

ЛЕКЦІЯ № 11

ЗАГАЛЬНІ ВІДОМОСТІ ПРО КОМП'ЮТЕРНИЙ ЗІР

ПЛАН

1. Сучасний погляд на комп'ютерний зір.
2. Типові задачі комп'ютерного зору.
3. Системи комп'ютерного зору.
4. Цифрове подання зображень.

ЛІТЕРАТУРА

Шапиро Л. Компьютерное зрение. Л. Шапиро, Дж. Стокман/ Пер. с англ. – М.:БИНОМ. Лаборатория знаний, 2006. – 752 с.

Форсайт Дэвид А., Понс Жан. Компьютерное зрение. Современный подход.: Пер. с англ. – М.: Издательский дом «Вильямс», 2004. – 928 с.

ВСТУП

Основна задача комп'ютерного зору зрозуміти, що зображено на зображенні



Бачить людина

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

Бачить комп'ютер

Комп'ютер повинен відповісти на будь яке питання про зображення, на яке може відповісти людина.

Що це означає?

ЗІР – це джерело семантичної (сислової) інформації про світ.

ЗІР – джерело інформації про відстані і розміри об'єктів.

Виділення об'єктів



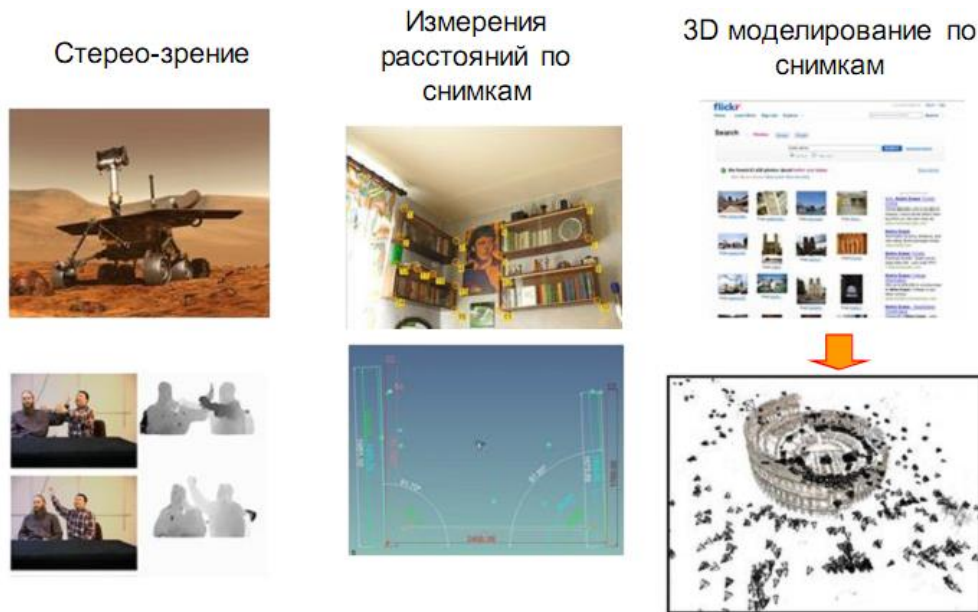
Характеристика об'єктів



Складна класифікація зображень



Виміри на зображеннях



1. Сучасний погляд на комп'ютерний зір

Комп'ютерний зір або **Комп'ютерне бачення** — теорія та технологія створення машин, які можуть проводити виявлення, стеження та класифікацію об'єктів.

Як наукова дисципліна, комп'ютерний зір належить до теорії та технології створення штучних систем, які отримують інформацію у вигляді зображень. Відеодані можуть бути представлені у вигляді багатьох форм, таких як відеопослідовність, зображення з різних камер або тривимірними даними з медичного сканера.

Як технологічна дисципліна, комп'ютерний зір прагне застосувати теорії та моделі комп'ютерного зору до створення систем комп'ютерного зору. Прикладами таких систем можуть бути:

- системи керування процесами (промислові роботи, автономні транспортні засоби);

- системи відео спостереження;

- системи організації інформації (наприклад, для індексації баз даних зображень);

- системи моделювання об'єктів або оточуючого середовища (аналіз медичних зображень, топографічне моделювання);

- системи взаємодії (наприклад, пристрої введення для систем людино-машинної взаємодії).

Комп'ютерний зір також може бути описаний як доповнення (але не обов'язково протилежність) біологічному зору. У біології вивчається зорове сприйняття людини і різноманітних тварин, в результаті створюються моделі роботи таких систем в термінах фізіологічних процесів. Комп'ютерний зір, з іншого боку, вивчає і описує системи комп'ютерного зору, які виконано апаратно або програмно. Міждисциплінарний обмін між біологічним та комп'ютерним зором виявився досить продуктивним для обох наукових галузей.

Підрозділи комп'ютерного зору включають відтворення дій, виявлення подій, стеження, розпізнавання образів, відновлення зображень.

Сучасний стан справ

Область комп'ютерного зору може бути охарактеризована як молода та різноманітна. І, хоча існують більш ранні роботи, можна сказати, що тільки з кінця 1970-х почалось інтенсивне вивчення цієї проблеми, коли комп'ютери змогли керувати обробкою великих наборів даних, таких як зображення. Однак, ці дослідження зазвичай починались з інших галузей, і, відповідно, **немає стандартного формулювання проблеми комп'ютерного зору**. Також, і це навіть більш важливо, немає стандартного формулювання того, як повинна вирішуватись проблема комп'ютерного зору. Замість того, існує маса методів для вирішення різноманітних строго визначених задач комп'ютерного зору, де методи часто залежать від задач і рідко коли можуть бути узагальнені для широкого кола застосування.

Важливою частиною в області штучного інтелекту, є автоматичне планування або прийняття рішень в системах, які можуть виконувати механічні дії, такі як пересування робота через деяке середовище. Цей тип обробки зазвичай потребує вхідних даних, що надаються системами комп'ютерного зору, діють як відеосенсор і надають високорівневу інформацію про середовище та роботу.

Інші галузі, які іноді описуються як ті, що належать до штучного інтелекту, і які використовуються відносно комп'ютерного зору, це розпізнавання образів та навчальні методи. *В результаті комп'ютерний зір іноді розглядається як частина галузі штучного інтелекту чи галузі комп'ютерних наук загалом.*

Фізика є іншою наукою, яка тісно пов'язана з комп'ютерним зором. Значна частина комп'ютерного зору має справу з методами, які потребують досконального розуміння процесу, в якому електромагнітне випромінювання, зазвичай в області видимого або інфрачервоного спектру, відображається поверхнею об'єктів та вимірюється датчиком зображення, щоб отримати відеодані.. Тому, комп'ютерний зір може розглядатись і як розширення фізики.

Третя галузь науки, яка відіграє важливу роль, — це *нейробіологія*, зокрема вивчення систем біологічного зору. За останнє століття були проведені великі дослідження очей, нейронів та структур мозку, що відносяться до обробки візуальних подразників як у людини, так і в різноманітних тварин. Це призвело до грубого, і водночас складного, опису того, як працюють «реальні» системи зору, що допомогло вирішити деякі задачі. Результати цих досліджень призвели до створення штучних систем, що імітують роботу і функціонування аналогічних біологічних систем на різних рівнях складності. Також, деякі методи вивчення, що були розроблені в галузі комп'ютерного зору, зобов'язані своєму походженню біології.

Ще однією галуззю, що пов'язана з комп'ютерним зором, є *обробка сигналів*. Багато методів обробки одномірних сигналів, зазвичай часових сигналів, можуть бути природнім шляхом розширені для обробки двовимірних чи багатовимірних сигналів в комп'ютерному зорі. Однак, через своєрідну природу зображень, існують багато методів, розроблених в галузі комп'ютерного зору, що не мають аналогів в області обробки одновимірних сигналів.

Пов'язані області

Комп'ютерний зір, Обробка зображень, Аналіз зображень, Зір робота і Машинний зір — тісно пов'язані області. Але досі точно не визначено, чи є вони розділами однієї, ширшої галузі. При детальному аналізі може здатись, що це лише різні назви однієї і тієї ж області. Щоб не виникло плутанини, прийнято розрізняти їх як напрямки, сфокусовані на певному предметі вивчення. Нижче наведено опис деяких з них, найбільш важливих.

Обробка зображень і Аналіз зображень в основному зосереджені на роботі з двовимірними зображеннями, тобто як перетворити одне зображення в інше. Наприклад, попіксельні операції збільшення контрастності, операції з виділення країв, усунення шумів чи геометричні перетворення, такі як обертання зображення. Дані операції припускають, що обробка/аналіз зображення діють незалежно від вмісту самих зображень.

Комп'ютерний зір зосереджується на обробці тривимірних сцен, спроектованих на одне чи декілька зображень. Наприклад, відновлення структури чи іншої інформації про тривимірну сцену по одному чи декільком зображенням. Комп'ютерний зір часто залежить від більш чи менше складних припущень відносно того, що представлено на зображеннях.

Машинний зір зосереджується на застосуванні, в основному промисловому, наприклад, автономні роботи і системи візуальної перевірки та вимірювання. Це означає, що технології датчиків зображення і теорії управління пов'язані з обробкою відеоданих для управління роботом і обробка даних в реальному часі здійснюється апаратно чи програмно.

Також існує область, названа *Візуалізацією*, яка початково була пов'язана з процесом створення зображень, але іноді мала справу з обробкою та аналізом. Наприклад, рентгенографія працює з аналізом відеоданих медичного призначення.

Нарешті, розпізнавання образів є областю, яка використовує різноманітні методи для отримання інформації з відеоданих, що, в основному, базуються на статистичному підході. Значна частина цієї області присвячена практичному застосуванню цих методів.

Приклади застосування комп'ютерного зору

Одним з найбільш важливих застосувань є *обробка зображень в медицині*. Ця область характеризується отриманням інформації з відеоданих для визначення медичного діагнозу пацієнту. В більшості випадків, відеодані отримують за допомогою мікроскопії, рентгенографії, ангіографії, ультразвукових досліджень та томографії.

Іншою прикладною областю комп'ютерного зору є *промисловість*. Тут інформацію отримують для підтримки виробничого процесу. Прикладом може слугувати контроль якості, коли деталі чи кінцевий продукт автоматично перевіряються на наявність дефектів.

Військове застосування є, мабуть, найбільшою областю комп'ютерного зору.

Одними з нових областей застосування є *автономні транспортні засоби*, включаючи підводні, наземні (роботи, машини), повітряні. Рівень автономності

вимірюється від повністю автономних (безпілотних) до транспортних засобів, де системи, що базуються на комп'ютерному баченні, підтримують водія чи пілота в різноманітних ситуаціях. Повністю автономні транспортні засоби використовують комп'ютерне бачення для навігації, тобто для отримання інформації про місце свого положення, для створення карти навколишнього оточення, для визначення перешкод.

Інші області застосування включають:
підтримку створення відеоефектів для кіно та телебачення;
спостереження.

2. Типові задачі комп'ютерного зору

Кожна з областей застосування комп'ютерного зору, що були описані вище, пов'язана з низкою задач; більш чи менше гарно визначені проблеми вимірювання чи обробки можуть бути вирішені з використанням багатьох методів. Деякі приклади типових задач комп'ютерного зору представлені нижче.

Розпізнавання

Класична задача в комп'ютерному зорі, обробці зображень і машинному зорі — це визначення того, чи містять відеодані деякий характерний об'єкт, особливість чи активність. Ця задача може бути достовірно і легко вирішена людиною, але досі не вирішена задовільно в комп'ютерному зорі в загальному випадку: випадкові об'єкти у випадкових ситуаціях.

Існуючі методи вирішення цієї задачі ефективні тільки для окремих об'єктів, таких як прості геометричні об'єкти (наприклад, багатокутники), людські обличчя, друковані чи рукописні символи, автомобілі і лише у визначених умовах, зазвичай це певне освітлення, фон і положення об'єкта відносно камери.

В літературі описане різноманіття проблем розпізнавання:

Виявлення: відеодані перевіряються на наявність визначеної умови. Наприклад, виявлення можливих неправильних клітин чи тканин в медичних зображеннях. Виявлення, що базується на відносно простих і швидких обчисленнях, іноді використовується для знаходження невеликих ділянок в зображенні, що аналізується, які потім аналізуються за допомогою заходів, що потребують більше ресурсів, для отримання правильної інтерпретації.

Розпізнавання: один чи декілька попередньо заданих чи вивчених об'єктів або класів об'єктів можуть бути розпізнані, зазвичай разом з їх двовимірним положенням на зображенні чи тривимірним положенням в сцені.

Ідентифікація: розпізнається індивідуальний екземпляр об'єкта. Приклади: ідентифікація визначеного людського обличчя чи відбитка пальців чи автомобіля.

Існує кілька спеціалізованих задач, що базуються на розпізнаванні, наприклад:

Пошук зображень за вмістом: знаходження всіх зображень серед великого набору зображень, які мають певний вміст. Вміст може бути визначено різними шляхами, наприклад в термінах схожості з конкретним зображенням (знайти всі зображення, що схожі на дане зображення), чи в термінах високорівневих критеріїв пошуку, що вводяться як текстові дані (знайти всі зображення на яких зображено багато будинків, які зроблені взимку і на яких нема машин).

Оцінка положення: визначення положення чи орієнтації визначеного об'єкта відносно камери. Прикладом застосування цієї техніки може бути сприяння руки робота при вилученні об'єктів з конвеєра на лінії складання.

Оптичне розпізнавання символів: розпізнавання символів на зображеннях друкованого чи рукописного тексту, зазвичай для перекладу в текстовий формат, найбільш зручний для редагування чи індексації (наприклад, ASCII).

Рух

Декілька задач, що пов'язані з оцінкою руху, в яких послідовність зображень (відеодані) обробляється для знаходження швидкості кожної точки зображення чи 3D сцени, або навіть самої камери, що робить зйомку. Прикладами таких задач є:

одометрія — визначення руху камери (переміщення і обертання) в тривимірному просторі на основі серії знімків.

стеження, тобто прямуванням за переміщенням об'єкта (наприклад, машин чи людей).

оптичний потік — визначення руху кожної точки зображення відносно площини зображення, тобто видимий рух, що є результатом руху як самої точки так і камери.

Відновлення сцени

Задані два або більше зображення сцени або відеодані. Відновлення сцени має за мету відтворення тривимірної моделі сцени. В найпростішому випадку, моделлю може бути набір точок тривимірного простору. Складніші методи відтворюють повну тривимірну модель.

Відновлення зображень

Задача відновлення зображень — це видалення шумів (шум датчика, розмитість об'єкта, що рухається тощо). Найпростішим підходом до вирішення цієї задачі є різноманітні типи фільтрів, таких як фільтри низьких чи середніх частот. Складніші методи використовують уявлення того, як повинні виглядати ті або інші ділянки зображення, і на основі цього їх перетворення.

Більш високий рівень видалення шумів досягається протягом первинного аналізу відеоданих на наявність різноманітних структур, таких як лінії чи межі, а потім управління процесом фільтрації на основі цих даних.

3. Системи комп'ютерного зору

Реалізація систем комп'ютерного зору сильно залежить від області їх застосування. Деякі системи є автономними і вирішують специфічні проблеми детектування та вимірювання, тоді як інші системи складають підсистеми більших систем, які, наприклад, можуть містити підсистеми контролю за механічними маніпуляторами, планування, інформаційні бази даних, інтерфейси людина-машина тощо. Реалізація систем комп'ютерного зору також залежить від того, чи є її функціональність наперед визначеною чи деякі її частини можуть бути вивчені і модифіковані в процесі роботи. Однак, існують функції, типові для багатьох систем комп'ютерного зору.

Отримання зображень: цифрові зображення отримуються від одного чи декількох датчиків зображення, які окрім різноманітних типів світлочутливих камер включають датчики відстані, радари, ультразвукові камери тощо. Залежно від типу датчика, отримані дані можуть бути звичайним 2D зображенням, 3D зображенням чи послідовністю зображень. Значення пікселів зазвичай відповідають інтенсивності світла в одній чи декількох спектральних смугах (кольорові чи зображення у відтінках сірого), але можуть бути пов'язані з різноманітними фізичними вимірюваннями, такими як глибина, поглинання чи відображення звукових або електромагнітних хвиль, або ядерним магнітним резонансом.

Попередня обробка: перед тим, як методи комп'ютерного зору можуть бути застосовані до відеоданих з метою вилучення певної частини інформації, необхідно обробити відеодані, щоб вони задовольняли деяким вимогам залежно від метода, що використовується. Приклади:

- повторна вибірка з метою, щоб переконатись, що координатна система зображення є правильною;

- видалення шумів з метою, щоб видалити спотворення, що вносяться датчиком;

- покращення контрастності для того, щоб потрібна інформація могла бути виявлена;

- масштабування для кращого розрізнення структур на зображенні.

Виокремлення деталей: деталі зображення різного рівня складності виділяються з відеоданих. Типовими прикладами таких деталей є:

- лінії та межі;

- локалізовані точки інтересу, такі як кути, краплі чи точки: складніші деталі можуть належати до структури, форми чи руху.

Детектування/Сегментація: на певному етапі обробки приймається рішення про те, які точки чи ділянки зображення є важливими для подальшої обробки. Прикладами є: виділення визначеного набору точок, що нас цікавлять;

Сегментація однієї або кількох ділянок зображення, які містять характерний об'єкт.

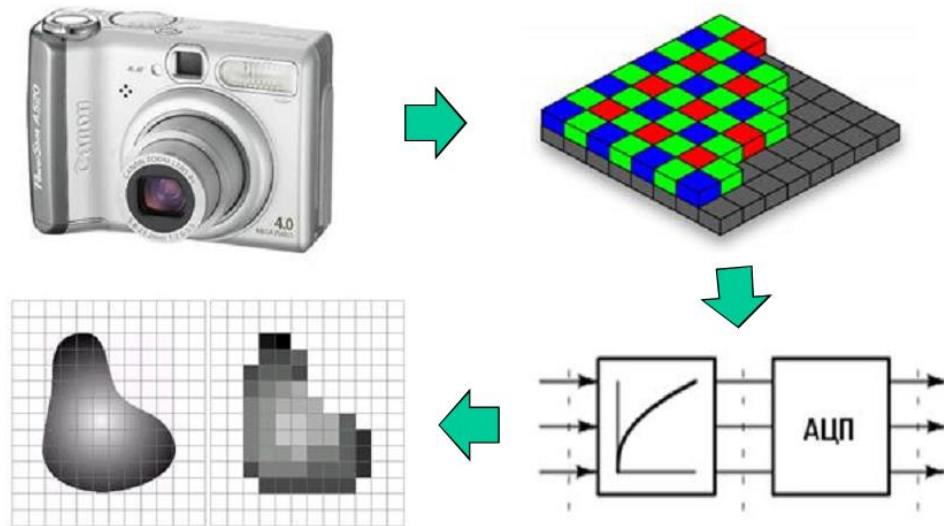
Високорівнева обробка: на цьому кроці вхідні дані зазвичай представляють невеликий набір даних, наприклад, набір точок чи ділянка зображення, в якій за припущенням знаходиться певний об'єкт. Прикладами є:

- перевірка того, що дані задовольняють умовам, що залежать від методу і застосування;

- оцінка характерних параметрів, таких як положення або розмір об'єкта;

- класифікація знайденого об'єкта за різними категоріями.

4. Цифрове подання зображень



Дискретизація растріванням та теорема відліків

Розподіл енергії відбитого від об'єктів випромінювання за просторовими координатами x , y і довжинами хвиль λ може бути описаний математично функцією $E(x,y,\lambda)$. Енергія випромінювання пропорційна квадрату амплітуди електричного поля \dot{y} , тому, являє собою дійсну позитивну величину. У системах, що створюють зображення, максимальна енергія зображення завжди обмежена.

Розміри зображення обмежені формуючою системою і середовищем, на яке воно записується. З метою спрощення будемо вважати, що всі значення енергії точок зображення відмінні від нуля тільки в визначеній прямокутній області $[X, Y]$. Таким чином, величина E є обмеженою функцією трьох обмежених перемінних: $E(x,y,\lambda)$. Будемо вважати цю функцію безпервною в області її визначення.

Відчуття *світлоти*, що виникає в зоровій системі людини, визначається *миттєвою яскравістю* колірного поля, тобто величиною

$$I(x, y) = \int E(x, y, \lambda) S(\lambda) d\lambda ,$$

де $S(\lambda)$ - спектральна чутливість людського зору. На рис.1.1,а. наведений приблизний графік функції $S(\lambda)$.

Слід зауважити що, саме відчуття світлоти нелінійно пов'язане з миттєвою яскравістю. Функціональна залежність зміни відчуття світлоти людиною від зміни яскравості наведена на рис.1.1,б.

Для *панхроматичної* системи спостереження функція $I(x,y)$ являє собою розподіл яскравості чи будь якої іншої фізичної величини, зв'язаної з яскравістю (оптичною щільністю, відбивною здатністю та такими іншими). Вона може бути описана виразом

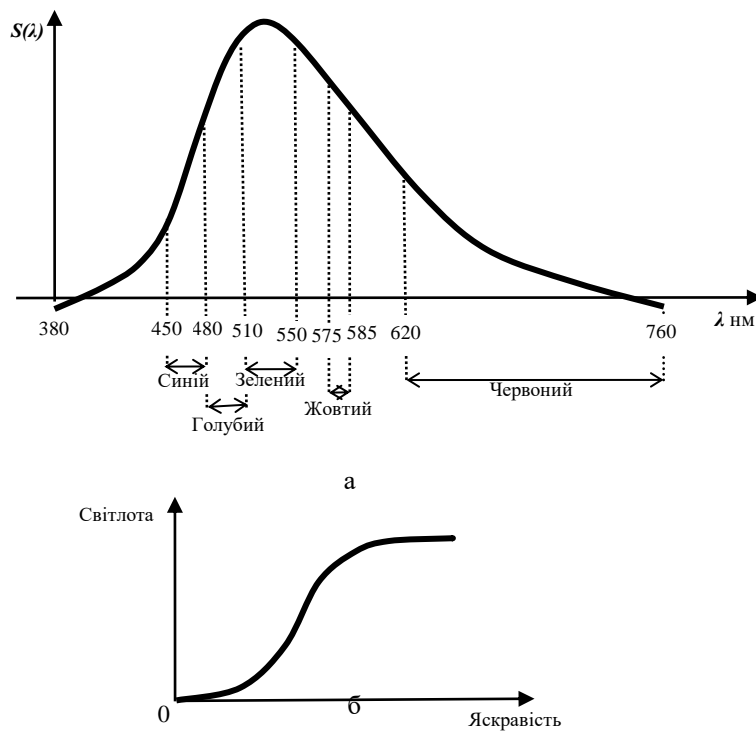


Рис.1.1. Спектральна чутливість людського зору (а) та залежність зміни відчуття світлоти людиною від зміни яскравості (б)

$$I(x, y) = \int E(x, y, \lambda) S(\lambda) d\lambda$$

де $S(\lambda)$ - спектральна чутливість датчика.

У багатозональних системах спостереження зображення для n -ної спектральної зони описується виразом

$$I_n(x, y) = \int E(x, y, \lambda) S_n(\lambda) d\lambda,$$

де $S_n(\lambda)$ - спектральна чутливість датчика n -го діапазону.

Колірні відчуття можна описати набором так званих координат кольору $\{I_R(x, y), I_G(x, y), I_B(x, y)\}$, які є пропорційними інтенсивностям червоного, зеленого і синього кольорів, суміш яких дає заданий колір. Можна використовувати інші системи колірних координат.

Для обробки безперервних зображень у цифрових процесорах необхідно, насамперед, перетворити їх у цифрову форму. Це перетворення являє собою перший етап обробки, що виконується за допомогою спеціальних пристроїв - перетворювачів аналог-код.

Перетворення безперервних сигналів у цифрові здійснюють найчастіше у вигляді послідовності процедур - дискретизації по просторі і квантування за значеннями. Дискретизація відноситься до класу лінійних перетворень сигналу, поелементне квантування - до класу поелементних нелінійних перетворень.

Дискретизація сигналу та відновлення зображень

Дискретизація сигналу зображення - це заміна безперервного в просторі (x,y) сигналу яскравості $I(x,y)$ послідовністю дискретних значень $F(i,j)$.

Найбільш зручним, з погляду організації обробки і природним способом дискретизації є подання сигналів у вигляді вибірок їхніх значень (відліків) в окремих, регулярно розташованих точках $(x=i\Delta x, y=j\Delta y)$. Щоб виділити цей спосіб дискретизації серед безліч інших можливих, будемо називати його *раструванням*. Послідовність точок $(x=i\Delta x, y=j\Delta y)$, у яких беруться відліки, називається *растром*.

Практично операція растрування здійснюється шляхом виміру значень сигналу за допомогою датчика, дію якого можна описати як згортку сигналу зображення з деякою двовимірною функцією $A_D(u,v)$:

$$F(i, j) \approx \iint I(i\Delta x - u, j\Delta y - v)A_D(u, v)dudv \quad (1.1)$$

Функція $A_D(u,v)$ являє собою характеристику розподілу чутливості елементарного приймача за полем зору. Її перетин площиною $A_D(u,v)=0$ задає *миттєве поле зору* елементарного дискретного фотоприймача.

Набір одержуваних значень $F(i,j)$ складає дискретне подання сигналу. Сукупність цих значень за звичай подається у вигляді матриці $F(i,j)$ відліків значень яскравості зображення в точках вузлів растра (рис. 1.2).

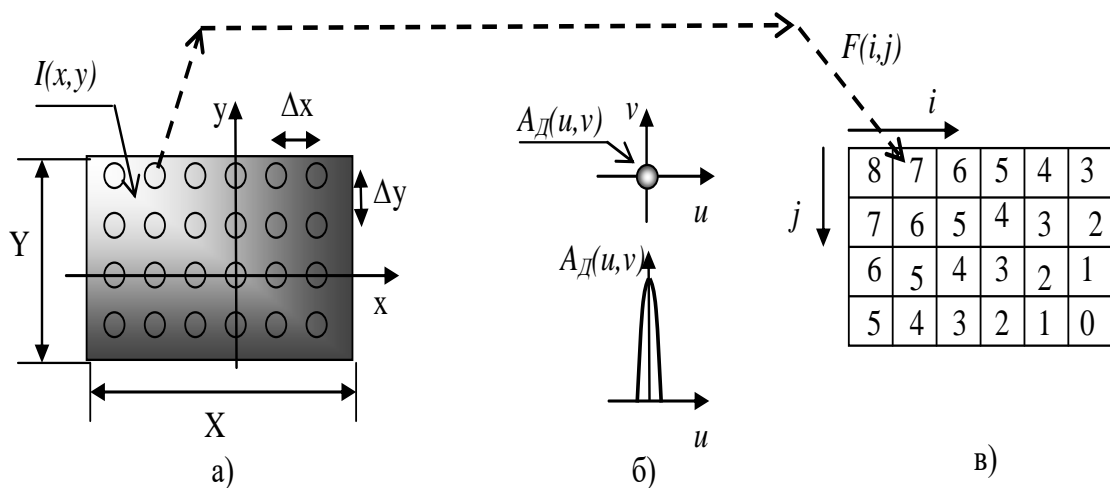


Рис.1.2. Вихідне зображення $I(x,y)$ з растром - (а), апертура дискретизації зображення $A_D(u,v)$ - (б) і результат дискретизації $F(i,j)$ - (в)

Функція $A_D(u,v)$ називається *апертурою дискретизації* [Л8].

Відновлення безперервного сигналу з отриманої послідовності його наближених значень $F(i,j)$ виконується шляхом *інтерполяції* його за цими значеннями:

$$I(x, y) \approx \sum_i \sum_j F(i, j)A_B(u + i\Delta x, v + j\Delta y) \quad (1.2)$$

за допомогою *інтерполюючої функції* $A_B(u,v)$. Ця функція називається *апертурою відновлення* [Л8]. Наприклад, в електронно-променевих пристроях запису і телевізійних дисплеях апертура відновлення - це розподіл яскравості світлової плями люмінофора в рухливій системі координат з початком у точці $(i\Delta x, j\Delta y)$.

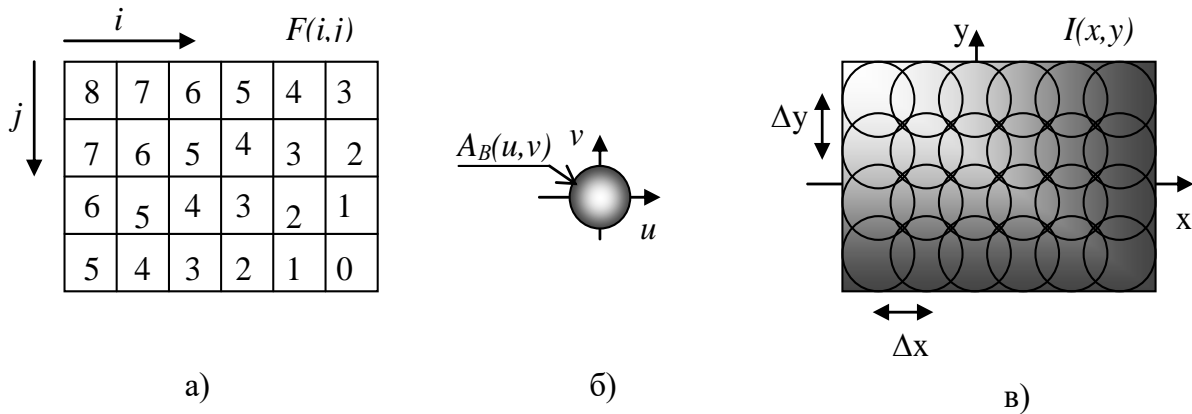


Рис.1.3. Дискретизоване зображення $F(i,j)$ - (а), апертура відновлення $A_B(\mathbf{u},\mathbf{v})$ - (б) і відновлене зображення $I(x,y)$ - (в)

Колами на рис1.3. "в" показано положення апертури відновлення. $\Delta x, \Delta y$ - крок растра відновлення по горизонталі і вертикалі.

Теорема відліків

Якщо виходити тільки з точності відновлення, то існує важливий клас сигналів (зображень) для яких вищенаведені подання є абсолютно точними. Це *сигнали з обмеженим просторовим спектром*. До них можна віднести і реальні зображення. Просторовий спектр Фур'є зображень, одержуваних в оптичних системах, обмежений через обмеженість розмірів лінз об'єктивів і взагалі через обмеженість розмірів найдрібнішої деталі зображення.

Нехай просторовий спектр сигналу відмінний від нуля на інтервалі $[-W, W]$. Для таких сигналів (1.1) і (1.2) переходять у точні рівності:

$$F(i, j) = \iint I(i\Delta x - u, j\Delta y - v) \delta(u, v) du dv \quad (1.3)$$

$$I(x, y) = \sum_i \sum_j F(i, j) \frac{\sin(u + i\Delta x, v + j\Delta y)}{(u + i\Delta x, v + j\Delta y)}. \quad (1.4)$$

Ці співвідношення, що означають можливість точного відновлення сигналів з обмеженим спектром по послідовності відліків, узятих на растрі з кроком $(\Delta x, \Delta y)$, називаються *теоремою відліків* [Л8].

Практично крок дискретизації повинний вибиратися з умови, щоб у найменшу деталь зображення потрапляло не менше двох точок растра по кожній координаті [Л1].

Рівність (1.4) означає, що ідеально дискретизоване зображення може бути відновлене без втрат шляхом інтерполяції на растрі відновлення з кроком $(\Delta x, \Delta y)$ за допомогою відновлюючого променя з розподілом виду $\sin(x)/x$ і шириною рівною $(\Delta x/2, \Delta y/2)$. Практично це означає, що для ідеального пристрою, що відновлює, крок відновлення (крок переміщення відновлюючого променя) повинний обиратися як половина ширини променя відновлення в площині відновлюваного зображення.

Більшесть реальних пристроїв цифрування зображень мають апертуру дискретизації, що має прямокутну форму із згладженими гаусоїдою краями.

Ефекти дискретизації і відновлення

Слід зазначити, що на практиці дискретизацію і відновлення за теоремою відліків здійснити не можливо.

Перше: жоден фотоприймач не має спрямованість по просторі порівнянної з дельта-функцією. Іншими словами, реальний фотоприймач бере не значення сигналу в точці растра, а середнє значення сигналу за деякою околицею цієї точки.

Реальна апертура дискретизації має цілком реальний розмір, що не дозволяє нескінченно зменшувати крок дискретизації без перекриття апертур. Тому, у першому наближенні можна вважати, що *реальна роздільна здатність* пристрою дискретизації дорівнює подвоєному мінімальному кроку дискретизації цього пристрою.

Друге: жодний пристрій, що відновлює, (монітор, друкуючий пристрій, фоторегістратор) не в змозі створити відновлюючу апертуру (відновлюючий промінь) з розподілом енергії у просторі виду $\sin(x)/x$. Реально апертура відновлення (розподіл енергії в промені, що відновлює,) має розподіл у виді унімодалної функції (наприклад гаусоїди). З цієї причини строге виконання на практиці рівності (1.4) не можливо. Тобто, реальні пристрої, що відновлюють, забезпечують трохи гірше відновлення при тій же растровій сітці відновлення і ширині відновлюючого променя. Звичайно це виявляється у вигляді деякої *втрати різкості* відновленого зображення в порівнянні з вихідним (тим, що було до цифрування).

Для зображень періодичної структури, якщо крок растрування Δx більше величини зворотної довжини спектра зображення, то виникає так званий *строб-ефект* [Л8]. Цим пояснюється явище зворотнього обертання, що спостерігається часто, у телебаченні на обертових об'єктах. Для двовимірних сигналів (зображень) строб-ефект виявляється в зменшенні просторової частоти періодичних структур, просторова частота яких близька до частоти дискретизації (рис.1.4).

Практично строб-ефект можна зустріти при скануванні поліграфічних растрових зображень – журнальних та газетних фотографій. Поліграфічна машина або друкуючий пристрій формує зображення з регулярно розташованих растрових елементів (рис.1.5) [Л2].

Поточне зворотнє значення яскравості (значення оптичної щільності) кожного елементу растру – це кількість зафарбованих

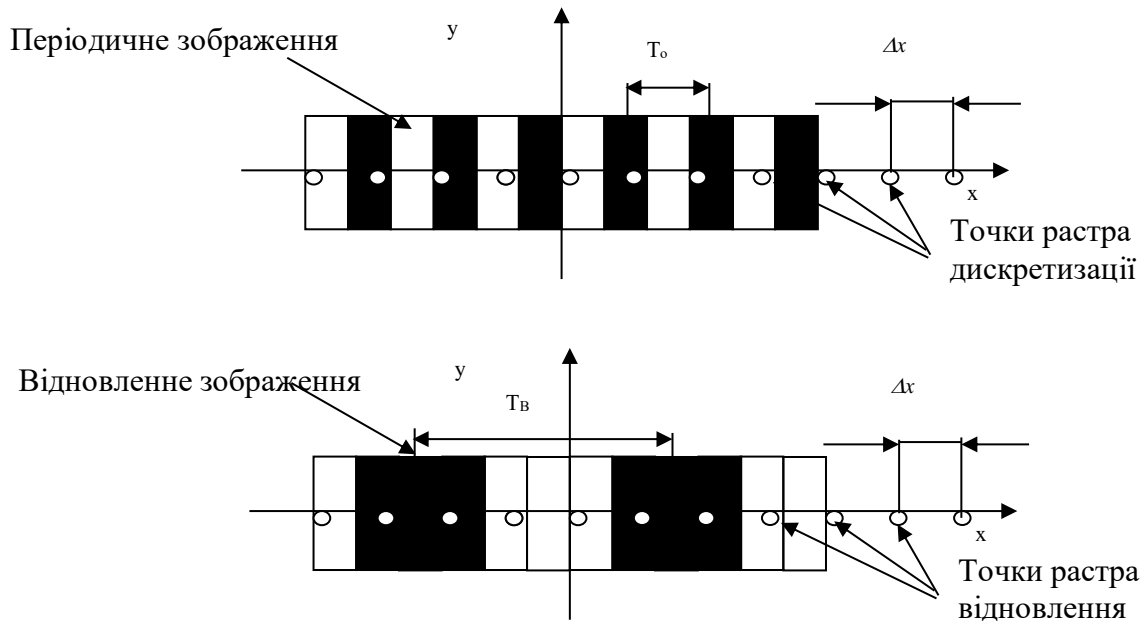


Рис.1.4. Утворення строб-ефекта при дискретизації періодичних зображень з кроком дискретизації трохи більшому, ніж потрібно за теоремою відліків

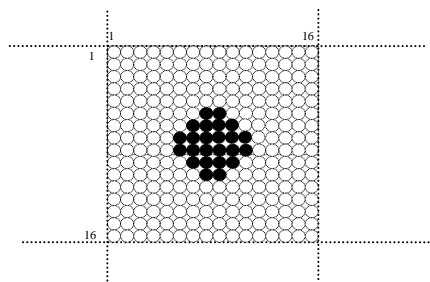


Рис. 1.5. Один елемент растру друкуючого пристрою при друці півтонових зображень з 256 градаціями яскравості

елементів фізичного растру друкуючої голівки. Форма зафарбованої плями для різних друкуючих пристроїв та різних способів розташування елементів у растрі може бути різною. На рисунку 1.6 показано гексогональне розташування елементів растру, яке забезпечує найкращу розрізненість на поліграфічних зображеннях при фіксованому кроці растру та при зафарбованій плямі еліптичної форми [Л2].

Крок з яким розташовані растрові елементи визначається лінеатурою растра. Типові значення кроку поліграфічного растра (у міліметрах) приблизно можна визначити за формулою $25,3/L$, де L може приймати значення: 25 – газетне фото; 75 – для високоякісних видань; 150 – над якісна поліграфія.

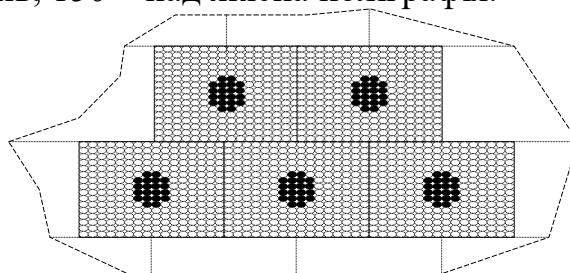


Рис.1.6. Гексагональне розташування елементів растру друкуючого пристрою

На рисунку 1.7. наведеній приклад появи строб-ефекту, який проявляється у вигляді періодичних плям з кроком пропорційним різниці між кроком дискретизації при скануванні (задається програмою управління сканером) та кроком елементів поліграфічного растру.



Рис.1.7. Приклад зображення із строб – ефектом в наслідок невірно обраного кроку дискретизації при скануванні поліграфічного зображення

Практично апаратура відеопрцесора при збільшенні завжди виконує інтерполяцію першого або більш високого порядку (тобто розширює апертуру відновлення до квадрата або кола відповідного розміру), тому зображення на екрані монітора ніколи не стає точковим. Чим більший порядок інтерполяції, тим менше помітний муар-ефект, але й різкість напівтонових меж об'єктів на зображеннях стає гіршою. На рисунку 1.8 показані можливі апертури відновлення (інтерполяційні функції).

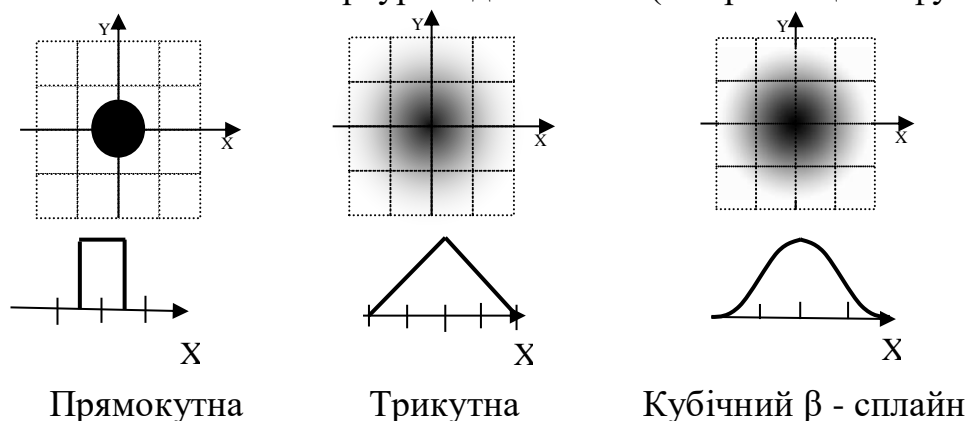


Рис.1.8 Апертури відновлення та відповідні інтерполяційні функції

Приклади відновлення зображень та одновимірних сигналів наведені на рисунку 1.9.

При відновленні зображень з цифрового подання перекручування виникають, якщо ефективна ширина апертури, що відновлює, (ширина променя, що відновлює,)

стає менше кроку растра відновлення. Таке відбувається, наприклад, при збільшенні зображення на екрані монітора. У цьому випадку, у відновленому після растрування зображенні з'являються сітки й інші періодичні компоненти. Це явище називають *муар-ефектом* [Л8]. Практично муар-ефект для різних пристроїв відображення виглядає по-різному. Так, для монітору, якщо не застосовувати інтерполяцію, зображення складається з окремих роздільних точок. Якщо використати при відновленні інтерполяцію першого порядку («найближчий сусід» або квадратний піксел), тоді муар-ефект проявляється у вигляді так званої «пик-

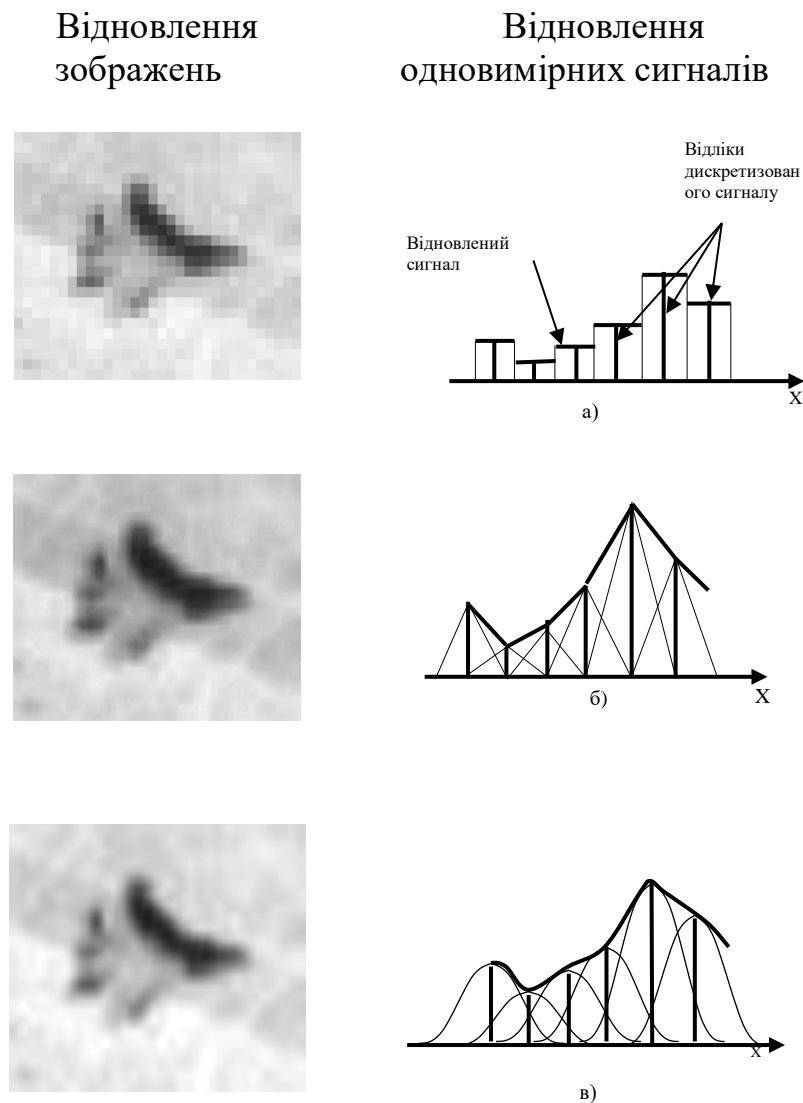


Рис.1.9. Відновлення сигналів різними апертурами відновлення

селізації» зображення – воно складається з помітних прямокутних елементів. Напівтонові межі об'єктів стають у вигляді «сходинок» (aliasing).

Квантування зображень

Поелементне квантування полягає в тому, що в області значень сигналу вибирається відрізок кінцевої довжини, котрий розбивається на інтервали квантування (рис.1.10,а). Значення, що попадають у кожен інтервал, позначаються одним числом - номером інтервалу.

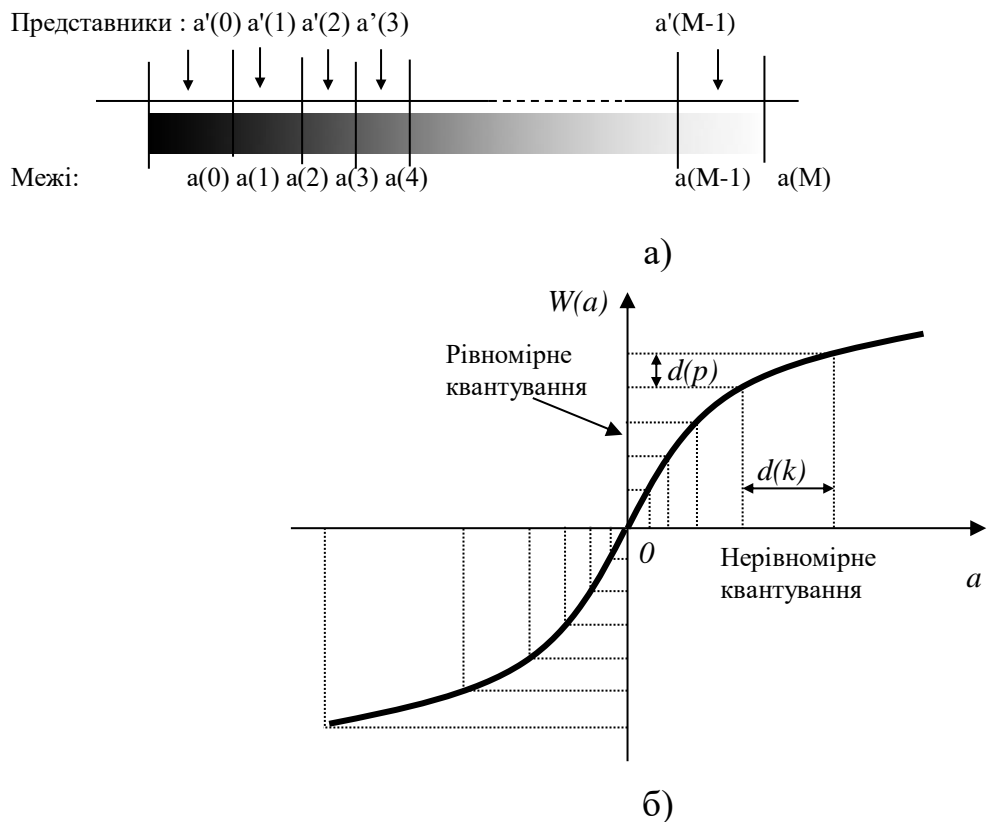


Рис.1.10. Нерівномірне квантування: а) - схема розташування границь і представників інтервалів квантування; б) - нерівномірне квантування за допомогою нелінійного викривлення

При відновленні сигналу номер замінюється значенням, що є представником даного інтервалу. Спосіб розбивки на інтервали і значення-представники інтервалів вибираються так, щоб задовольнялися вимоги до точності представлення безперервного сигналу цифровим.

Слід зазначити, що задача поелементного квантування виникає в цифровій обробці сигналів не тільки при перетворенні безперервних сигналів у цифрові, але і на різних стадіях обчислень у цифрових процесорах при переході від однієї форми представлення чисел до іншої. (наприклад, від формату чисел з плаваючою комою, до байтового формату для збереження їх в архівних запам'ятовуючих пристроях).

Описаний підхід до оптимального квантування, дозволяє одержати чисельне рішення, найкраще погоджені саме з задачею квантування в цифровому процесорі. При виборі ж оптимального квантування безперервного сигналу для введення в цифровий процесор зручніше трохи видозмінити постановку задачі, щоб одержати її аналітичне рішення, яке можна було б втілити в пристрої перетворення аналогового сигналу в цифровий.

Справа в тім, що існуючі пристрої квантування (АЦП) звичайно здійснюють рівномірне квантування сигналів, при якому границі інтервалів квантування розміщуються рівномірно в заданому діапазоні значень сигналу, а представники рівнів квантування розташовуються посередині між цими границями. При відновленні безперервного сигналу з квантованого, також звичайно використовують цифро-аналогові перетворювачі з рівномірно квантованим вхідним сигналом.

Використовуючи такі пристрої, оптимальне (нерівномірне в загальному випадку) квантування можна забезпечити, якщо перед квантуванням сигнал піддати відповідним чином обраному нелінійному перетворенню (попередньому викривленню) (рис.1.11,а). При відновленні безперервного сигналу сигнал на виході пристрою, що відновлює, з рівномірним квантуванням необхідно піддати відповідної нелінійної корекції (рис.1.11,б).

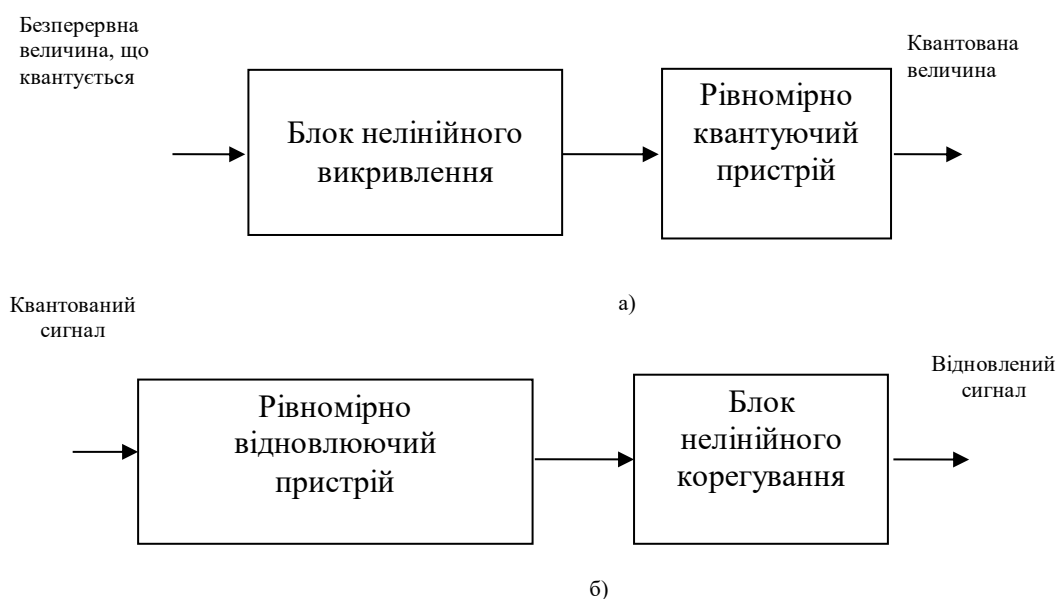


Рис.1.11. Схеми оптимального квантування (а) і відновлення (б) з нелінійним попереднім викривленням і корекцією

При відновленні квантованих зображень можуть спостерігатися наступні викривлення. Якщо рівнів квантування не достатньо, на відтвореному зображенні замість плавних переходів напівтонів будуть спостерігатися скачкоподібні переходи. Межі ланок сусідніх напівтонів будуть помітними (див.рис.1.12). Ці межі називаються *хибними контурами* [Л8]. Одночасно, об'єкти на зображенні, які мали *слабкий контраст* (невелику різницю напівтонів фону і об'єкту) взагалі не будуть спостерігатися на відновленому після квантування зображенні (будуть втрачені).

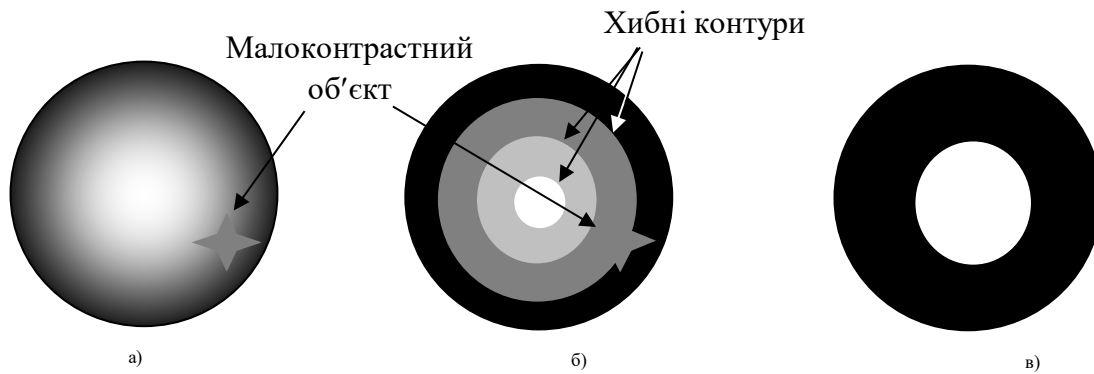


Рис.1.12. Спотворення відновлених після квантування зображень при недостатньому числі рівнів квантування:

- а) безперервне зображення;
- б) відновлене зображення після квантування на 4 рівні;
- в) відновлене зображення після квантування на 2 рівня

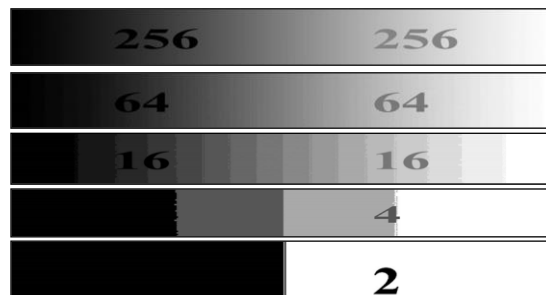


Рис.1.13. Зображення «Шкала сірого» при різному числі рівнів квантування

Рисунок 1.13. демонструє передачу напівтонів при різному числі рівнів квантування. Цифри на рисунку показують застосоване число рівнів квантування. Невеликий контраст цифр приводить до втрати їх зображення при малому числі рівнів.

Колірні координати



Колірна модель має на увазі метод для визначення кольорів.

Найбільш часто використовуються три колірні моделі: RGB (червоний, зелений, синій) і CMYK (блакитний, пурпурний, жовтий, чорний). HSB (тон, насиченість, яскравість).

Модель RGB будується за принципом сполучення червоного (Red), зеленого (Green) і синього (Blue) кольорів різної яскравості та використовується при виводі зображення на екран монітору. Ці компоненти називаються *первинними адитивними* і дозволяють одержати більшість кольорів видимого спектру, а при додаванні дають білий колір (рис.Д1 в колірному додатку).

При обробці кольорових RGB-об'єктів графічна програма привласнює кожному елементу зображення (пікселу) значення інтенсивності, що може змінюватися в межах від 0 (чорний) до 255 (білий). При чому, чим більше значення байта колірний складовий, тим яскравіше цей колір. При накладенні однієї складової на іншу яскравість сумарного кольору також збільшується.

Адитивні кольори використовуються в системах висвітлення, слайд-рекордерах (пристроях запису на фотоплівку), відеосистемах і моніторах. У режимі RGB здійснюється сканування зображень, ця ж модель покладена в основу конструкції екрана монітора.

Модель CMYK використовується для підготовки зображень до чотирьохфарбного друку (із застосуванням блакитного, пурпурного, жовтого і чорного кольорів). Ці фарби називаються *тріадними* (чи *субтрактивними*) і в сумі дають чорний колір. Графічний редактор кожному пікселу CMYK-зображення привласнює значення, що визначають процентний уміст тріадних компонентів. Причому найясніші тони характеризуються їх низьким вмістом, а самі темні (тіні) — відповідно, більш високими значеннями. Чистому білому кольору відповідають нульові значення всіх чотирьох складових.

У той час, як відтворення кольорів за допомогою моделі RGB залежить від джерела світла, в основі моделі CMYK лежить здатність друкувальних фарб до *світлопоглинання* (абсорбції). При проходженні білого світла через світлопроникну фарбу частина спектра поглинається. Не поглинений колір відбивається і попадає назад в око людини.

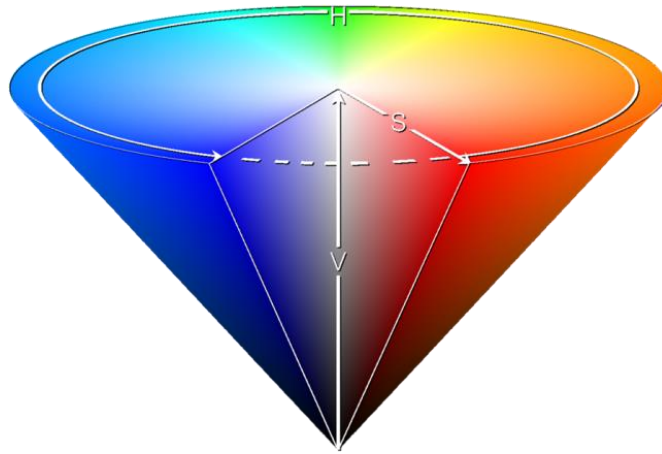
Як правило, перехід до формату CMYK виконується на завершальній стадії обробки зображення.

В основі *моделі HSB* лежить сприйняття кольорів людським оком (рис.Д2). У ній усі кольори визначаються трьома базовими параметрами:

Колірний тон (hue) — це довжина світлової хвилі, відбитої чи минулої через об'єкт. Він займає визначене положення на стандартному колірному колі, для його опису використовується назва кольору, і він характеризується величиною кута в діапазоні 0...360 градусів.

Насиченість (saturation), чи *хроматичність* — це ступінь чистоти кольору. Насиченість визначає співвідношення сірого кольору і поточного колірному тону; вона виражається у відсотках від 0% (сірий) до 100% (цілком насичений). На стандартному колірному колі насиченість збільшується від центра до його границь.

Яскравість (brightness) характеризує відносну світлоту кольору. Звичайно вона вимірюється у відсотках у діапазоні від 0% (чорний) до 100% (білий).



Модель HSB зручно застосовувати при створенні (малюванні) зображення, а по закінченні роботи зображення можна перетворити в модель RGB чи CMYK.

Типовим прикладом колірної моделі такого типу є CIE (Commission Internationale de l'Éclairage — Міжнародна комісія з освітленості). Вона заснована на реакції людського ока на RGB і спроектована таким чином, щоб точно представляти сприйняття кольору людиною. Ця модель використовується для того, щоб визначати так називані *апаратно-незалежні кольори*, що можуть правильно сприйматися та відтворюватися пристроями будь-якого типу: сканерами, моніторами і принтерами.

Моделі такого типу широко застосовуються, тому що вони описують широкий діапазон кольорів.

Поширено наступні моделі: CIE-XYZ, CIE-L*a*b і CIE-LUV.

Колірні охоплення

Колірним охопленням називають усе різноманіття кольорів, що може бути відтворено монітором чи вивідним пристроєм. Видимий спектр електромагнітного випромінювання містить у собі величезна безліч кольорів, що здатне розрізнити людське око. Колірне охоплення будь-якого пристрою є лише підмножиною видимого спектра. Серед усіх колірних моделей найбільш широке колірне охоплення має модель RGB. Колірне охоплення моделі RGB містить у собі підмножину кольорів, що можуть бути відтворені на екрані телевізора чи монітора. Деякі кольори, такі як “чистий блакитний” чи “чистий жовтий”, не можуть бути точно відтворені на екрані.

Самим вузьким колірним охопленням характеризується модель CMYK вона відтворює тільки ті кольори, що можуть бути надруковані за допомогою тріадних фарб (рис.Д3 в додатку).

Стиснення зображень

Зображення у машинному поданні - двовимірний масив $N \times M$, де N - його ширина, M - висота. При скануванні звичайно використовують розділення від 72 до 2400 dpi (dots per inch - точок на дюйм). Найбільше часто - 300 dpi. Якщо взяти лист папера 21/29 см. із зображенням і відсканувати його в RGB Truecolor, то нестиснене зображення буде займати ~ 27300000 байтів чи 26 Мбайт.

Аерокосмічні зображення можуть бути подані в набагато більших обсягах. Так смуга огляду датчика МСУ-8 подається 6000 елементами розділення в трьох частинах діапазону світлових хвиль (два у видимій частини спектра та один у ближньої інфрачервоної). Кодування в кожному піддіапазоні – 11 біт. Приймаючи кадр квадратним, отримаємо, що зображення кадра буде займати $6000 \times 6000 \times 3 \times 11 = 1188$ млн. біт (приблизно 150 Мбайт). Зображення, які можуть бути більш детальними, мають розмір до 1.5 Гбайт.

Методи стиснення застосовуються для будь-яких масивів даних, а не тільки для зображень. Будемо розглядати статичне стиснення, тобто масив даних для стиснення цілком сформований. Методи статичного стиснення часто підрозділяють на послідовне і ентропійне. Послідовний стиснення використовує в роботі наявність повторюваних ділянок даних. Ентропійне використовується з метою скорочення до мінімуму надмірності інформації. Послідовне застосування цих методів дозволяє одержати гарний результат.

Формати файлів зображень

Формати файлів визначають тип інформації, що зберігається у файлі, сумісність цього файлу з різними програмами і можливості обміну даними з іншими користувачами. Вибір найкращого формату для збереження файлу після сканування залежить від декількох факторів — програми, застосовуваної для наступної обробки зображення, вільного місця на диску, необхідності обміну даними з іншими користувачами і способу кінцевого виводу даних.

Багато програм обробки зображень і малювання — наприклад, Adobe Photoshop — дозволяють зберігати зображення в оригінальному форматі, що розпізнається тільки самими цими програмами. Оригінальні формати файлів зручні, тому що файли мають менший розмір і є можливість зберігати складні типи даних (наприклад, у Photoshop шари, маски вибірки, векторні шляхи і канали).

У ENVI оригінальним форматом є формат IMG. Часто зручніше зберігати зображення в оригінальному форматі, особливо якщо виведення зображення буде здійснюватися безпосередньо з цієї програми. Якщо потім виникне необхідність компонувати зображення сторінки, перенести його в інше середовище чи передати замовнику, то завжди можна перезаписати його у форматі, зручному для наступного користувача.

Зберігайте файл у стандартному форматі, що читається більшістю програм:
для виводу на друк зручні TIFF чи EPS,
для обробки — PICT і BMP,
для збереження — TIFF чи PICT.

Якщо файл напевно буде передаватися з однієї програми в іншу і є можливість вибору між різними сумісними форматами, то зберігайте файл у форматі, що містить найбільш повну інформацію про зображення.

Якщо зображення використовується як для друку, так і для обробки, краще створити принаймні дві копії файлу — одну у форматі TIFF чи EPS для друку, а іншу у форматі, найбільш придатному для середовища обробки.

Загальна структура файлів растрових зображень

Растрові файли, відрізняючи друг від друга деталями, мають загальну структуру. Тут будуть розглянуті компоненти типового растрового файлу. Розглянемо загальну структуру растрових файлів.

Растрові файли містять заголовок, растрові дані й іншу інформацію, у тому числі і про колірну палітру.

Основними компонентами простого растрового файлу є:

Заголовок

Растрові дані

Програмну інформацію, що не розміщується в заголовку, розміщують у кінцівці файлу. Якщо застосовується палітра (набір кольорів присутніх у зображенні), то її можна зберегти в заголовку файлу, але зручніше розмістити її в середині файлу, після заголовка. Крім того, палітру можна зберігати після даних зображення на місці кінцівки файлу чи безпосередньо в її складі. Таблиці рядків розгорнення і таблиці колірної корекції можуть розташовуватися після заголовка як перед даними зображення, так і після них:

Заголовок

Палітра

Таблиця рядків розгорнення

Таблиця колірної корекції

Растрові дані

Таблиця колірної корекції

Кінцівка

Організація піксельних даних у виді колірних площин ілюструється рис.1.16.

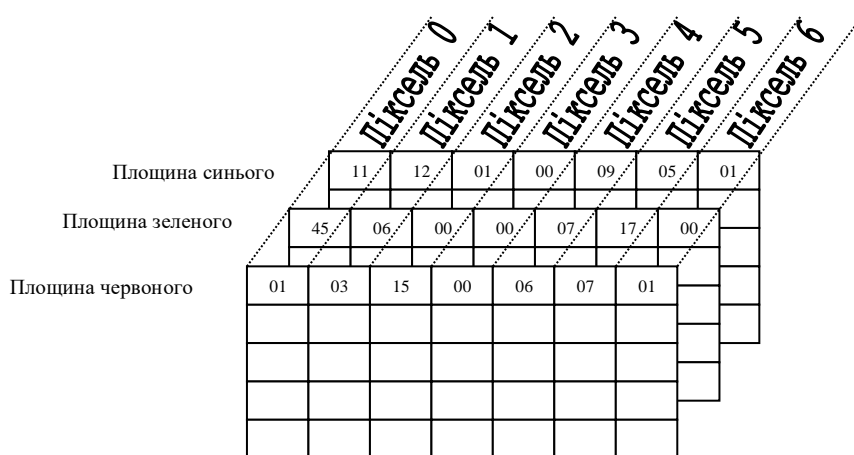


Рис.1.16. Організація піксельних даних у виді колірних площин

Векторні зображення й робота з ними

Векторні зображення являють собою списки об'єктів, що мають координати в загальній системі координат. Об'єкти зображуються графічними примітивами (крапкою, лінією, прямокутником, колом, ...), яки характеризуються стилем зображення до якого, в свою чергу, відносяться різні види заливання, штрихування, товщина ліній, колір, типи символів й ін. Використовуючи загальну систему координат об'єкти відображаються у вигляді графічних примітивів на площині зображення з використанням стилів.

Крім геометричних і стильових характеристик об'єктам присвоюються їхні атрибутивні характеристики, які властиві даним типам об'єкта й унікальні для даного об'єкта. Для запису атрибутивних характеристик об'єкта використовують бази даних. Це дозволяє робити аналіз, селекцію об'єкта за різними критеріями фактично векторне зображення являє собою бази даних, де об'єкти мають ще й свій вид (рис.1.23).

Векторне зображення являє собою списки наступного виду:

Заголовок: тип формату, яка версія й т.д.

Об'єкт 1, тип, тіло об'єкта.

Об'єкт 2, тип, тіло об'єкта.

.....

Об'єкт N, тип, тіло об'єкта.

Тіла можуть винесені окремо й у списку вказується лише місце розташування кожного тіла у файлі.

Ідент. номер	Висота	Адреса	Кількість поверхів	Власник
00001	8	вул. Богунська, 56	1	Приватн.
00002	25	пр. Миру, 23	6	Державн.

Рис. 1.23. База даних об'єктів (у вигляді таблиці)

Якщо необхідно виділити якісь групи об'єктів окремо, то раціонально створювати зображення в окремих таблицях бази даних - шарах.

Векторні графічні примітиви, які відображаються як певні об'єкти бувають трьох основних типів: площинні, лінійні та точкові.

Векторні дані можуть бути застосовані для виділення точкових, лінійних та площинних об'єктів на знімку; для вибору еталонів при класифікації об'єктів; для виділення границь областей, що представляють інтерес (region of interes – ROI).

На рис. 1.27. наведений приклад сумісного використання растрового зображення (космічного знімка) і векторного зображення (позначки анотації) в програмному комплексі обробки космічних знімків ENVI.

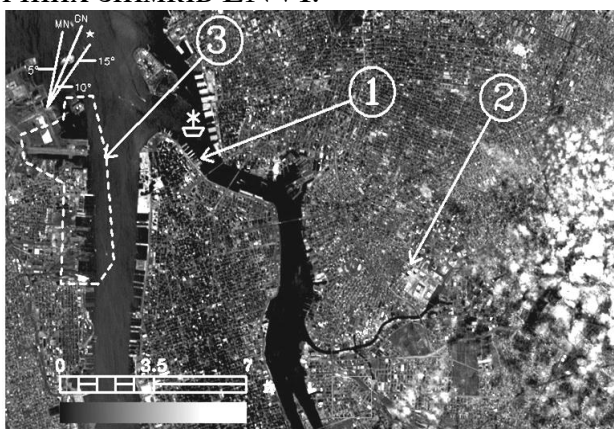


Рис.1.27. Приклад сумісного використання растрового і векторного зображень