text{при } x(i_1, i_2) > T \end{cases}$$

де T – єдиний параметр обробки, називаний порогом.

Рівні вихідної яскравості y_0 й y_1 можуть бути довільними, вони лише виконують функції міток, за допомогою яких здійснюється розмітка одержуваної маски – віднесення її точок до класів H_0 або H_1 відповідно. Часто їхні значення відповідають рівням чорного й білого. Якщо існує більше двох класів, то при граничній обробці повинен бути заданий набір порогів, що відокремлюють яскравості різних класів друг від друга. Центральним питанням граничної сегментації є визначення значень порогів, яке повинне виконуватися автоматично. Запропоновано багато методів обчислення порогів, в основному вони базуються на аналізі гістограми вихідного зображення.

Нехай $w(x)$, $0 < x < 255$, – гістограма напівтонового зображення, тобто яскравості змінюються в діапазоні від 0 до 255. Ідея гістограмного методу визначення порога ґрунтується на припущеннях про те, що розподіли ймовірностей для кожного класу унімодальні (тобто містять по

одному домінуючому піку), а граничних пікселів, які розділяють ділянки різних класів на зображенні, набагато менше, ніж пікселів приналежним обом класам. Цим припущенням відповідає гістограма, яка має багатомодальний характер. Окремі моди (пагорби на гістограмі) відповідають різним класам, а поділяючі їхні западини – нечисленним по кількості точок границь. Пороги сегментації перебувають при цьому по положенню западин. Рис. 9.1 ілюструє сказане вище стосовно до сегментації на два класи.

На практиці скористатися такими спрощеними міркуваннями для вибору порога вдається вкрай рідко, тому що реальні гістограми звичайно мають безліч дрібних піків і сильно порізані. У зв'язку із цим застосовується згладжування гістограми як одномірного масиву даних будь-яким низькочастотним фільтром, описаним раніше. Другою перешкодою є те, що границі між однорідними ділянками на зображенні бувають розмитими, внаслідок чого рівень гістограми в тих її частинах, які відображають точки границі, зростає. Це приводить до зменшення провалів у гістограмі або навіть їхньому зникненню.

Якщо на темному тлі присутні світлі й білі предмети, гістограма буде мати три піки. Сегментацію можна виконати за допомогою двох порогів. Аналогічно виконується сегментація по кольорах. Спочатку будуються гістограми окремих кольорів або вихідне зображення перетворюється в інше колірне подання, наприклад HSI, Lab. Потім виконується сегментація по I , L або ab .

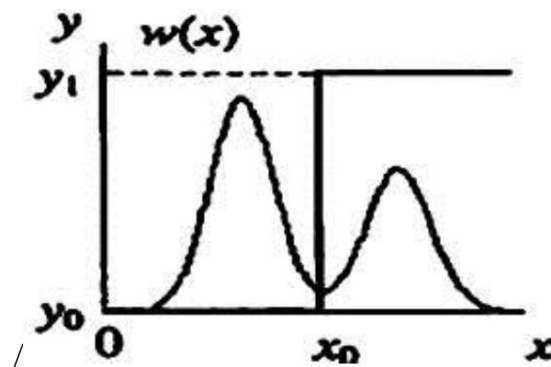


Рисунок 9.1 – Визначення порога по гістограмі при сегментації за яскравістю

Алгоритми, які адаптують своє поведіння, ґрунтуючись на статистиках оброблюваного зображення, мають більшу стійкість до зміни характеристик зображення.

Алгоритм сегментації методом k -середніх

Крок 1. Завантажити зображення, задати число класів k , на яке потрібно розбити зображення. Визначити ознаки, по яких виконується

сегментація, наприклад для напівтонових зображень можна використати яскравість. Всі пікселі розглядаються як набір векторів $x_i, i = 1, \dots, p$.

Крок 2. Визначити k векторів $m_j, j = 1, \dots, k$, які оголошуються початковими центрами кластерів. Для цього вибрати значення $m_j, j = 1, \dots, k$ (можна випадковим образом або обчислити).

Крок 3. Оновити значення середніх векторів $m_j, j = 1, \dots, k$ (центрів кластерів). Для цього обчислити відстань від кожного $x_i, i = 1, \dots, p$ до кожного $m_j, j = 1, \dots, k$. Віднести (приписати) кожен x_i до кластера j^* , відстань до центра якого m_{j^*} мінімально. Перерахувати середні значення m_j по всіх кластерах.

Крок 4. Повторювати кроки 2, 3 поки центри кластерів не перестануть змінюватися.

Сегментація за принципом Вебера

Опишемо алгоритм сегментації напівтонових зображень, заснований на принципі Вебера. Цей принцип має на увазі, що розходження рівнів сірого $W(I)$, які можуть бути замічені людським оком, можна описати нелінійною функцією градацій сірого I . Найпростішою функцією Вебера є

$$W(I) = \begin{cases} 20 - \frac{12I}{88}, & 0 \leq I \leq 88; \\ 0,002(I - 88)^2 + 8, & 88 \leq I < 138; \\ \frac{7(I - 138)}{255 - 138} + 13, & 138 \leq I \leq 255. \end{cases}$$

Відповідно до принципу Вебера людське око навряд чи розрізняє відмінності рівнів сірого між $[I(n), I(n) + W(I(n))]$, тому можна об'єднати рівні сірого з діапазону $[I(n), I(n) + W(I(n))]$, замінивши їх одним значенням яскравості.

Алгоритм нелінійного перетворення рівнів сірого для напівтонових зображень

Крок 1. Установити номер першого класу $n=1$ і початковий рівень сірого $I(n) = 0$.

Крок 2. Обчислити значення $W(I(n))$, яке відповідає яскравості $I(n)$ по формулі Вебера.

Крок 3. У вихідному зображенні I установити значення яскравості $I(n)$ для всіх пікселів, яскравість яких перебуває в діапазоні $[I(n), I(n) + W(I(n))]$.

Крок 4. Знайти пікселі, значення яскравості яких вище $G = I(n) + W(I(n)) + 1$. Якщо такі пікселі є, збільшити номер класу $n=n+1, I(n) = G$, перейти до кроку 2. Якщо таких немає, закінчити роботу. Зображення сегментоване на n класів, кожен клас показаний яскравістю $W(I(n))$.

Даний спосіб сегментації можна реалізувати, побудувавши таблицю

LUT, обчисливши нові яскравості за допомогою функції Вебера й замінивши вихідні значення яскравості знайденими номерами класів миттєво.

При сегментації кольорових зображень можна використати статистичні характеристики в околиці пікселя: відхилення кожної з колірних компонентів від середнього значення по околиці; дисперсію кожного колірної компонента в околиці пікселя й т.д.

Для простого випадку розбивки за яскравістю можна використати рекурсивний алгоритм нарощування областей (Region Growing).

Алгоритм сегментації Вежневца.

Вежневец розробив наступний ітеративний алгоритм сегментації.

Крок 1. Почати обхід зображення з лівого верхнього кута. Оголосити лівий верхній піксель зображення новим класом C_1 . Для пікселів першого рядка обчислити відхилення від класу лівого пікселя й порівняти із заданим порогом. Якщо воно менше порога, додати піксель до класу сусіда, інакше – завести новий клас C_{i+1} . Приналежність до класу можна відзначити номером цього класу в матриці такого ж розміру, як, наприклад, матриця, що відповідає компоненті R (червоні кольори).

Крок 2. Перший піксель кожного наступного рядка порівняти із класом верхнього сусіднього пікселя. Далі порівнювати піксель із класами двох сусідів: лівого й верхнього.

Якщо відхилення від обох порівнюваних класів більше порога, то завести новий клас, якщо відхилення більше тільки для одного класу, то додати піксель до того класу, відхилення від якого менше порога.

У випадку якщо відхилення припустиме для обох класів, можливі два варіанти:

$L(g(C_i) - g(C_j)) < \delta$ – об'єднати ці два класи (якщо це не той самий клас) і додати до об'єднаного класу поточний піксель;

$L(g(C_i) - g(C_j)) > \delta$ – додати піксель до того із двох класів, від якого відхилення мінімально.

Як міру L можна використати будь-яку функцію відстані, наприклад різниця в RGB-просторі.

Відзначимо, що алгоритм використовує чотири зв'язних пікселя, але можлива реалізація й для восьми зв'язних областей. Результат сегментації представлений на рис. 9.2.

Сегментація за кольорами шкіри

Використання кольорів шкіри як властивості особи при знаходженні осіб на зображеннях добре зарекомендувало себе на попередньому етапі в процесі розпізнавання людей, тому що кольори шкіри не залежить від орієнтації особи, аналіз кольорів пікселів в обчислювальному сенсі ефективний. У літературі описано багато підходів до сегментації кольорових зображень із метою виділення на фотографіях областей, які відповідають кольорам шкіри людей. Далі в знайдених областях

виконується детекція особи людини і його розпізнавання.



Рисунок 9.2 – Вихідне зображення (а) і результат його сегментації (б) при $\delta=50$

Як приклад виконаємо сегментацію зображення по кольорах шкіри. Рішення даного завдання складається в знаходженні критерію оцінки близькості кольору кожного пікселя до відтінку шкіри. Визначити поняття «відтінок шкіри» математично неможливо, тому що воно засновано на людському сприйнятті кольорів, міняється при зміні освітлення, відрізняється в різних рас. Проте існує ряд емпіричних підходів до наближеного рішення цього завдання. Їхня суть полягає в описі деякої моделі розподілу кольорів, за допомогою якої й оцінюється подібність із кольорами шкіри. Розробка моделі кольорів шкіри розбивається на три етапи:

- 1) накопичування навчальних даних, використовуючи зображення, на яких указують області «шкіра» й «не шкіра». За цими даними накопичується статистика відтінків шкіри (іноді гістограма відтінків, до шкіри не стосовних, тобто «не шкіра»);
- 2) обробку отриманої статистики й вибір параметрів моделі кольору шкіри для наступного використання; вибір критеріїв оцінки приналежності пікселів до області «шкіра»;
- 3) обробку зображень із використанням отриманих критеріїв.

Найпростішим (але не завжди таким, що дає коректні результати) варіантом такої сегментації є поділ колірних моделей RGB або інших за порогом.

Розглянемо граничні критерії, які в якості моделі кольору шкіри використовують багатогранник, заданий у певному колірному просторі, тобто задається набір правил для контекстно незалежної сегментації. Будемо вважати, що кольори пікселя (R, G, B) зараховується до області «шкіра», якщо виконані наступні умови:

$$R > 95, i G > 40, i B < 20, i \max\{R, G, B\} - \min\{R, G, B\} > 15,$$

$$i |R - G| > 15, i R > G \& R > B$$

або

$$R > 220, i G > 210, i B > 170, i |R - G| (15, i R > B, i G > B.$$

Інший варіант:

$$\left(\frac{r}{g} > 1,185 \right), \text{ и } \left(\frac{r \cdot b}{(r + g + b)^2} > 0,107 \right), \text{ и } \left(\frac{r \cdot g}{(r + g + b)^2} > 0,112 \right)$$

де $r = R / (R + G + B)$, $G = G / (R + G + B)$, $B = B / (R + G + B)$.

Після сегментації за кольорами шкіри (рис. 9.3) необхідно виконати морфологічну або медіанну фільтрацію.



Рисунок 9.3 – Приклад сегментації за кольорами шкіри

Простота реалізації – це безсумнівна перевага граничних підходів до сегментації, однак є й недоліки. Необхідно підбирати колірний простір і правила емпірично, при цьому можлива неточна сегментація шкіри, важко підібрати єдині критерії до людей різних рас.

9.2 Статистичний підхід до текстурної сегментації

У цифровій обробці зображень для опису текстури застосовуються два основних підходи: статистичний і структурний. Статистичні методи дозволяють охарактеризувати текстуру області як гладку, грубу й зернисту. Структурні методи визначають й описують взаємне розташування найпростіших повторюваних елементів зображення, наприклад відрізків паралельних ліній, що проходять із постійним кроком, кліток на шахівниці.

Розглянемо статистичний підхід до сегментації на прикладі зображення, яке містить два типи областей представлених різними текстурами (рис. 9.4).

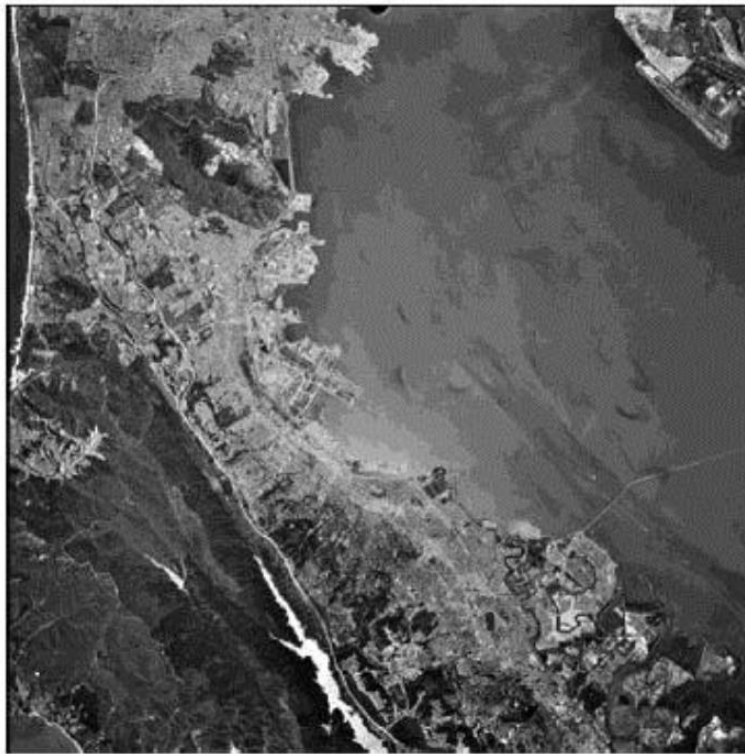


Рисунок 9.4 – Зображення з областями текстури різних типів, які відповідають суші та воді

Результатом сегментації є розбивка зображення на водну поверхню й сушу. Це неможливо зробити методами бінарізації, тільки шляхом аналізу параметрів текстури в околиці кожного пікселя.

Алгоритм сегментації за допомогою текстурного аналізу

Крок 1. Читати зображення.

Крок 2. Визначити параметри текстури на зображенні. Припускаючи, що яскравість у пікселях зображення – це випадкова величина z , їй відповідає ймовірність розподілу $p(z_i)$, $i = 0, 1, 2, \dots, L-1$, узятя з гістограми (L – число рівнів яскравості).

Центральний момент порядку n випадкової величини z_i дорівнює

$$\mu_n(z) = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i) \quad (9.1)$$

де m середнє значення z (середня яскравість зображення),

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (9.2)$$

З (9.1) витікає, що $\mu_0 = 1$ й $\mu_1 = 0$. Для опису текстури важливий другий момент, тобто дисперсія $\sigma^2(z) = \mu_2(z)$. Вона є мірою яскравісного контрасту, який можна використати для обчислення ознак відносної гладкості. Наприклад,

$$R = 1 - \frac{1}{1 + \sigma^2(z)} \quad (9.3)$$

дорівнює нулю для областей постійної яскравості (де дисперсія нульова) і наближається до одиниці для більших значень $\sigma^2(z)$. Оскільки для напівтонових зображень із яскравістю від 0 до 255 значення дисперсії великі, доцільно нормувати дисперсію до інтервалу $[0,1]$. Для цього необхідно поділити $\sigma^2(z)$ на $(L-1)$.

Значення стандартного відхилення

$$s = \sigma(z) \quad (9.4)$$

також часто використовується як характеристика текстури.

Третій момент є характеристикою симетрії гістограми:

$$\mu_3(z) = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i) \quad (9.5)$$

Для оцінки текстурних особливостей використовується функція ентропії, яка визначає розкид за яскравістю сусідніх пікселів:

$$e = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \quad (9.6)$$

де $p(z_i)$ – імовірність поточної яскравості в околиці точки; L – кількість рівнів яскравості; e – значення ентропії в поточній точці.

Для опису текстури також використовується міра однорідності, яка оцінює рівномірність гістограми:

$$U = \sum_{i=0}^{L-1} p^2(z_i) \quad (9.7)$$

У табл. 9.1 наведені значення описаних вище характеристик, підібрані для гладкої, грубої й періодичної текстур.

Таблиця 9.1 – Параметри трьох різних типів текстур

Текстура	Середнє	Стандартне відхилення	R (норм)	Третій момент	Однорідність	Ентропія
Гладка	82,64	11,79	0,002	-0,105	0,026	5,434
Груба	143,56	74,63	0,0079	-0,151	0,005	7,783
Періодична	99,72	33,73	0,017	0,750	0,013	6,674

Крок 3. Створити маску для виділення більшої текстури. Нехай на зображенні є текстури двох типів: велика й дрібна (більше зерниста). Дрібна відповідає водній зоні. Для відділення однієї області від іншої створити маску, яка видаляє маленькі об'єкти. Для цього використати функцію визначення зв'язної множини пікселів на бінарному зображенні й обчислити площі отриманих об'єктів. Використати тип зв'язності вісім. Якщо кольори сусідів збігаються, то вони належать до одного об'єкта, у протилежному випадку – до різних. Всі об'єкти із площею менше заданої величини S видалити.

Результат текстурної фільтрації на базі обчислення ентропії по формулі (9.6) у вікні 9×9 представлений на рис. 9.5. Бінарна маска водної поверхні після видалення об'єктів малої площі і результат сегментації суші представлені на рис. 9.6.

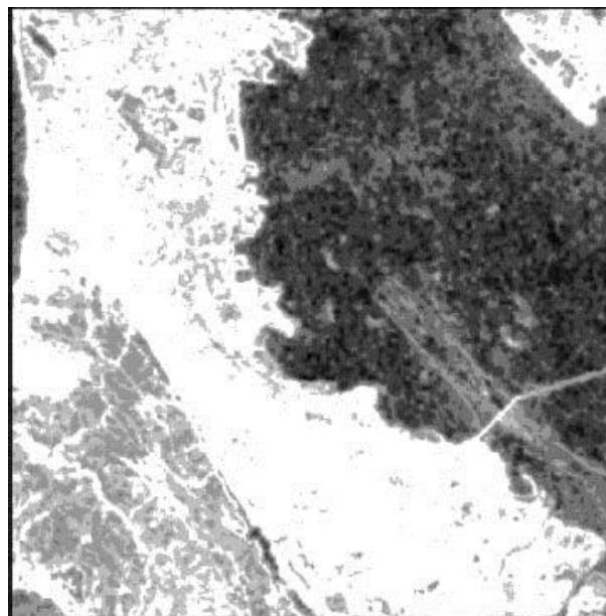


Рисунок 9.5 – Результат текстурної фільтрації на базі обчислення ентропії за формулою (9.6) у вікні 9×9 . Області двох текстур відмічені темними (вода) і світлими (суша) відтінками

Для сегментації двох зазначених типів текстур можна використати різні текстурні ознаки, наприклад набори ознак, що обчислюють за формулами (9.2)...(9.4), (9.3)...(9.5), (9.2)...(9.6), (9.4)...(9.7), (9.6)...(9.7), (9.5)...(9.6), (9.2)...(9.3), (9.3)...(9.7), (9.4)...(9.6) і ін.

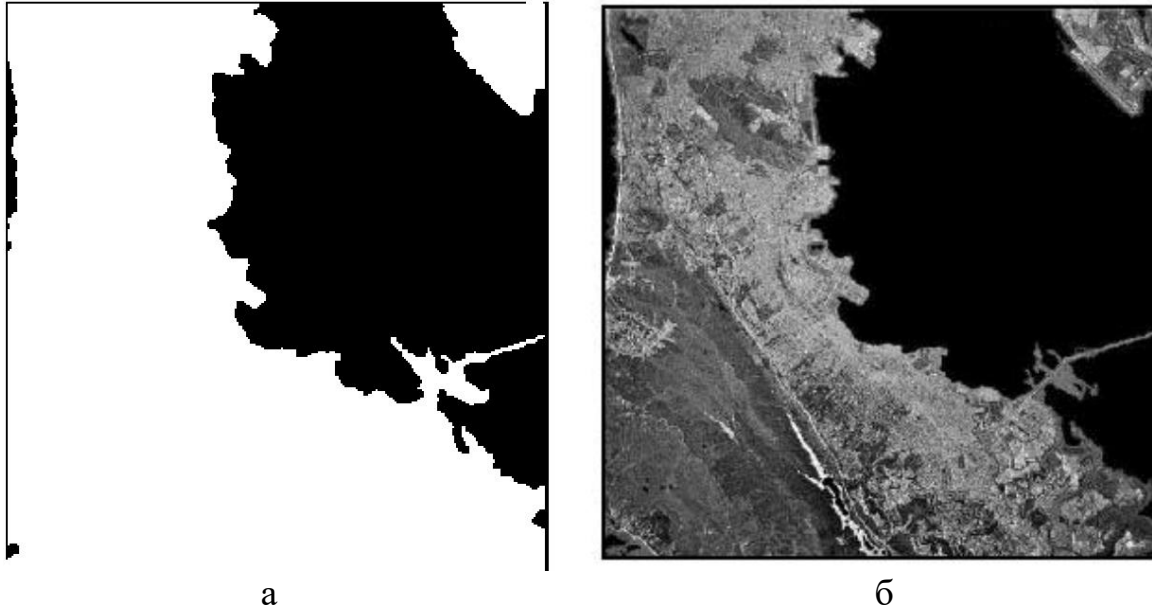


Рисунок 9.6 – Маска водної поверхні після видалення об'єктів малої площі (а) і результат сегментації суші (б)

9.3 Сегментація методом морфологічного водорозділу

У методі морфологічного водорозділу напівтонове зображення розглядається як цифрова модель місцевості, де значення яскравості трактуються як висоти щодо деякого рівня, тобто зображення – це матриця висот. Якщо на таку місцевість ллє дощ, утвориться безліч басейнів. Вода заповнює маленькі басейни, потім з переповнених басейнів виливається й басейни поєднуються в більші відповідно до висот рівня води. Місця об'єднання басейнів відзначаються як лінії водорозділу. У підсумку вся місцевість може бути затоплена. Результат сегментації залежить від моменту припинення надходження води. Якщо процес зупинити рано, зображення буде сегментовано на дрібні області, якщо пізно – на дуже великі.

У такій інтерпретації всі пікселі діляться на три типи: локальні мінімуми; на ті, що знаходяться на схилі, тобто ті, з яких вода скачується в той самий локальний мінімум; локальні максимуми, тобто ті, з яких вода скачується більш ніж в один мінімум.

При сегментації за допомогою даного методу потрібно визначити водозбірні басейни й лінії водорозділу на зображенні шляхом обробки локальних областей залежно від їх яскравісних характеристик.

При реалізації методу водорозділу обчислюється функція сегментації

(це стосується зображень, де об'єкти розміщені в темних областях і є важко помітними); обчислюються маркери переднього плану зображень на підставі аналізу зв'язності пікселів кожного об'єкта; обчислюються фонові маркери, які представляють собою пікселі, які не є частинами об'єктів; модифікується функція сегментації на підставі значень розташування маркерів тла й маркерів переднього плану.

Одним з найважливіших кроків сегментації по водорозділах є виділення на тлі зображення однорідних за яскравістю об'єктів (у вигляді плям). Області, що характеризуються малими варіаціями яскравості, мають малі значення градієнта. Тому на практиці метод сегментації по водорозділах звичайно застосовується не до самого зображення, а до його градієнтного подання.

Контрольні запитання

1. Що таке сегментація зображення ?
2. В чому полягає сегментація на основі граничної обробки ?
3. Як здійснюється обчислення порогу сегментації на два класи за гистограмою ?
4. В чому полягає сегментація зображення методом k-середніх ?
5. Які існують алгоритми сегментації зображень за принципом Вебера ?
6. Як здійснюється сегментація за кольорами шкіри ?
7. В чому полягає статистичний підхід до текстурної сегментації зображень ?
8. В чому суть сегментації зображень методом морфологічного водорозділу ?

10 ЦИФРОВА ОБРОБКА КОСМІЧНИХ ЗНІМКІВ

10.1 Особливості зображень на космічних знімках у різних ділянках спектра

Дешифрування космічних знімків, отриманих у різних діапазонах, відрізняються один від одного, як за формою, так і за змістом. Різний підхід до дешифрування знімків, у першу чергу, визначається відмінністю фізичних основ їх одержання.

Видима ділянка спектра. Основними об'єктами простежування із супутників у видимій ділянці спектра є хмари та відкриті ділянки земної поверхні на освітленій стороні Землі.

Можливість виявлення хмарності на фоні підстильної поверхні полягає в їхній різній здатності відбивати падаючу на них пряму та

розсіяну радіацію, завдяки чому на знімках, отриманих у видимій ділянці спектра, вони характеризуються різною яскравістю.

У свою чергу, яскравість поверхні об'єкта визначається інтенсивністю відбитого світлового потоку і залежить від його освітленості та коефіцієнта відбиття. Яскравість поверхні об'єктів, які мають однакові відбивні властивості, зростає при збільшенні освітленості і слабшає при її зменшенні. Природно, що яскравість об'єктів з різними відбивними властивостями різна при одній і тій же освітленості. Найважливішим фактором, що визначає освітленість поверхні об'єкта, є висота Сонця. Залежно від зміни висоти Сонця суттєво змінюється і освітленість поверхні. Серед інших не менш важливих факторів, які впливають на освітленість, необхідно вказати на прозорість атмосфери, на нахил поверхні, її шорсткість і ін.

Наявність в атмосфері аерозолів у вигляді крапель води, кристаликів льоду, пилу, димових часток і ін., помітно впливає на прозорість атмосфери. Зі збільшенням вмісту аерозолів зменшується освітленість прямим і зростає освітленість розсіяним світлом. Оскільки атмосфера практично завжди замутна, то при низькому положенні Сонця відбуваються істотні втрати в загальній освітленості горизонтальної поверхні.

Верхня межа хмар, як це показують результати літакового зондування, нечасто буває однорідною по висоті. Тут можуть спостерігатися у вигляді величезних утворень потужні купчасті і купчасто-дощові хмари, які піднімаються, вертикальна протяжність яких декілька кілометрів, а горизонтальні розміри 10...15 км і більше. Обернені до Сонця похилі ділянки хмарного покриву і особливо бічні поверхні потужних купчастих і купчасто-дощових хмар на знімках, отриманих у видимій ділянці спектра, виглядають більш яскравими, ніж горизонтальні поверхні з такими ж відбивними властивостями.

Нерівність (шорсткість) верхньої межі суцільного хмарного покриву і відкритих ділянок поверхні приводить до створення великої кількості тіней. Кількість тіней визначається шорсткістю поверхні та кутом падіння сонячних променів. Чим більше шорсткість поверхні і менше кут падіння сонячних променів, тим більша кількість тіней може бути виявлена на знімку. У якості параметра, який характеризує здатність об'єкта відбивати падаючу на нього радіацію, використовується коефіцієнт відбиття, або альbedo. Відомо, що альbedo різне для різних довжин хвиль падаючої радіації.

Установлено, що відбивні властивості хмар тісно пов'язані з їхньою поглинальною здатністю і пропущенням прямої і розсіяної сонячної радіації, а також з відбивною здатністю підстильної поверхні, яка перебуває під хмарами. Альbedo хмар в основному визначається їхньою вертикальною потужністю, водністю, фазовим станом, висотою Сонця. У

табл. 10.1 наведені середні значення альbedo різних хмар, виміряні по зображеннях, отриманих у видимій ділянці спектра.

Таблиця 10.1 – Середнє альbedo різних хмар, виміряне по зображеннях, отриманих у видимій ділянці спектра (хмарами закрито більше 80% виділеної площі)

Форма хмар	A, %
Купчасті гарної погоди над сушею	29
Пір'ясто-шаруваті над сушею	32
Пір'ясті над сушею	36
Шаруваті (тонкі)	42
Масиви шарувато-купчастих (усередині суцільного хмарногопокриву над океаном)	60
Шаруваті (потужні)	64
Шарувато-купчасті (значні над сушею)	68
Купчасті і шарувато-купчасті (значні над сушею)	69
Пір'ясто-шаруваті (потужні, з хмарністю та опадами, що пролягають нижче)	74
Купчасто-дощові (дрібні, вершини на висоті 6 км)	86
Купчасто-дощові (великі і потужні)	92

Таким чином, у відповідності із середніми значеннями альbedo (і зі зростаючими значеннями водності), по відбивних характеристиках хмари можна розташувати в наступному порядку:

- пір'ясті,
- шаруваті,
- шарувато-купчасті та купчасті,
- купчасто-дощові.

На відміну від хмар об'єкти земної поверхні звичайно не відбивають помітної кількості радіації зі своїх внутрішніх областей у вигляді розсіяного світла.

Порівняльний аналіз альbedo хмар (табл. 10.1) та підстильної поверхні (табл. 10.2) показує, що розбіжності між ними коливаються в широких межах. Якщо альbedo хмар нижнього ярусу в середньому перевищує альbedo підстильної поверхні на 50...60 %, то значення цього параметра у хмар верхнього ярусу в середньому відрізняється від альbedo підстильної поверхні, порівняно мало, на 19...20%. Найменша відмінність (2...8%) між альbedo хмар верхнього ярусу підстильної поверхні, при великих висотах Сонця.

У зв'язку із цим можуть виникати більші труднощі у виявленні хмар верхнього ярусу над відкритими ділянками поверхні Землі по знімках, отриманих у видимому діапазоні спектра. У тому випадку, коли в хмарному покриві спостерігаються просвіти або розриви, загальне альbedo видимого з штучного супутника Землі (ШСЗ) хмарного покриву

зменшується. Якщо хмарний покрив складається із дрібних хмарних елементів, розділених такими ж невеликими просвітами, то при загальному зниженні яскравості просвітина знімку видимого діапазону не виявляються.

Таблиця 10.2 – Альbedo різних видів підстильної поверхні

Вид підстильної поверхні	A, %
Водна поверхня	2...6
Чорнозем	11
Деревна рослинність	10...18
Хвойний ліс	12
Трав'яний покрив	18...20
Забруднений вологий сніг	20...30
Поверхня глинистої пустелі	29...31
Пісок білий	34...40
Лід (залежно від згуртованості та засніженості)	30...70
Сніг, який щойно випав	80 (в умовах Арктики 96...98)
Снігова поверхня Антарктиди (середнє значення за рік)	83...84

Інфрачервона ділянка спектра. Поглинання інфрачервоної (ІЧ) радіації водяною парою, вуглекислим газом, озonom і іншими компонентами атмосфери відомо. Установлено, що найбільша прозорість атмосфери спостерігається у вікнах прозорості. На сучасних ШСЗ інфрачервона зйомка проводиться в довгохвильовому вікні прозорості 8...12 мкм. Відомо, що хмари (за винятком тонких шаруватих – Сі) непрозорі для ІЧ променів з довжиною хвилі 8...12 мкм, тому радіометром вимірюється енергія, випромінювана поверхнею Землі в районах, вільних від хмар, або верхньою межею хмарного покриву. Інтенсивність радіації, яка надходить на супутник, вимірюється в межах елементарного кута зору супутника.

Отримані сигнали можуть перераховуватися в значення *еквівалентної температури* для кожного елемента відрізнєння або за допомогою особливих пристроїв перетворюватися на зображення випромінюючих поверхонь.

Інфрачервоні зображення виходять завдяки наявності *енергетичних контрастів* між підстильною поверхнею та хмарами, тому ці знімки називаються тепловими. На таких теплових інфрачервоних знімках об'єкти, які мають більш високу температуру й, отже, що випромінюють більшу кількість променистої енергії, мають більш яскравий тон зображення, ніж об'єкти з меншою температурою.

Хмари в більшості випадків мають температуру нижчу, ніж підстильна поверхня, і на стандартних теплових знімках хмарність мала б

більш темний тон зображення, ніж підстильна поверхня. Для того щоб наблизити зорове сприйняття теплових знімків до знімків у видимій ділянці спектра у практиці використовують негативні відбитки інфрачервоних знімків. На цих знімках більш холодна хмарність має світлий тон зображення, а тепла підстильна поверхня, – темний тон зображення. Найбільш яскраві ділянки на ІЧ знімках відповідають хмарам, що досягли найбільших висот, і які мають найбільш низьку температуру верхньої границі. Менш яскраві ділянки – відповідають хмарам з більш високою температурою верхньої границі або тонким хмарам верхнього та середнього ярусів і хмарним масивам із просвітами, випромінювання від яких зміщується з випромінюванням від земної поверхні.

При потужних приземних інверсіях, які виникають уночі у холодну пору року внаслідок сильного вихолоджування, верхня границя шаруватої хмарності може мати більш високу температуру, ніж підстильна поверхня. У цьому випадку хмари на ІЧ знімках виглядають темніше безхмарного простору, їх прийнято називати чорними шаруватими хмарами.

На ІЧ зображеннях поверхня суші в основному має темно-сірий тон, але внаслідок її температурної неоднорідності виражена різними півтонами. Так, низинні ділянки місцевості за тоном зображення відрізняються від ділянок, розташованих більш високо над рівнем моря. Наявність великого теплового контрасту між сушею і водною поверхнею дозволяє на знімках розрізнити берегові лінії морів, обриси озер і великих рік, зміну тону водної поверхні залежно від зміни дня і ночі.

Мікрохвильова ділянка спектра. Можливість виявлення хмарності на надвисокочастотних (НВЧ) знімках, так само як і в ІЧ діапазоні, заснована на наявності контрастів, тобто різниці в потоках випромінювання різних об'єктів. Однак наявність цих контрастів на НВЧ зображеннях обумовлена не однією температурою випромінюючих поверхонь, як на ІЧ знімках, а сукупністю багатьох факторів, і, зокрема, коефіцієнта випромінювання.

У зв'язку з тим, що спектральний коефіцієнт випромінювання суші в 2,0...2,5 рази більше коефіцієнта випромінювання морської поверхні, на НВЧ знімках завжди чітко видна границя вода-суша, тобто обриси материків і островів. Суша характеризується яскраво-білим тоном зображення, морська вода (поверхня океану) – темно-сірим. Чітко на знімках видна і границя вода-лід, тому що коефіцієнт випромінювання льоду також значно більше коефіцієнта випромінювання води. У мікрохвильовому діапазоні випромінювання підстильної поверхні поглинається хмарністю лише частково, а власне випромінювання хмар формується у всій їх товщі і залежить, крім середньої температури хмарного шару, від його фізичного стану, зокрема від вмісту в ньому води і її хімічних властивостей. Цей факт має велике значення. Енергетичні контрасти між хмарами та сушею або негативні, або (у випадку потужної

хмарності та опадів) близькі до нуля, а між морською поверхнею і хмарами завжди позитивні. Тому хмарні утворення виявляються невидимими на тлі „гарячої” суші і відмінно проглядаються над „холодною” морською поверхнею (при наявності певних енергетичних контрастів).

Оскільки кристалічні хмари прозорі для мікрохвильового випромінювання підстильної поверхні і самі не випромінюють електромагнітну енергію в цьому діапазоні, то на НВЧ знімках вони не видні, яким би щільним і потужним не був їхній шар. Інтенсивність випромінювання атмосфери і хмар, від якої залежить тон їх зображення на НВЧ знімках, на довжині хвилі 0,8 см пропорційна кількості водяної пари та вмісту води в атмосфері. Зі збільшенням вмісту води в атмосфері росте інтенсивність випромінювання. Тому чим більше вертикальна потужність рідинної частини хмари, тим світліше тон його зображення на НВЧ знімку. Самий яскравий тон зображення (за винятком материків) мають хмари, з яких випадають помірні та сильні опади. Ріст радіояскравісних температур зі збільшенням вмісту води в атмосфері відбувається лише до певної межі, після якої, як би не збільшувався зміст води в атмосфері, зростання радіояскравісних температур припиняється. На довжині хвилі 0,8 см такою межею є значення радіояскравісної температури 255...260 К, яка спостерігається при інтенсивності випадання опадів близько 10 мм за годину.

10.2 Етапи і зміст обробки космічних знімків

У задачах вивчення Землі з космосу істотна роль приділяється багатозональним космічним зйомкам. Великий обсяг інформації, який у цьому випадку надходить з ШСЗ, приводить до необхідності створення і розвитку раціональних методів і засобів її обробки. По сучасних представленнях така обробка повинна полягати, принаймні, із **чотирьох етапів**:

- службової обробки,
- просторової прив'язки,
- радіаційної корекції,
- тематичної інтерпретації.

Службова обробка повинна полягати у введенні первинної інформації в електронну обчислювальну машину (ЕОМ), її анотуванні та проведенні апаратурних корекцій. Вихідними даними для останніх служать матеріали геометричних і радіометричних калібрувань знімальної апаратури.

У результаті **просторової прив'язки** повинні бути отримані трансформовані в проекцію карти зображення із вдрукованою в них координатною сіткою. Для вивчення динаміки природних утворень необхідні матеріали повторних зйомок з високою точністю просторової

прив'язки (до декількох елементів відрізнення). Одержати таку точність можливо лише на основі спільного використання апріорних даних про траєкторію і орієнтацію ШСЗ і апостеріорних даних про геодезичні координати контурних точок знімків.

Радіаційна корекція полягає в трансформації на борті ШСЗ потоку випромінювання до деяких стандартних умов вимірів (наприклад, до однакових умов освітленості на нижній границі атмосфери).

Метою **тематичної інтерпретації** є розпізнавання на зображеннях природних об'єктів або явищ. Тематична обробка у свою чергу може містити в собі розв'язання задач конструювання нових зображень і їх інтерпретацію. При конструюванні нових зображень доцільні такі операції, як перетворення Фур'є, підкреслення контурів, складання зображень, що є комбінацією вихідних багатозональних зображень, і т.п. Основна мета таких операцій – підвищення образотворчих властивостей знімків.

Зазначені етапи обробки поєднують за змістом робіт в **2 групи**:

1) **попередня обробка** знімків полягає в корекції та покращенні супутникових зображень і містить:

- геометричну корекцію супутникових зображень,
- радіометричне калібрування знімків,
- радіометричну корекцію впливу атмосфери,
- відновлення пропущених пікселів,
- покращення зображення шляхом зміни контрасту,
- фільтрацію та ін.

Слід зазначити, що деякі методи покращення зображень (фільтрація, зміна контрасту) мають на увазі зміну спектральних характеристик знімка, тому після їхнього застосування не можна застосовувати методи тематичної обробки, які засновані на аналізі значень спектральної яскравості пікселів (класифікації, арифметичні перетворення каналів і ін.).

2) **тематична обробка** знімків містить:

- колірні перетворення,
- створення індексних зображень,
- аналіз головних компонентів,
- спектральний поділ,
- класифікацію та ін.

Необхідність **попередньої обробки** визначається особливостями дистанційного зондування.

При проведенні **просторової прив'язки** потрібно виконати операцію геометричної корекції.

Геометрична корекція – це усунення на зображенні геометричних викривлень (орторектифікація) і, власне, географічна прив'язка. Причин геометричних викривлень існує декілька і діють вони спільно. Для різних типів космічних знімків комбінація цих причин різна.

1) Кривизна поверхні Землі.

- геометричні викривлення знімків, спричинені кривизною поверхні Землі, виходять у результаті того, що точки сканованої місцевості не лежать в одній площині і спостереження ведеться не в надирі, а під кутом до поверхні Землі (рис. 10.1). Тому при віддаленні від центральної лінії сканування (де зйомка ведеться в надирі) викривлення форми та розміру об'єктів збільшується. Це призводить до викривлення форми об'єктів. Цим типом викривлення можна знехтувати, якщо кут огляду сканера невеликий,
- викривлення масштабу. Для знімків, зроблених оптико-механічним сканером (MODIS, AVHRR, ETM і MSS–Landsat, Aster (TIR)), масштаб при віддаленні від центральної лінії знімка стає дрібнішим. Тобто, якщо взяти два пікселя знімка: один із центральної області знімка, а другий з бічної, то піксель із бічної області буде містити більшу площу Землі, хоча кутовий розмір їх однаковий. Для знімків, отриманих за допомогою ПЗЗ камери (супутники SPOT, IRS, Ikonos, датчик Aster (VNIR, SVIR)) масштаб при віддаленні від центральної лінії знімка не змінюється.

2) Нерівності рельєфу.

Нерівності рельєфу викликають ті ж викривлення, що й кривизна поверхні землі, але задача усунення їх складніше через те, що форми рельєфу складніші, ніж форма Землі, яка близька до сфери. Оскільки космічні знімки роблять із великої висоти, то вплив форм рельєфу незначний, тому даний тип викривлень ураховують лише для гірських областей (рис. 10.2).

3) Обертання Землі.

Оскільки сканування Землі з космосу відбувається не миттєво, як при фотозйомці, то за час сканування сцени обертання Землі викликає зміну умов зйомки (за 1 хвилину Земля повертається на $0,25^\circ$).

4) Рух космічного апарата в процесі формування зображення.

На якість і властивості знімків впливає форма та висота орбіти супутника. Наприклад, кругова орбіта забезпечує однакову висоту зйомки земної поверхні, а, отже, для однієї й тієї ж апаратури – однакове охоплення і відрізнення знімків.

Радіометричне калібрування знімків. Знімки, спочатку одержувані із супутників, записані у вигляді так званих „сирих значень” яскравості DN (Digital Number) або, як говорять, у вигляді RAW файлів (raw – сирий). Дані в такому форматі не можна адекватно зіставляти з даними інших зйомок. Завдання радіометричного калібрування полягає в перекладі цих значень у фізичні одиниці. Формула для калібрування багатозональних знімків в оптичному діапазоні наступна:

$$V^*_{\lambda} = K_{\lambda} D_N + C_{\lambda} \quad (10.1)$$

де V^*_{λ} – енергетична яскравість для спектральної зони λ ; D_N – „сирі” значення яскравості; K_{λ} – калібрувальний коефіцієнт; C_{λ} – калібрувальна константа, відповідна до мінімальної величини реєстрованої яскравості.

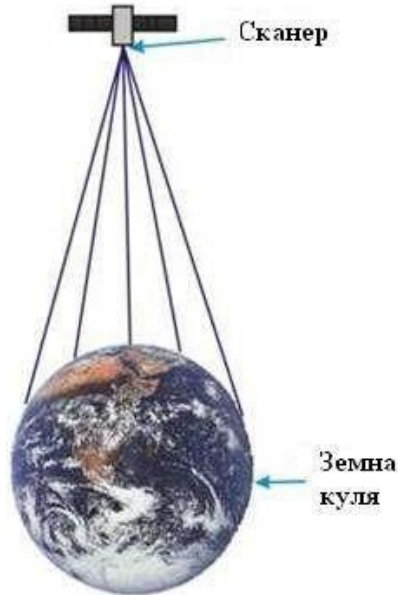


Рисунок 10.1 – Ілюстрація кривизни поверхні Землі

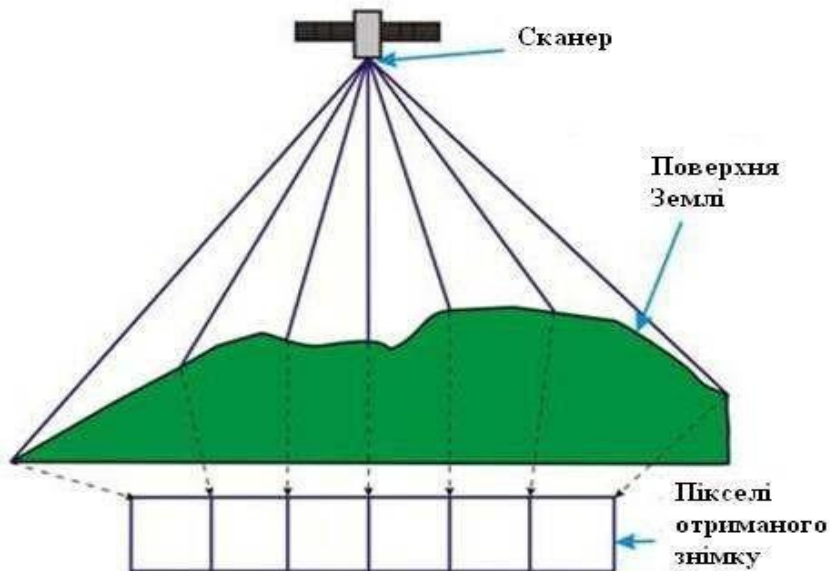


Рисунок 10.2 – Ілюстрація нерівностей рельєфу

Радіометрична корекція впливу атмосфери на знімок. При проходженні через атмосферу електромагнітні хвилі поглинаються та

розсіюються. Причиною поглинання і розсіювання є озон, водяна пара, вуглекислий газ, кисень, метан, пил, дим. Хмарність також є перешкодою при зйомці в оптичному діапазоні (рис. 10.3).



Рисунок 10.3 – Заважальні атмосферні чинники

Поглинання, розсіювання і хмарність викликають викривлення значень яскравості пікселів на знімках. Потрібна атмосферна корекція.

Способи атмосферної корекції.

- 1) Хмари та тумани добре видні на тлі води, тому що в червоній та ІЧ ділянках спектра поверхня води за своїми оптичними характеристиками близька до абсолютно чорного тіла. Тому вміст в атмосфері водяної пари і аерозолів (димув) можна оцінити по знімках, які включають ділянки морів і океанів.
- 2) Існують математичні методи побудови моделей стану атмосфери з урахуванням типів розсіювання в атмосфері, пори року, метеорологічних даних. Для уточнення таких моделей використовують наземний вимір відбивної здатності об'єктів під час польоту супутника.

Відновлення пропущених пікселів. „Пропущені пікселі” можуть виникнути під час зйомки або передачі даних. Також трапляється заміна значень яскравості цілого рядка значеннями сусіднього рядка. Такі явища можуть стати перешкодою при тематичній обробці знімка. Пропущені пікселі можна відновити шляхом інтерполяції з певною похибкою (рис. 10.4).

При обробці знімків часто **покращують зображення шляхом зміни контрасту.**

Контраст зображення – це різниця між максимальним і мінімальним значеннями яскравості.

Слабкий контраст – найпоширеніший дефект зображень (рис. 10.5).

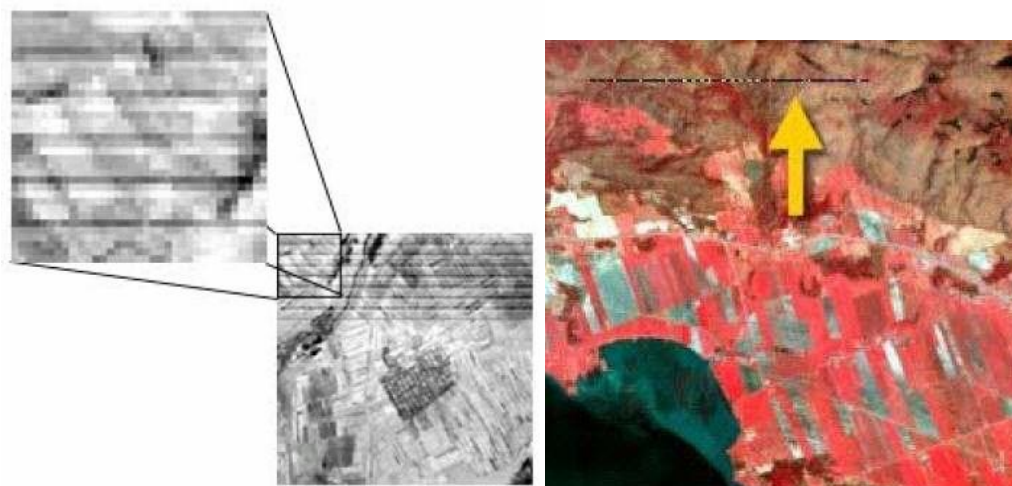


Рисунок 10.4 – Відновлення пропущених пікселів

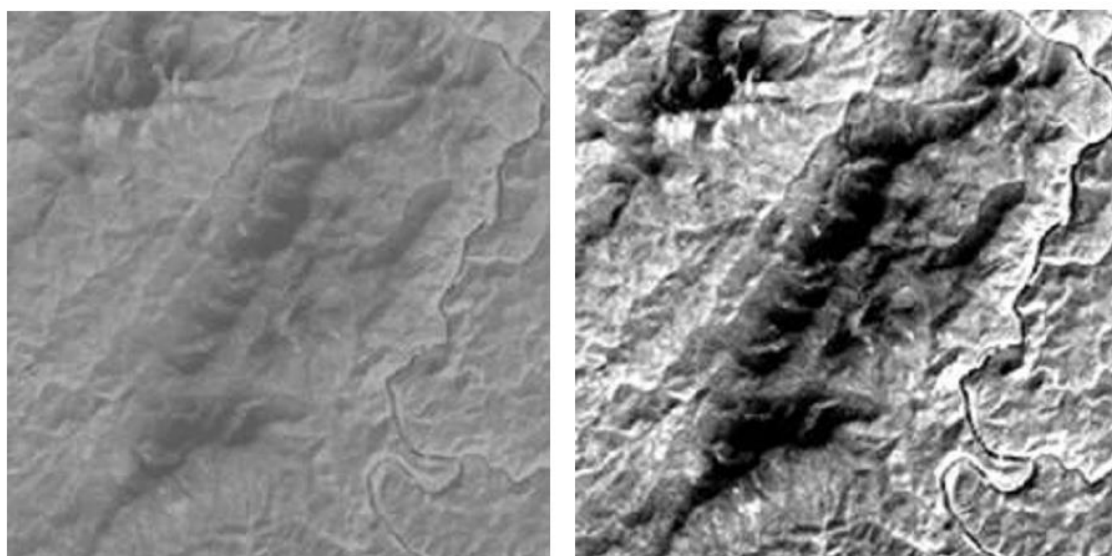


Рисунок 10.5 – Неконтрастне (ліворуч) і контрастне (праворуч) зображення

Операція підвищення контрастності знімка може бути використана при візуальному дешифруванні знімків (наприклад, щоб підкреслити границі об'єктів).

Для підвищення контрасту існує декілька методів цифрової обробки, які описані раніше:

1) Лінійне розтягування гистограми.

Усім значенням яскравості привласнюються нові значення з метою охоплення всього можливого інтервалу зміни яскравості (0, 255). Вихідний знімок: діапазон значень яскравості від 126 до 165, середнє = 140 (рис. 10.6).

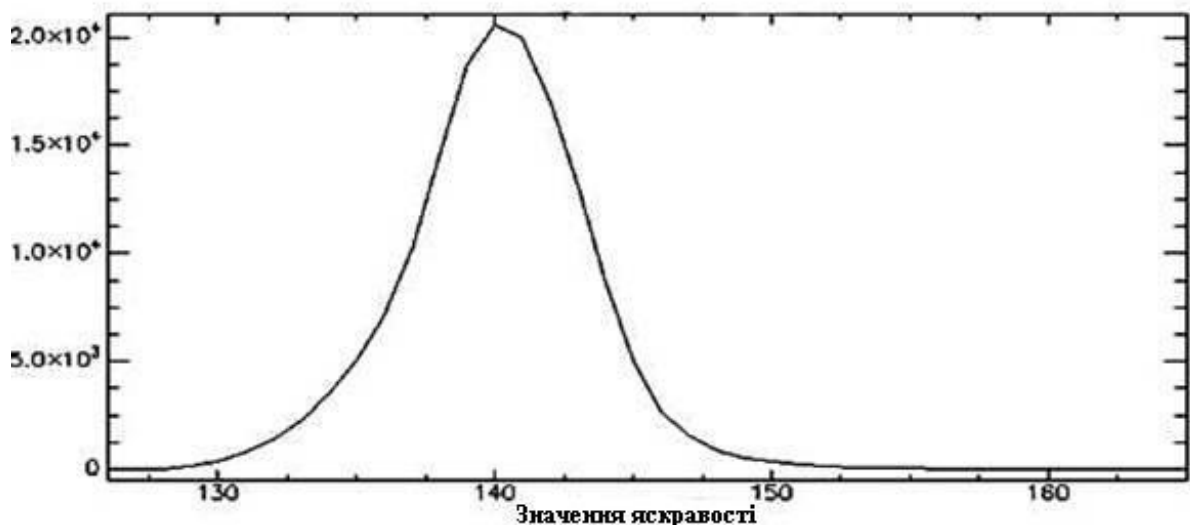
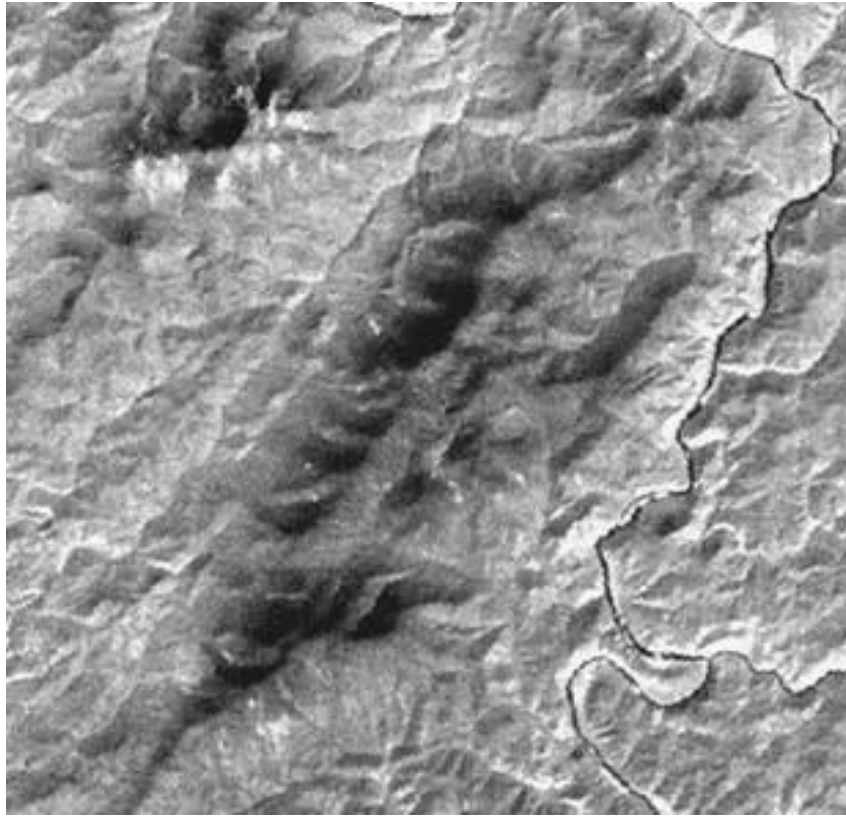


Рисунок 10.6 – Вихідний знімок для перетворення і його гистограма

Знімок після лінійного розтягування гистограми: діапазон значень яскравості від 0 до 255, середнє = 136,7 (рис. 10.7).

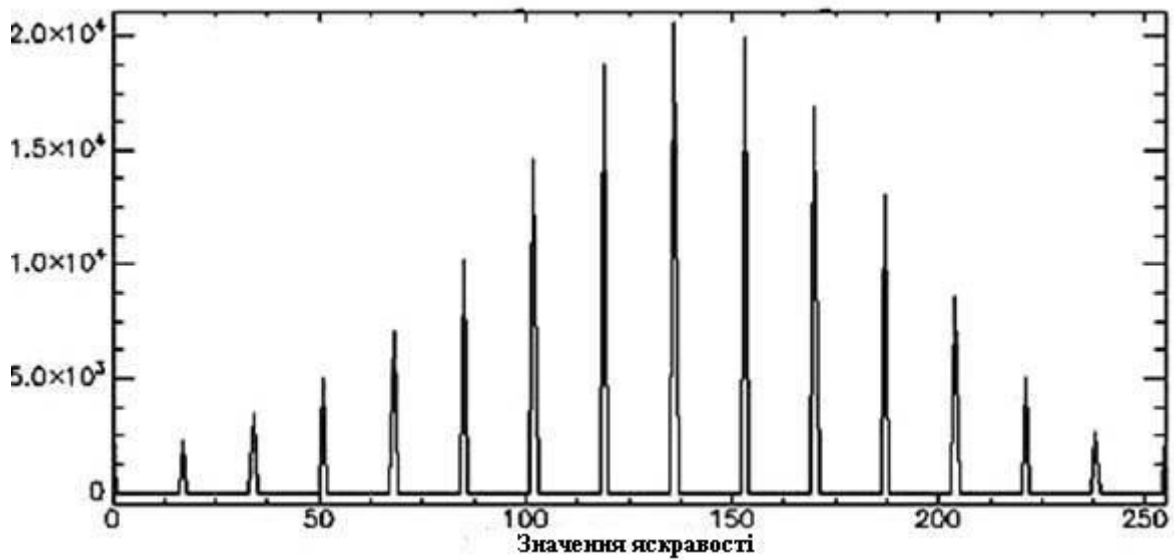
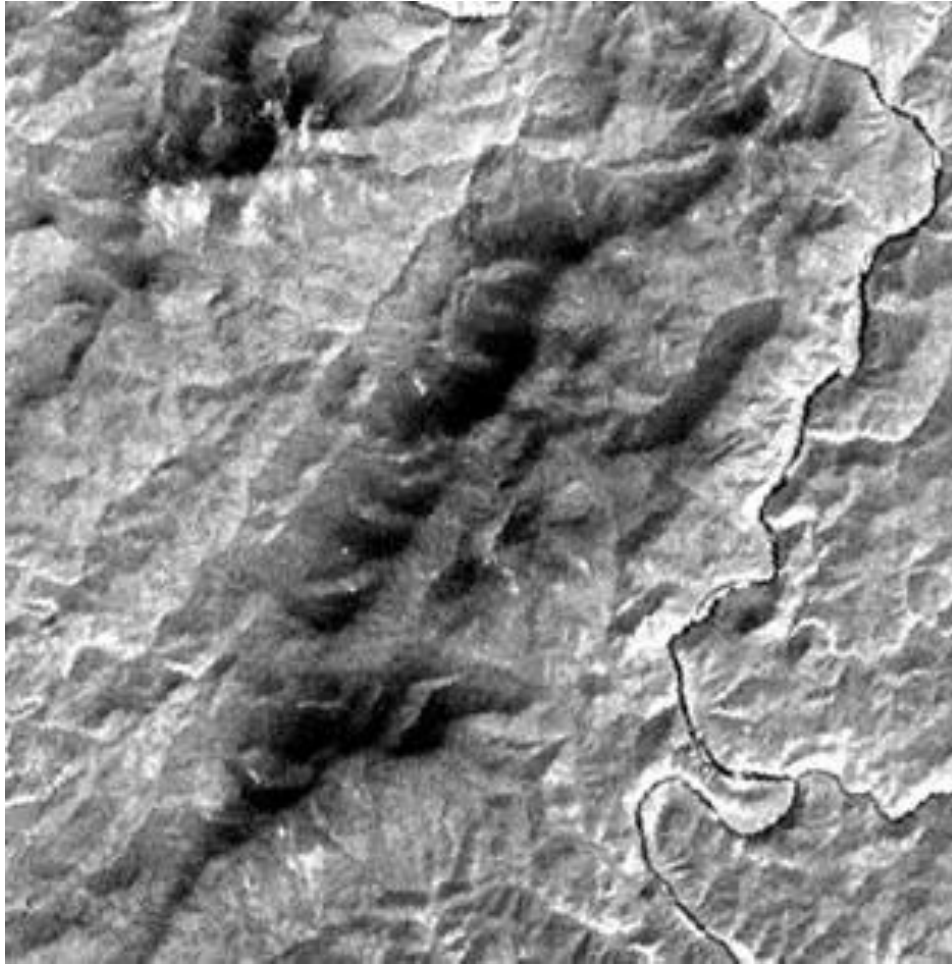


Рисунок 10.7 – Знімок після лінійного розтягування гистограми і його гистограма

2) *Нормалізація гистограми.*

На весь можливий інтервал зміни яскравості розтягується не вся

гістограма, а її найбільш інтенсивна ділянка.

Знімок після нормалізації гістограми: діапазон значень яскравості від 0 до 255, середнє = 121,5. Причому розтягнута середня, найбільш інформативна частина гістограми (рис. 10.8).

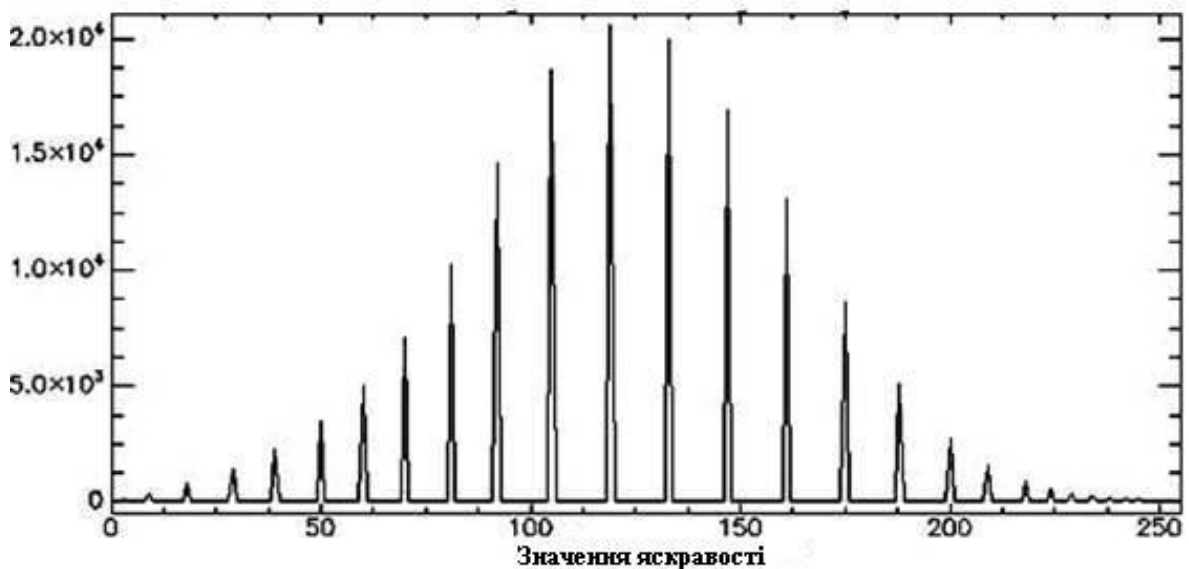
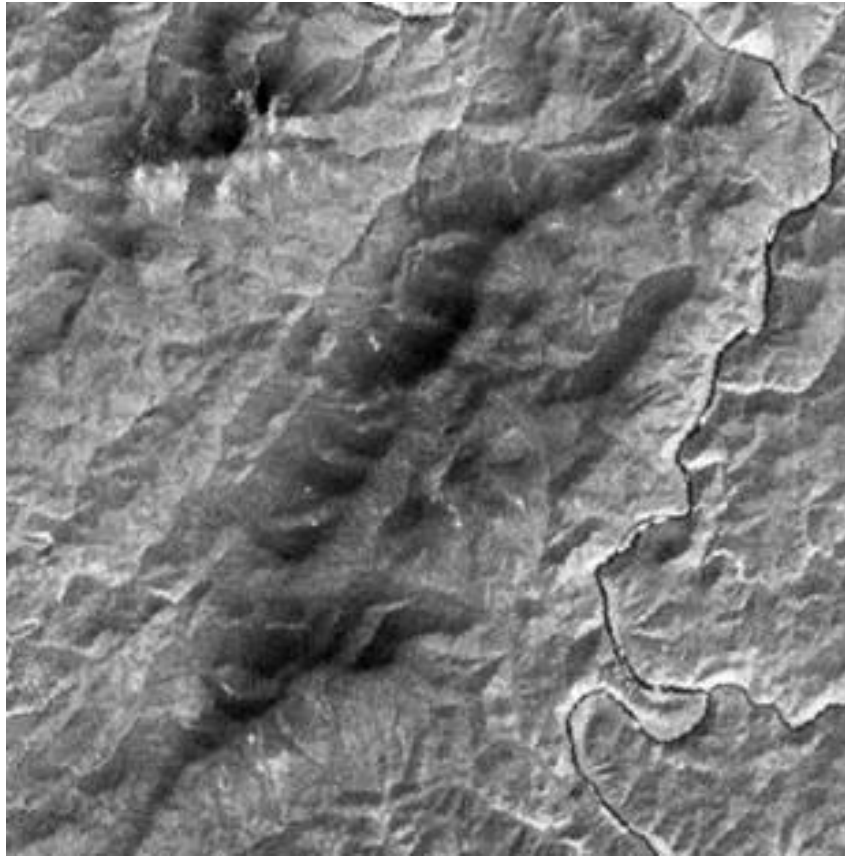


Рисунок 10.8 – Знімок після нормалізації гістограми і його гістограма

3) *Вирівнювання гістограми (лінеаризація, еквалізація).*

У процесі вирівнювання відбувається зміна значень яскравості пікселів таким чином, щоб для кожного рівня яскравості була однаковою або близькою кількістю пікселів.

Знімок після вирівнювання гистограми: діапазон значень яскравості від 0 до 255, середнє = 115,9 (рис. 10.9).

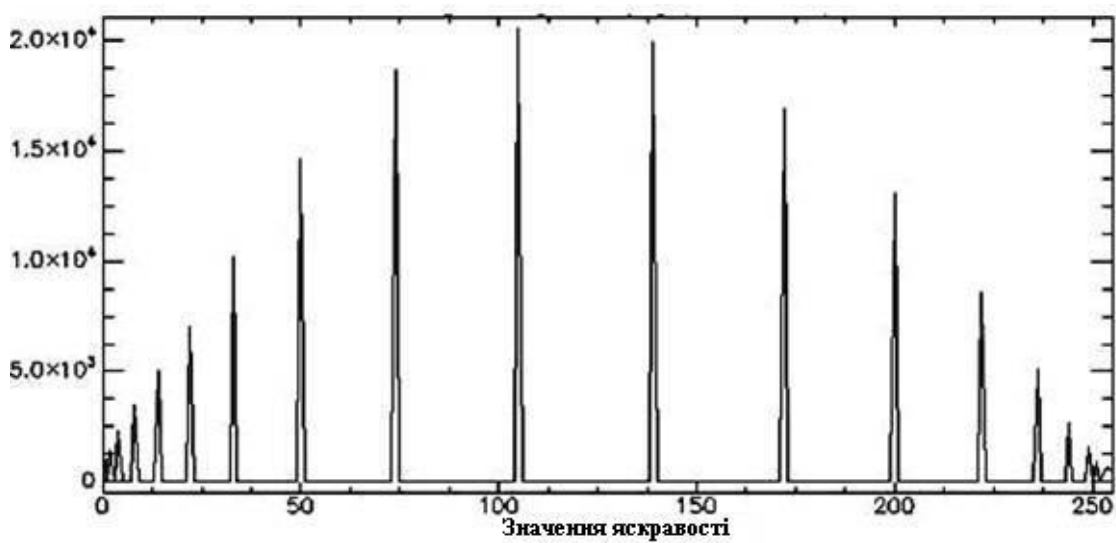


Рисунок 10.9 – Знімок після вирівнювання гистограми і його гистограма

При покращенні зображень часто застосовують *методи фільтрації*.

Фільтрація – це перетворення, яке дозволяє підсилити відтворення тих або інших об'єктів, придушити небажану вуаль на зображенні, усунути інші випадкові перешкоди (шум).

Один з найпростіших способів фільтрації – перетворення в ковзному вікні. При такому перетворенні перераховуються значення яскравості всіх пікселів зображення. Перерахування відбувається для кожного пікселя в такий спосіб: якщо даний піксель є центральним у вікні, яке „рухається” по знімку, йому дається нове значення, яке є функцією від значень оточуючих його у вікні пікселів. Розмір вікна може бути, наприклад, 3×3 або 5×5 пікселів. Щораз вікно зміщується на 1 піксель і рухається доти, поки не пройде весь знімок. Для всіх пікселів вікна встановлюються вагові коефіцієнти, виходячи із мети дешифрування.

Якщо вагові коефіцієнти всіх пікселів у вікні рівні, зображення буде згладжене (рис. 10.10).

0.1	0.1	0.1
0.1	0.1	0.1
0.1	0.1	0.1

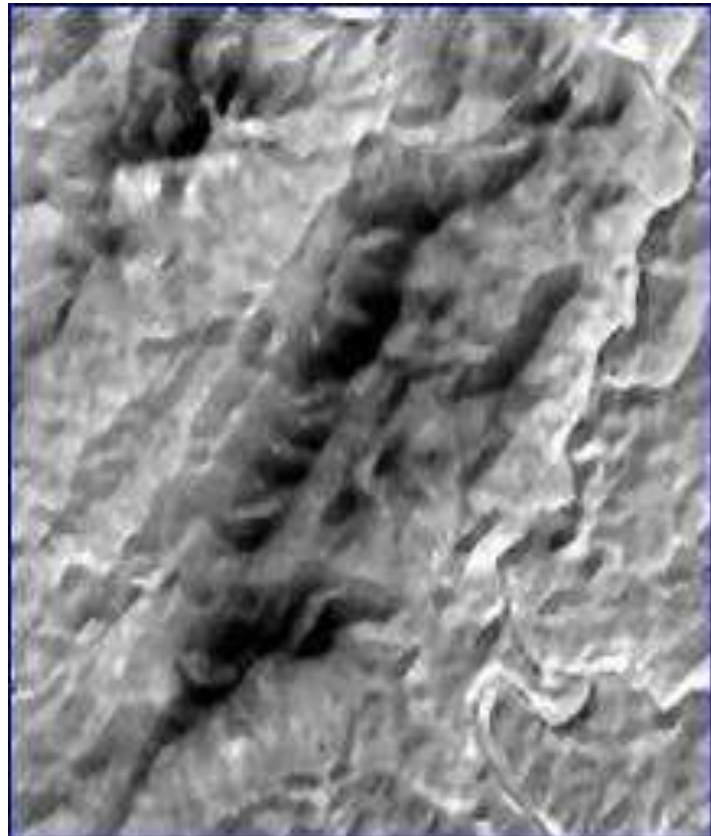


Рисунок 10.10 – Згладжений знімок

Якщо ваговий коефіцієнт центрального пікселя у вікні значно більший від інших, то зображення стане більш різким (рис. 10.11).

Якщо вагові коефіцієнти вертикально розташованих пікселів у вікні будуть значно більші від інших, то на зображенні виділяться вертикальні

лінії (рис. 10.12).

Методи інтерпретації зображень прийнято підрозділяти на:

- візуальні,
- візуально-інструментальні,
- машинні.

-1	-1	-1
-1	3	-1
-1	-1	-1



Рисунок 10.11 – Підвищення різкості знімка

-1	3	-1
-1	3	-1
-1	3	-1

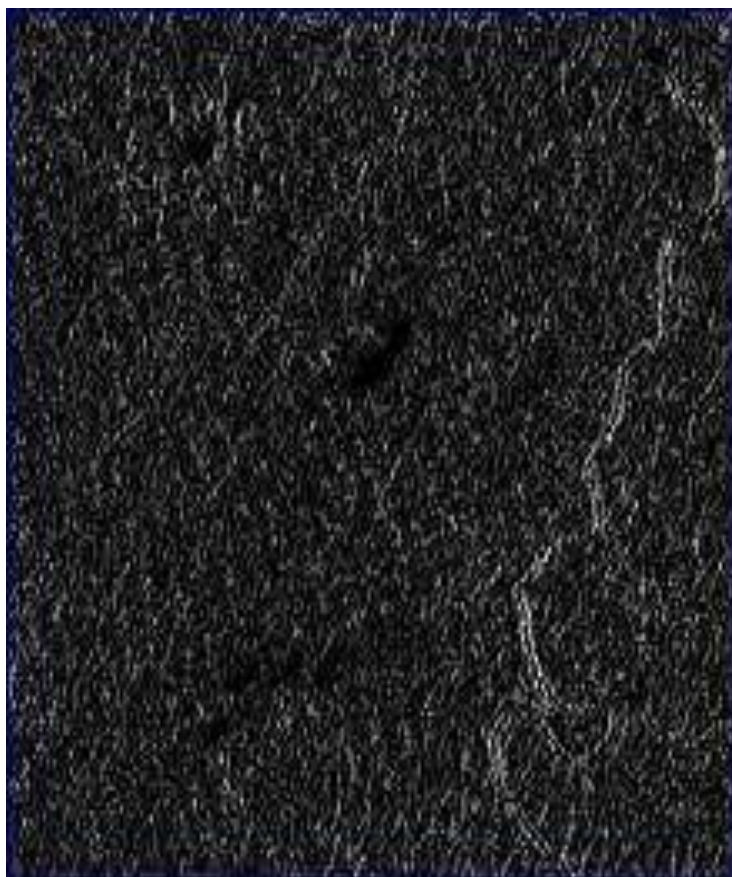


Рисунок 10.12 – Виділення ліній на знімку

У кожному разі розробка методів інтерпретації починається з вибору і вивчення *дешифрувальних ознак* природних утворень. Такими ознаками можуть бути: обмірювані радіаційні характеристики (або їх функції), рисунок зображення, непрямі ознаки, які впливають із ландшафтних особливостей зображеної території. Вибір дешифрувальних ознак здійснюється на основі відбивних характеристик систем природних утворень і умов трансформації цих характеристик атмосферою.

Безпосередньо інтерпретація містить у собі так звану *задачу ухвалення рішення*. При візуальному дешифруванні ухвалення рішення ґрунтується головним чином на знаннях і досвіді фахівця. При інтерпретації за допомогою ЕОМ звичайно використовуються критерії, які впливають із теорії оптимальних статистичних рішень. Для визначення параметрів (метеорологічних, екологічних, біометричних показників, швидкості стоку і т.п.) результати дешифрування повинні бути піддані спеціальній обробці.

Успіх інтерпретації багатозональних знімків суттєво залежить від оптимального вибору спектральних діапазонів зйомки відповідно до відбивних властивостей шуканих природних об'єктів і явищ. У той же час вибір цих діапазонів повинен базуватися на розв'язанні задач інтерпретації. *Дешифруванням* називається процес розпізнавання

хмарності, снігу, льоду і об'єктів земної поверхні по космічних знімках. При дешифруванні необхідно, за можливістю, одержати оптимальне число різноманітних характеристик, що дозволяють дати правильну якісну та кількісну оцінки об'єкта, зображеного на знімку.

Тематична обробка космічних знімків – це процес дешифрування аборозпізнавання об'єктів і явищ на космічних знімках.

Велику роль при інтерпретації зображень відіграє **колір**. Кольорове зображення на моніторі комп'ютера виходить шляхом додавання трьох основних кольорів. За основні прийняті кольори, відповідні до монохроматичного випромінювання із довжиною хвилі 0,7 мкм (червоний – R); 0,5461 мкм (зелений – G); 0,4358 (синій – B). Це відоме вже представлення колірної моделі RGB (рис. 10.13).

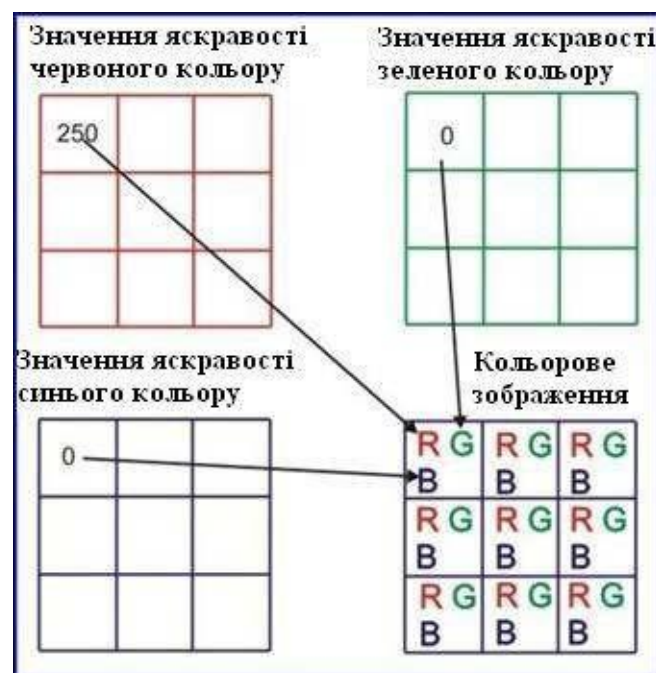


Рисунок 10.13 – Варіант представлення колірної моделі RGB

У файлі кольорового зображення для кожного пікселя записано три числа, які означають інтенсивності трьох основних кольорів, діапазон значень від 0 до 255. Практично будь-який колір, видимий людиною, може бути представлений як сума трьох основних кольорів.

Якщо значення яскравості $R=255$, $G=0$, $B=0$, то піксель буде пофарбований у червоний колір.

Якщо значення яскравості рівні між собою $R=G=B$, піксель буде пофарбований у сірий колір.

Якщо значення яскравості R, G, B не рівні між собою, то залежно від значень яскравості R, G, B піксель буде пофарбований у певний колір (рис. 10.14).

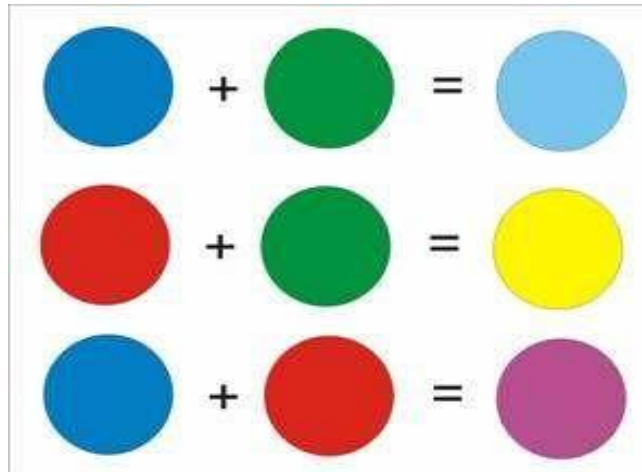


Рисунок 10.14 – Додавання кольорів у колірній моделі RGB

Застосовуються й інші методи представлення кольорових зображень.

Метод колірних перетворень. Якщо вивести на екран дисплея зображення одного каналу багатоспектрального знімка, то воно буде пофарбовано в сірі тони (напівтонове зображення). Щоб одержати кольорове зображення потрібно скласти разом три канали багатоспектрального знімка, один з яких буде червоний (R), інший зелений (G), третій синій (B) (рис. 10.15).

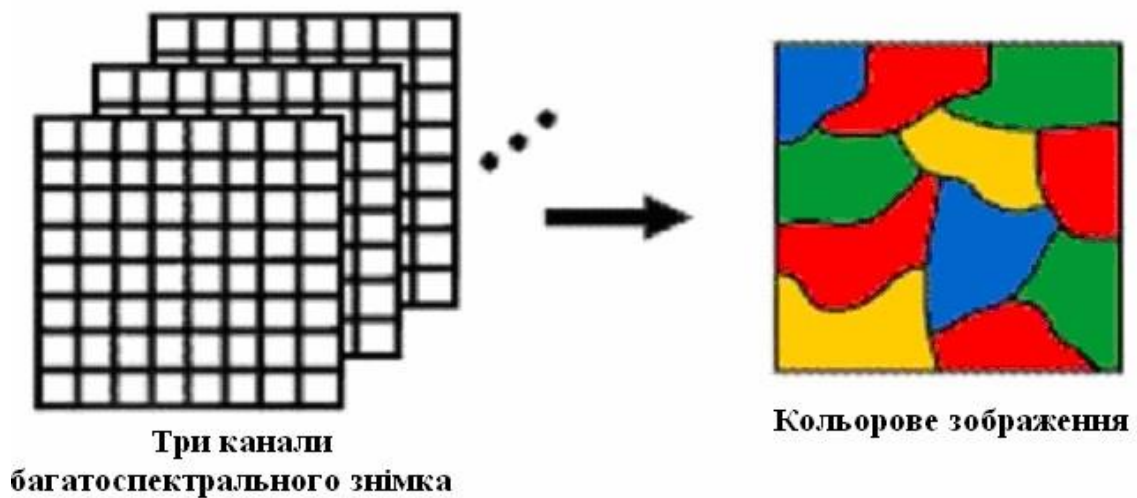


Рисунок 10.15 – Кольорове зображення

Часто фарбують зображення в натуральні та неправильні кольори або псевдокольори. Якщо в канали R, G, B зображення підставлені R, G, B канали знімка, то зображення представлено в натуральних кольорах (рис. 10.16).



Рисунок 10.16 – Приклад фарбування зображення в натуральні кольори (коментарі в табл. 10.3)

Таблиця 10.3 – Характеристика зображення (рис. 10.16)

Канали зображення	Спектральні канали знімка	Колір води на зображенні	Колір рослинності на зображенні	Колір ґрунту на зображенні
R	Червоний	Синій	Зелений	Коричнево-сірий
G	Зелений			
B	Синій			

Якщо в канали R, G, B зображення підставлені інші канали знімка, то зображення представлено в неправильних або псевдокольорах (рис. 10.17). Механізм добору кольору зображення наступний. Якщо в канал зображення R підставити канал знімка, у якому значення яскравості пікселів високі, а в канали G і B канали з низькими значеннями яскравості, то зображення буде пофарбовано переважно в червоні тони і так далі.

Таблиця 10.4 – Характеристика зображення (до рис. 10.17)

Канали зображення	Спектральні канали знімка	Колір води на зображенні	Колір рослинності на зображенні	Колір ґрунту на зображенні
R	Близький ІЧ	Чорний	Червоний	Сіро-блакитний
G	Червоний			
B	Зелений			

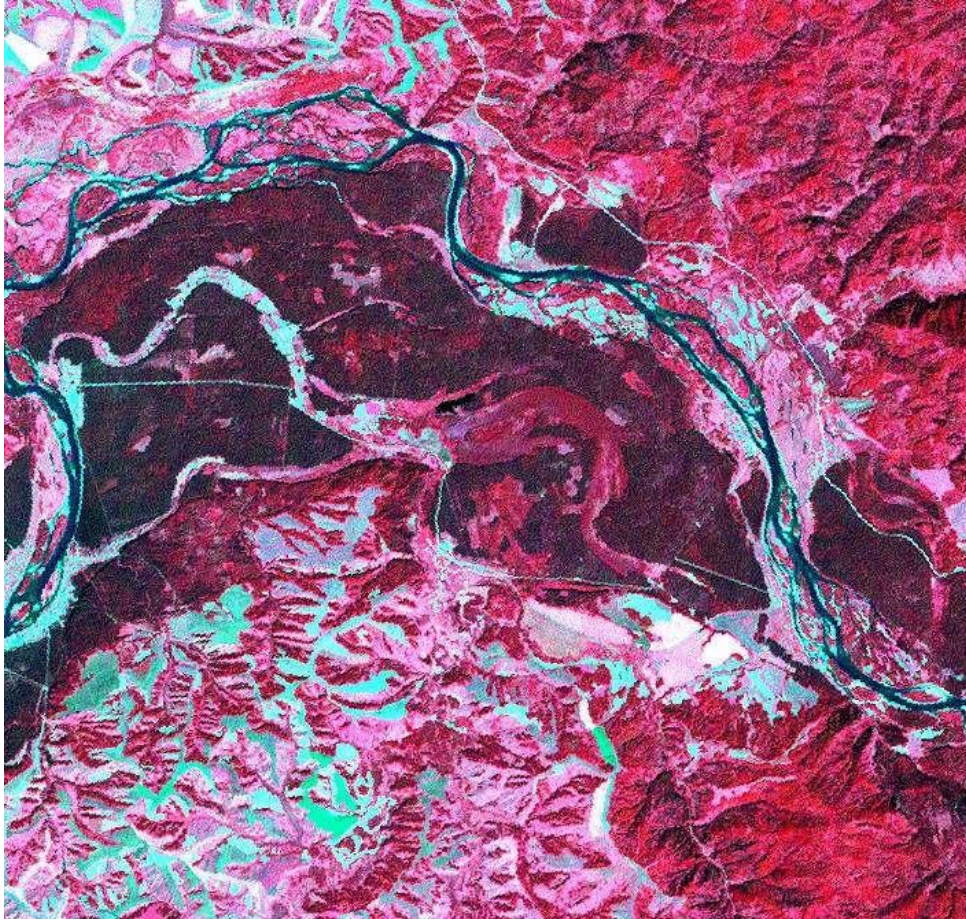


Рисунок 10.17 – Приклад фарбування зображення в неправильні (псевдо) кольори (коментарі в табл. 10.4)

Сутність об'єктів доцільно визначати на зображеннях у натуральних кольорах, а розділяти та оконтурювати об'єкти зручніше на зображеннях у псевдокольорах. Правильно підібрана шкала дозволяє операторові візуально виділяти об'єкти, які мало відрізняються за яскравістю на напівтоновому зображенні.

Індексні зображення. Для одержання індексного зображення, значення яскравості кожного пікселя обчислюється шляхом застосування арифметичних операцій над значеннями яскравості цього пікселя з різних каналів знімка. При вивченні об'єктів по багатозональних знімках часто важливі не абсолютні значення, а характерні співвідношення між значеннями яскравості об'єкта в різних спектральних зонах. На таких зображеннях більш яскравіше і контрастніше виділяються шукані об'єкти в порівнянні з вихідним знімком (рис. 10.18).

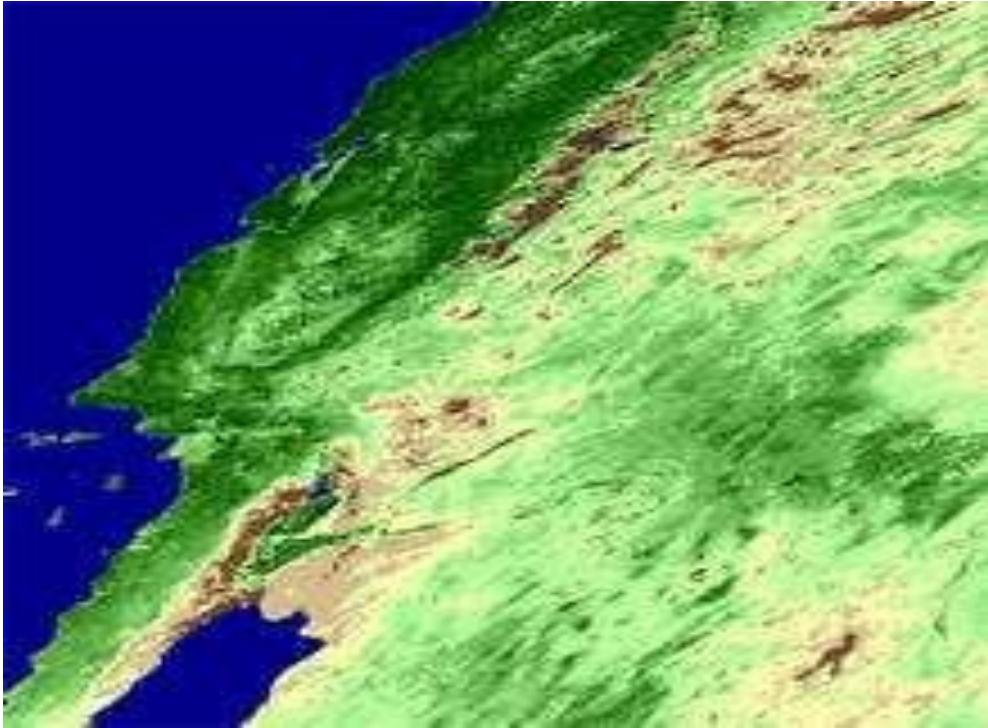


Рисунок 10.18 – Приклад індексного фарбування зображення (вегетаційний індекс)

Як приклад наведемо декілька індексів, їх призначення і порядок розрахунків:

- індекс вмісту оксиду заліза, призначений для виявлення вмісту окислів заліза, являє собою відношення значень яскравості в червоному (Ч) каналі до значень яскравості в синьому каналі (С) – $Ч / С$;

- індекс вмісту глинистих мінералів, призначений для виявлення вмісту глинистих мінералів, являє собою відношення значень яскравості в межах середнього інфрачервоного каналу (СІЧ) – $СІЧ1 / СІЧ2$, де СІЧ1 це діапазон від 1,55 до 1,75 мкм, СІЧ2 – від 2,08 до 2,35 мкм;

- індекс вмісту залізистих мінералів, призначений для виявлення вмісту залізистих мінералів, являє собою відношення значення яскравості в середньому інфрачервоному (СІЧ1 – див. вище) каналі до значення яскравості в близькому інфрачервоному каналі (БІЧ) – $СІЧ1 / БІЧ$;

- індекс червонокольоровості (RI), призначений для виявлення вмісту оксиду заліза в ґрунті, заснований на відмінності відбивної здатності червоноколірних мінералів у червоному (Ч) і зеленому (З) діапазонах – $RI = (Ч - З) / (Ч + З)$;

- водний індекс (WI), застосовується для визначення вмісту води в рослинності по гіперспектральних знімках – $WI = 0,90 \text{ мкм} / 0,97 \text{ мкм}$;

- нормалізований диференціальний індекс снігу (NDSI), використовується для виділення територій, покритих снігом (для снігу $NDSI > 0,4$), NDSI – це відносна величина, яка характеризується

відмінністю відбивної здатності снігу в червоному (Ч) і короткохвильовому інфрачервоному (КІЧ) діапазонах – $NDSI = (Ч - КІЧ) / (Ч + КІЧ)$;

– нормалізований диференціальний вегетаційний індекс (NDVI), показує наявність і стан рослинності. Хлорофіл листя рослин відбиває випромінювання в близькому інфрачервоному (БІЧ) діапазоні електромагнітного спектра й поглинає в червоному (Ч). Відношення значень яскравості в цих двох каналах дозволяє чітко відокремлювати і аналізувати рослинні від інших природних об'єктів – $NDVI = (БІЧ - Ч) / (БІЧ + Ч)$. Значення NDVI варіюють у межах від -1 до 1:

густа рослинність –	0,7;
вираджена рослинність –	0,5;
відкритий ґрунт –	0,025;
хмари –	0;
сніг і лід –	-0,05;
вода –	-0,25;
штучні матеріали (бетон, асфальт) –	-0,5.

Існують різні способи *класифікації*. Розглянемо один зі способів, а саме, метод аналізу головних компонентів.

Аналіз головних компонентів – це метод аналізу багатоспектральних корельованих даних. Поняття корельовані дані означає, що при зростанні значення яскравості пікселів в одному спектральному каналі зростають значення яскравості і в інших спектральних каналах.

На рис. 10.19 показаний приклад двовимірного простору спектральних ознак, при зростанні значення в першому каналі зростає значення в другому, це означає високу кореляцію між цими каналами.

Видно, що область розподілу значень розташована під кутом до осей графіка, тому ні по одній із цих осей не відображається весь діапазон значень об'єкта. Це значить, що при синтезі кольорового зображення знімок буде бідний кольорами. Наприклад, усі види рослинності будуть зображені близькими нерозрізненними відтінками кольору. Таке розташування значень яскравості характерно для більшості природних об'єктів (рослинності, ґрунтів, гірських порід).

Якщо осі координат розгорнути так, щоб одна з них йшла паралельно полю розподілу значень, а друга розташовувалася ортогонально, то уздовж кожної з осей діапазон значень буде максимальним, що збільшує можливості дешифрування знімка (рис. 10.20...10.22).

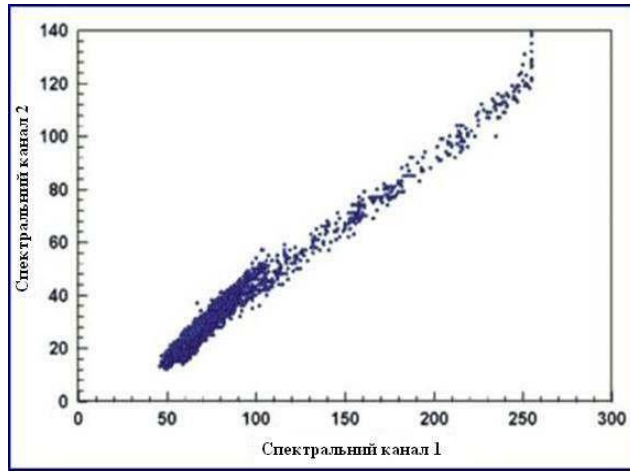
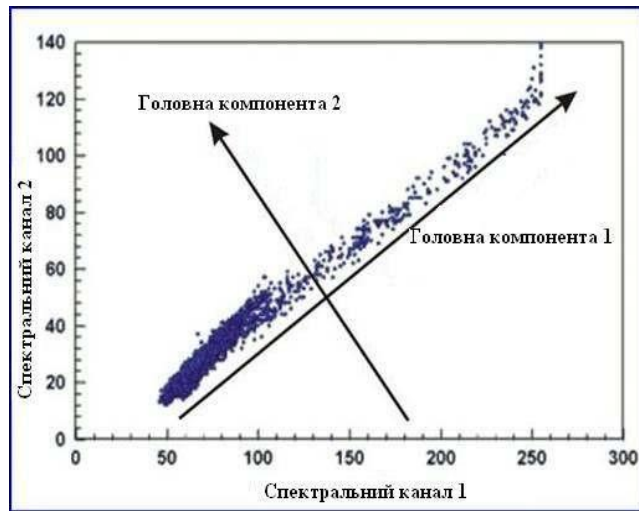
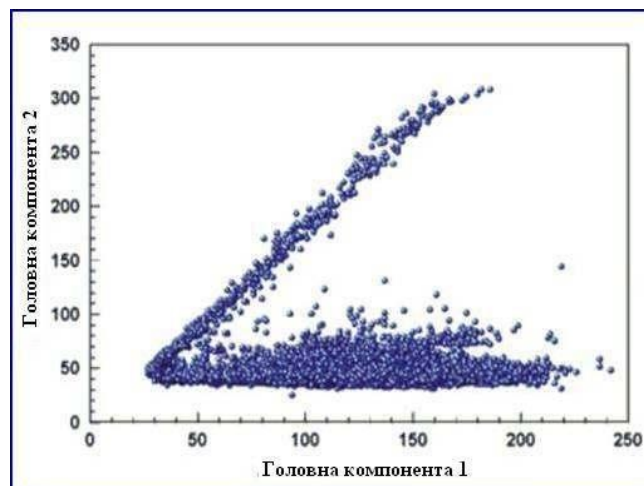


Рисунок 10.19 – Поняття корельованих даних



а



б

Рисунок 10.20 – Поняття головних компонентів



Рисунок 10.21 – Фрагмент знімка, представленого в неправильних (псевдо) кольорах



Рисунок 10.22 – Той самий фрагмент знімка, оброблений за методом ГОЛОВНИХ КОМПОНЕНТІВ

Можливості методу аналізу головних компонентів полягають у наступному:

1) якщо знімок містить більше трьох спектральних каналів, можна створити кольорове зображення із трьох головних компонентів, оскільки в типовому багатозональному зображенні звичайно перші дві або три компоненти здатні описати практично всю мінливість спектральних характеристик. Інші компоненти найчастіше піддані шумовим впливам. Відкидаючи ці компоненти можна зменшити обсяг даних без помітної втрати інформації;

2) якщо об'єкти малого розміру і низького контрасту погано дешифруються на вихідних знімках, вони часто добре виявляються на зображеннях окремих спектральних компонентів;

3) перетворення даних за методом головних компонентів проводять для серії різночасних знімків, наведених у єдину систему координат, для виявлення динаміки, яка яскраво проявляється в одній або двох компонентах.

Метод спектрального поділу. Один піксель знімка може відображати від декількох квадратних метрів до тисяч квадратних метрів поверхні Землі, і містити інформацію не про один об'єкт, а про групу об'єктів, які розташовані на відповідній території. Метод спектрального поділу застосовують для розпізнавання на знімках об'єктів, розмір яких значно менший від розміру пікселя. Суть методу полягає в наступному: змішані спектри аналізують, порівнюючи їх з відомими чистими спектрами, наприклад, зі спектральних бібліотек чистих матеріалів (рис. 10.23). Відбувається кількісна оцінка співвідношення даного відомого (чистого) спектра і домішок у спектрі кожного пікселя. Після виконання такої оцінки можна одержати зображення, розфарбоване так, що колір пікселя буде означати, який компонент переважає в спектрі цього пікселя (рис. 10.24).

При *дешифруванні* застосовують методи *класифікації*.

Класифікація – це комп'ютерне дешифрування знімків або процес автоматизованого поділу всіх пікселів знімка на групи (класи), які відповідають різним об'єктам.

По використанню статистичних параметрів розрізняють параметричні та непараметричні класифікації. При виконанні параметричних класифікацій передбачається, що значення яскравості пікселів в отриманих класах розподілені за нормальним (гаусовим) законом. Крім того, можуть бути враховані різні статистичні параметри, і чим більше їх враховано, тим точніше класифікація. Непараметричні класифікації не враховують розподіл значень спектральної яскравості усередині класу. Їх використовують або для вкрай простих випадків, або для класифікації складних неоднорідних об'єктів.

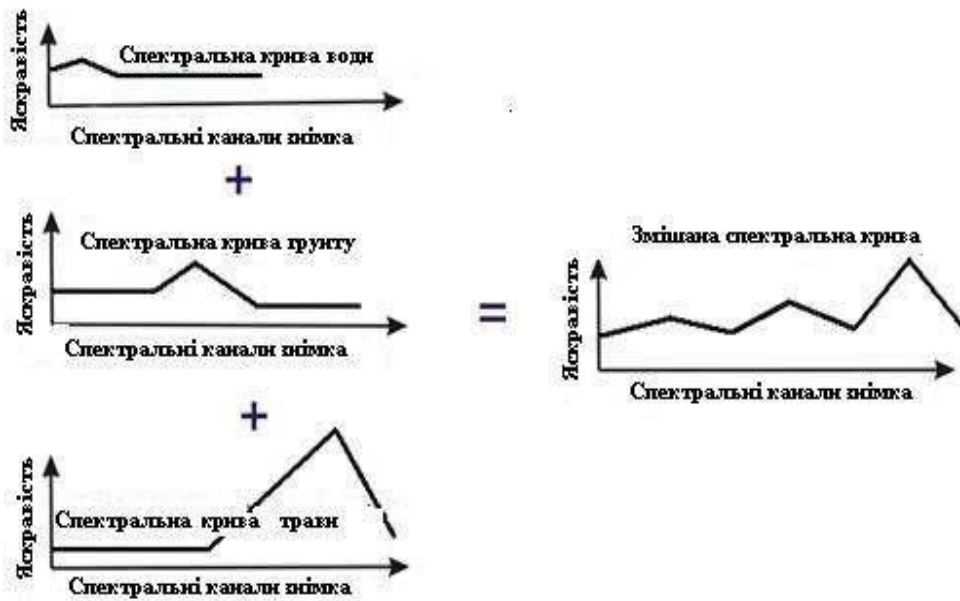


Рисунок 10.23 – Спектральна бібліотека чистих матеріалів і змішаний спектр

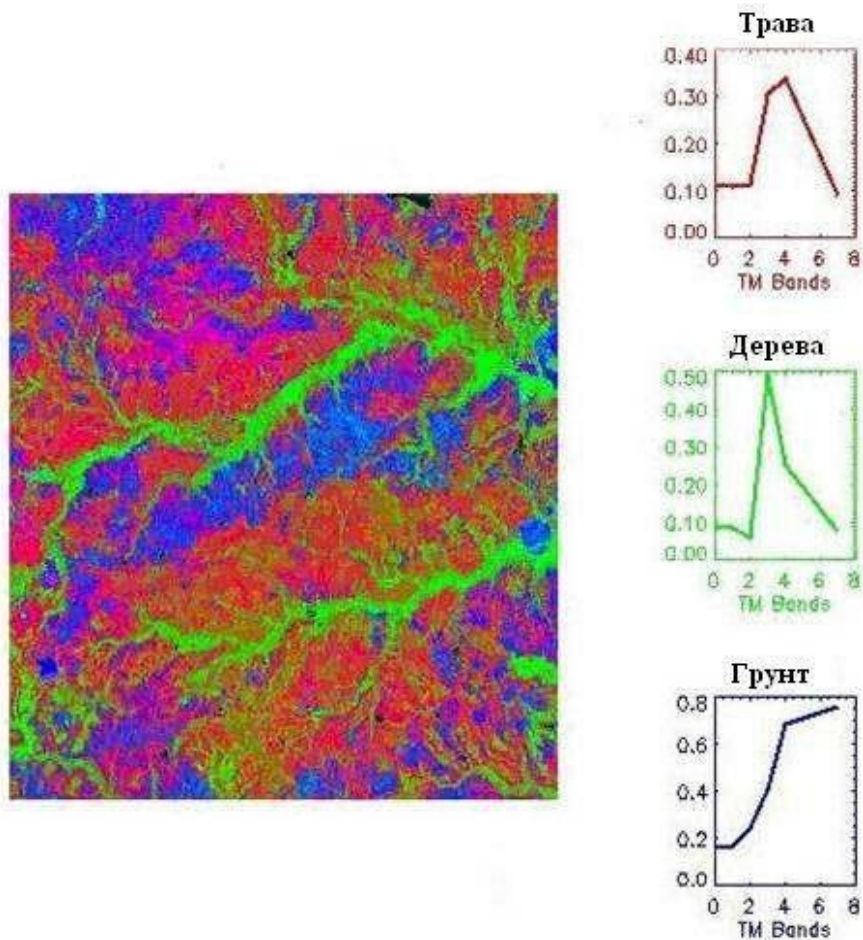


Рисунок 10.24 – Застосування методу спектрального поділу

10.3 Програмні засоби цифрової обробки супутникових зображень

Matlab (від англ. MatrixLaboratory – матрична лабораторія) – це пакет прикладних програм для вирішення завдань технічних обчислень, безліч спеціалізованих бібліотек й однойменна мова програмування. Мова Матлаб – інтерпретатор, він реалізований для більшості сучасних операційних систем.

Використання Matlab прискорює дослідницьку роботу в кілька разів, дозволяє зосередитися саме на перевірці алгоритму, а не на його реалізації, помилки реалізації відслідковуються легше.

У Matlab можна створювати спеціальні бібліотеки функцій (toolbox), що розширюють його функціональність. Наприклад, у базовій версії є бібліотеки Digital Signal Processing Toolbox, Image Processing Toolbox, Wavelet Toolbox, Filter Design Toolbox, які використовуються для цифрової обробки сигналів і зображень.

Пакет Image Processing Toolbox звільняє розроблювача від пошуку рішень, кодування й налагодження алгоритмів, дозволяючи зосередити зусилля на рішенні основного практичного завдання. Matlab і пакет Image Processing максимально пристосовані для навчання, розвитку нових ідей й їхньої перевірки.

Image Processing Toolbox (<http://matlab.exponenta.ru/imageprocess/book5/index.php>) являє собою бібліотеку основних функцій обробки зображень. У їхнє число входять:

- читання й запис різних форматів зображень, відеопотоків зображень;
- засоби побудови інтерактивного графічного інтерфейсу користувача (GUI Builder) для прискорення розробки GUI-шаблонів, які дозволяють налаштовувати його під завдання користувача;
- перетворення в різні типи зображень і колірні моделі, зміна палітри зображень;
- геометричні перетворення;
- просторові перетворення зображень;
- обробка виділеної області зображення;
- морфологічні операції;
- ковзна й блокова обробка;
- лінійна й нелінійна фільтрація;
- поліпшення якості зображень;
- відновлення зображень;
- аналіз зображень.

Більшість функцій бібліотеки представлені в системі у вигляді *m-*

файлів. В них оптимально реалізовані найбільше часто використовувані алгоритми обробки зображень. Коди цих функцій відкриті.

Відмітною рисою Matlab є орієнтація на матричні операції. Навіть константа трактується як матриця розміром 1×1 . Така орієнтація дуже зручна для обробки зображень.

Переваги системи Matlab:

- велика кількість вбудованих операцій при роботі з матрицями. Це спрощує як програмування різних завдань, пов'язаних з рішенням систем лінійних алгебраїчних рівнянь, так і роботу із зображеннями;
- багатий набір бібліотек, які розвивають постійно, різних математичних методів (чисельні методи, диференціальні рівняння, обробка сигналів, обчислювальна геометрія й т.д.);
- в інтернеті доступна безліч додаткових бібліотек і функцій для Matlab;
- багато функцій описані в книгах;
- більші можливості по швидкій і простий (для розроблювача) візуалізації різних даних (двовірних і тривірних графіків функцій, множин і т.д.);
- є інтерпретатор з постійно доступним командним рядком (у потрібний момент програму можна призупинити й переглянути дані, отримані на певному кроці, відобразити їх графічно, модифікувати, виправити параметри викликаної функції й т.д.), що полегшує налагодження й розуміння помилок у програмі;
- існує можливість використовувати Matlab у програмах, написаних на C++, і навпаки, у Matlab можна використовувати бібліотеки, написані на C++. У такий спосіб можна тестувати алгоритми, написані на C++ за допомогою середовища Matlab і використовувати у своїх програмах засоби Matlab (візуалізацію даних й інші рутинні операції). Наприклад, додавати у свою програму 3D-графіку (для налагодження) досить трудомістко, а підключившись до Matlab є можливість передати туди дані й побудувати необхідні графіки.

Приведемо приклад коду найпростішої програми обробки зображення мовою Matlab (результат обробки представлений на рис. 10.25):

```
A=imread('cameraman.tif'); B=A/2; imshow(B);
```



a)



б)

Рисунок 10.25 – Вихідне зображення (*a*) і результат обробки системою Matlab (*б*)

Контрольні запитання

1. Які особливості зображень є на космічних знімках у видимому діапазоні електромагнітних хвиль ?
2. Які особливості зображень є на космічних знімках у інфрачервоному діапазоні електромагнітних хвиль ?
3. Які особливості зображень є на космічних знімках у мікрохвильовому діапазоні електромагнітних хвиль ?
4. В чому полягає службова обробка космічних знімків ?
5. В чому полягає просторова прив'язка космічних знімків ?
6. В чому полягає радіаційна корекція космічних знімків ?
7. В чому полягає тематична інтерпретація космічних знімків ?
8. Який зміст попередньої обробки космічних знімків ?
9. Який зміст тематичної обробки космічних знімків ?
10. В чому полягає геометрична корекція космічного знімка ?
11. В чому полягає радіометричне калібрування космічних знімків ?
12. В чому полягає радіометрична корекція впливу атмосфери на космічний знімок ?
13. В чому полягає відновлення пропущених пікселів на космічних знімках ?
14. В чому полягає покращення зображень на космічних знімків ?
15. В чому полягає фільтрація космічних знімків ?
16. В чому полягає суть колірних перетворень зображень на космічних знімках ?
17. Що таке індексні зображення і як вони формуються з космічних знімків ?
18. Як здійснюється аналіз космічного знімка методом головних компонентів ?
19. Як здійснюється аналіз космічного знімка методом спектрального поділу ?

ВИКОРИСТАНІ ДЖЕРЕЛА

1. Яне Б. Цифровая обработка изображений / пер. с англ. А.М.Измайловой. Москва : Техносфера, 2007. 584 с.
2. Шапиро Л., Стокман Дж. Компьютерное зрение / пер. с англ. А.А.Богуславского; под ред. С.М.Соколова. Москва : БИНОМ. Лаборатория знаний, 2006. 752 с.
3. Гонсалес Р., Вудс Р. Цифровая обработка изображений / пер. с англ. ЗАО «РИЦ Техносфера»; под ред. П.А.Чочиа. Москва : Техносфера, 2005. 1072 с.
4. Методы компьютерной обработки изображений / Гашников М.В., Глумов Н.И., Ильясова Н.Ю. и др.; под ред. В.А. Сойфера. Москва : ФИЗМАТЛИТ, 2003. 784 с.

Навчальне електронне видання

ПЕРЕЛИГІН Борис Вікторович

ЦИФРОВА ОБРОБКА ЗОБРАЖЕНЬ

Конспект лекцій

Видавець і виготовлювач

Одеський державний екологічний університет

вул. Львівська, 15, м. Одеса, 65016

тел./факс: (0482) 32-67-35

Е-mail: info@odeku.edu.ua

Свідоцтво суб'єкта видавничої справи

ДК № 5242 від 08.11.2016