

# ЛЕКЦІЯ № 8

## ОСНОВИ КОМП'ЮТЕРНОГО ЗОРУ

### ПЛАН

1. Сучасний погляд на комп'ютерний зір.
2. Цифрове подання зображень.
3. Основні характеристики якості цифрових зображень.
4. Попередня обробка зображень
5. Виділення об'єктів на зображенні
6. Бібліотека комп'ютерного зору OpenCV

### 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

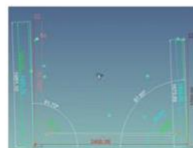


## Виміри на зображеннях

Стерео-зрение



Измерения расстояний по снимкам



3D моделирование по снимкам



## 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).

На сьогоднішній день, **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.

*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.

### **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:

*одоμέтрія* — 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.

## **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.

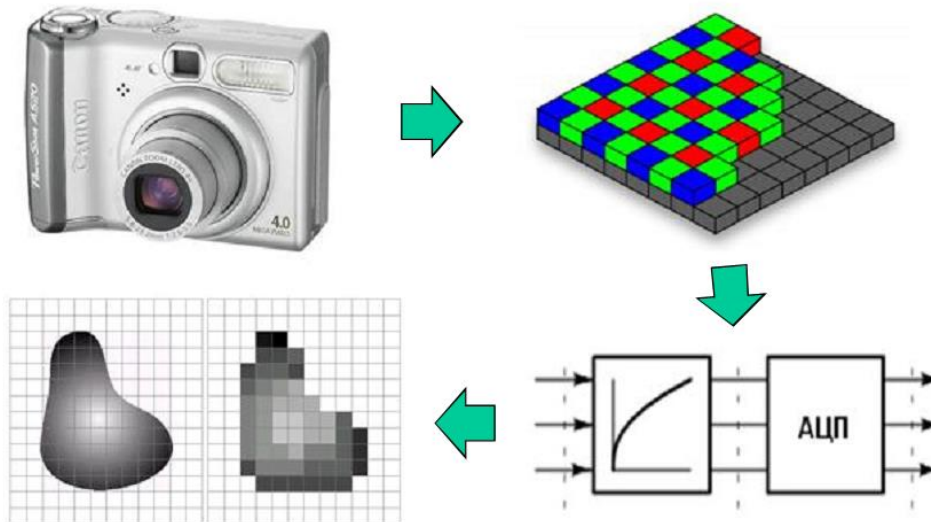
**Високорівнева обробка:** на цьому кроці вхідні дані зазвичай представляють невеликий набір даних, наприклад, набір точок чи ділянка зображення, в якій за припущенням знаходиться певний об'єкт. Прикладами є:

- 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  $\lambda$  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 безперервною 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

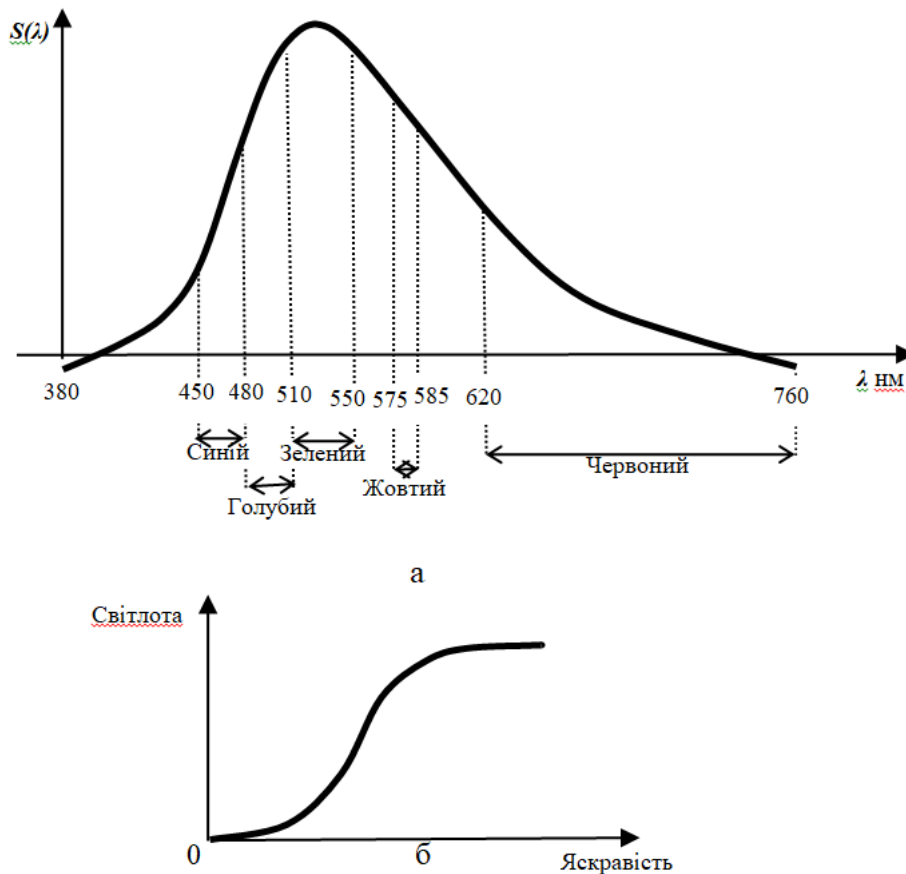


Рис.1.1. Спектральна чутливість людського зору (а) та залежність зміни відчуття світлоти людиною від зміни яскравості (б)

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

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.



## 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 присвоює 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.

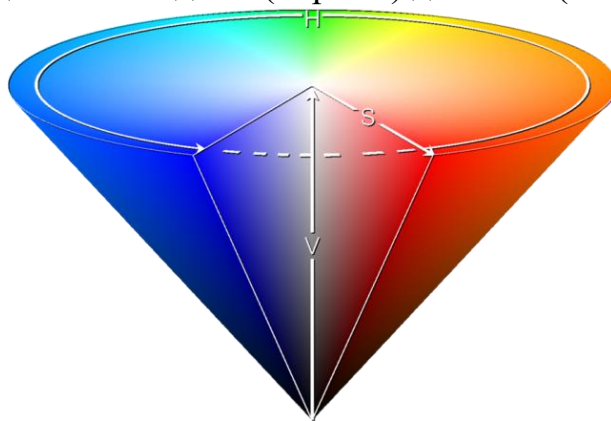
В основі *моделі HSB* лежить сприйняття кольорів людським оком (рис.Д2). У ній усі кольори визначаються трьома базовими параметрами:

*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 колірною 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.

### 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 Truescolor, 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, яки 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, відрізняючись один від одного 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.

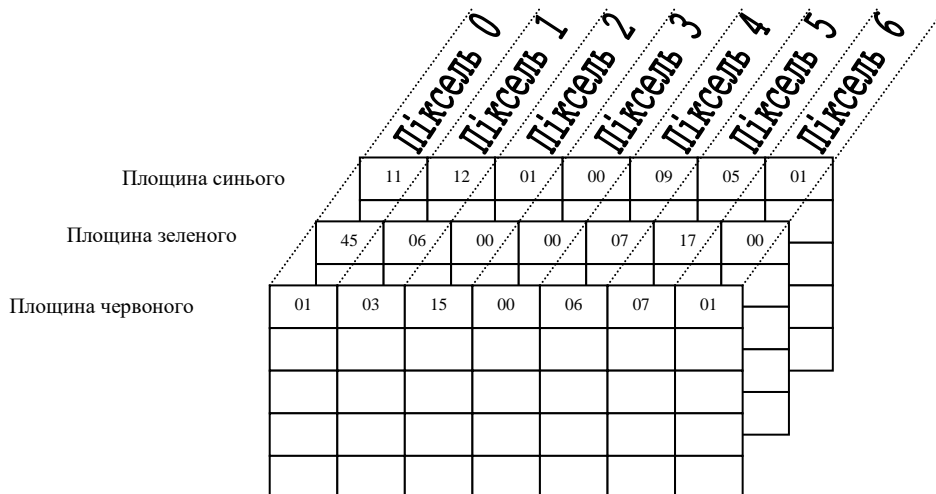


Рис.1.16. Організація піксельних даних у виді колірних площин

### 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, ...), які 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,	1	Приватн.
00002	25	56 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 interest – 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.

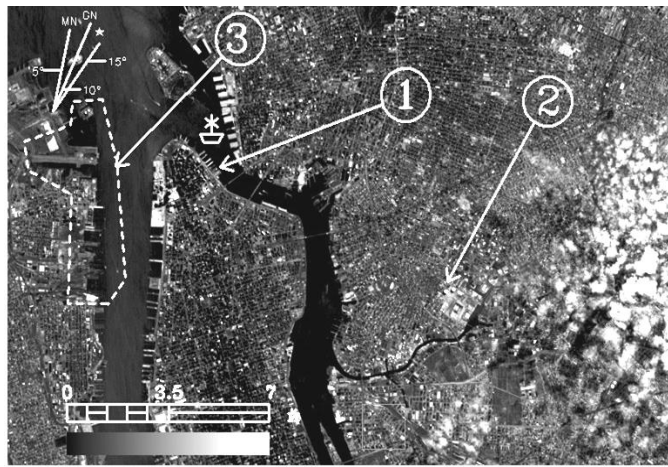


Рис.1.27. Приклад сумісного використання растрового і векторного зображень

### 3. Основні характеристики якості цифрових зображень.

**Геометричні розміри** кількість пікселів у рядку і стовпці.

**Масштаб (мірило) зображення** – відношення розмірів об'єкта на зображенні, виконаних без спотворень, до його реальних значень.

**Детальність зображення** – властивість зображення, що характеризує здатність відображати найменшу за лінійними розмірами деталь (см, мм).



**Яскравість зображення** – це характеристика відбиття світлової енергії від об’єктів і фону, що зафіксована на зображенні.

**Контрастність** - відношенням яскравості найбільш світлої ділянки до яскравості найбільш темної ділянки. Вона є безрозмірною величиною та вимірюється коефіцієнтом  $K = 0 \dots 1$ .

Наявність чи відсутність **викривлень** на зображенні – спотворення на зображеннях.

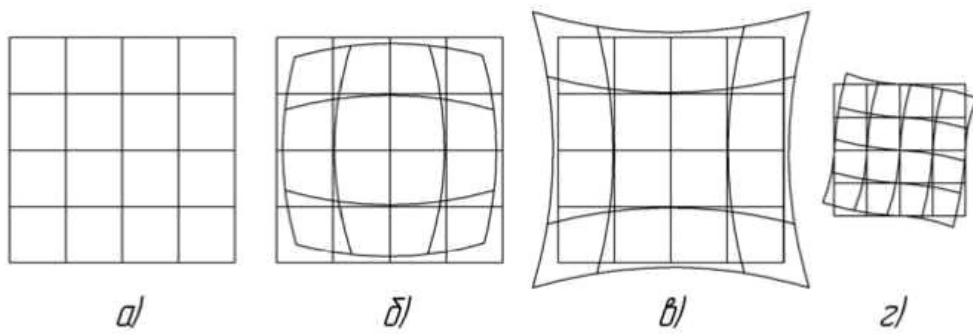


Рисунок 7.7 – Спотворення форми об'єкта дисторсією лінзи:  
а – тест-об'єкт (квадратна сітка); б – бочкоподібна дисторсія;  
в – подушкоподібна дисторсія; г – анізотропна дисторсія

**Кольоровість** – наявність чи відсутність кольору та спосіб його подання



#### 4. Попередня обробка зображень

Що може погіршуватися на зображенні?



Темне (світле) або слабо контрастне.



Шуми



Нерізкість



Нерівномірна  
Освітленість



Неправильні  
кольори

#### 4.1. Радіометрична корекція цифрових зображень

Слабкий контраст - найбільш розповсюджений дефект зображень, обумовлений обмеженістю діапазону відтворених яскравостей, що нерідко сполучається з нелінійністю характеристики передачі рівнів. У багатьох випадках контраст можна підвищити, змінюючи яскравість кожного елемента зображення.



**Гістограма** – функція (графік) розподілу яскравості на зображенні. Кожен стовпець – частота повторення яскравості на зображенні.



0

255



0

255

Не повністю використовується весь діапазон яскравостей.

Нерівномірне заповнення діапазону яскравостей, групування біля певних значень.



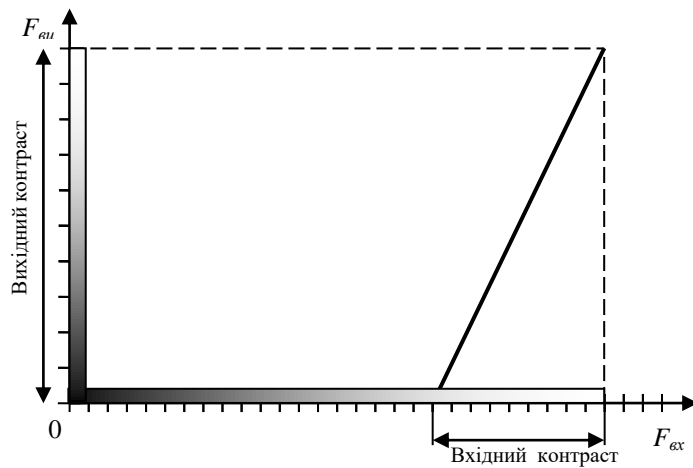
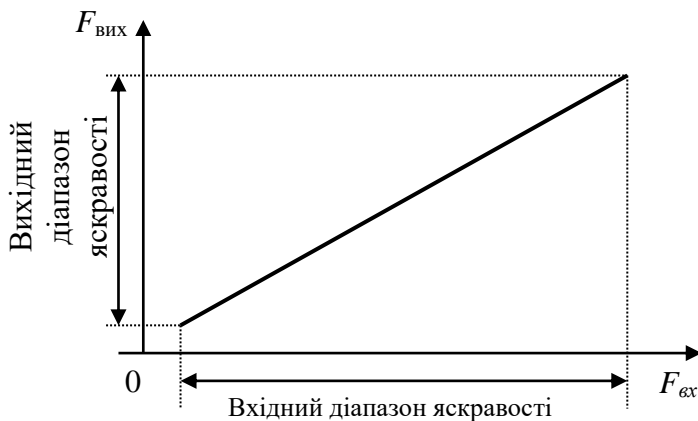


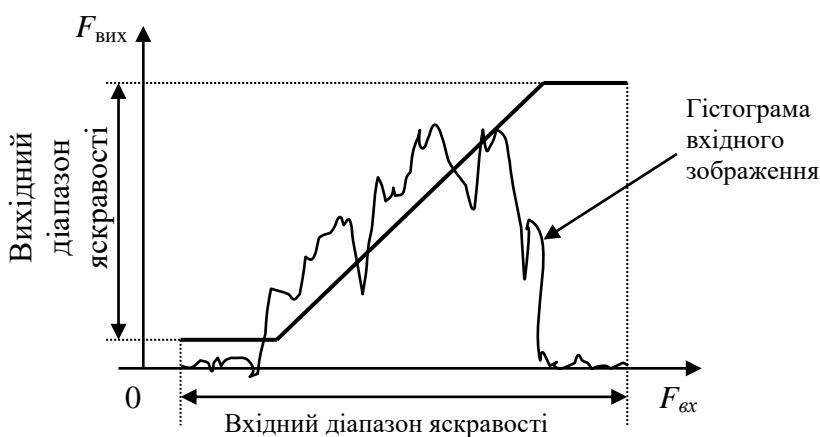
Рис.2.6. Перетворення контрасту.

Рис.2.6 ілюструє характеристику передачі рівнів, що потрібно для підвищення контрасту типових безперервних малокоонтрастних зображень. Корекцію таких зображень можна здійснити користуючись фотографічними методами, однак реалізувати довільну характеристику передачі рівнів з високою точністю, звичайно важко. У випадку цифрових зображень одержати необхідну характеристику передачі рівнів відносно просто.

Очевидно, що в діапазоні яскравостей вихідного зображення деякі рівні не будуть використовуватися, тому окремі перепади яскравості вихідного зображення будуть перевищувати відповідні перепади вхідного зображення. Унаслідок цього можуть ви-



а) лінійне масштабування діапазонів;



б) масштабування діапазонів з обмеженням.

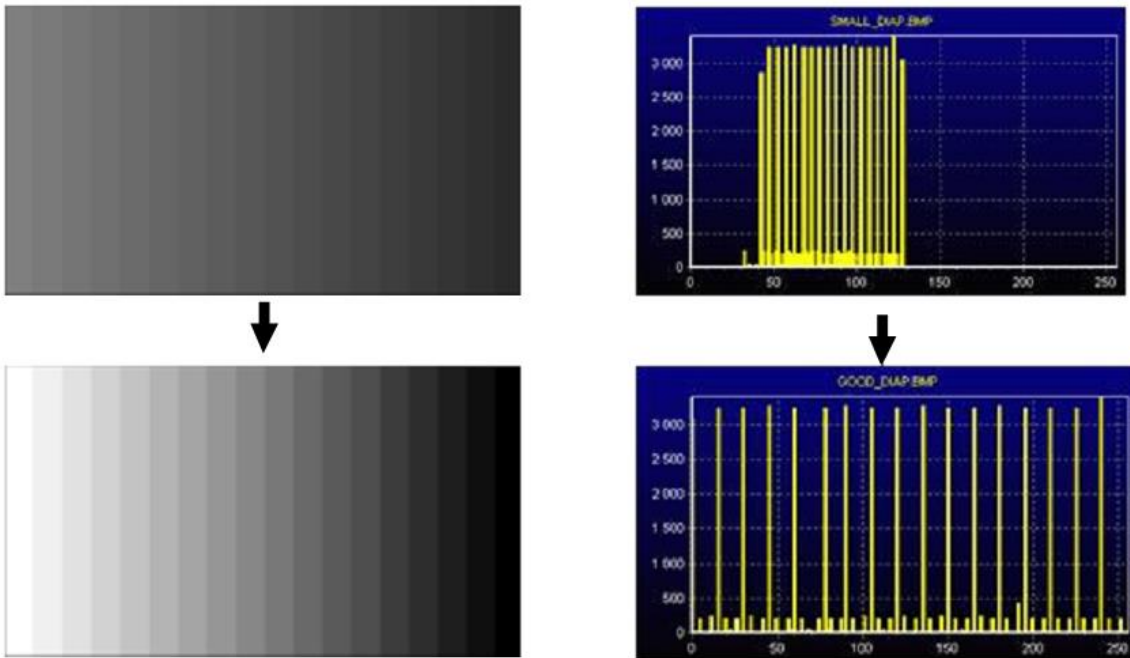


Рис.2.7. Масштабування діапазону контрасту.

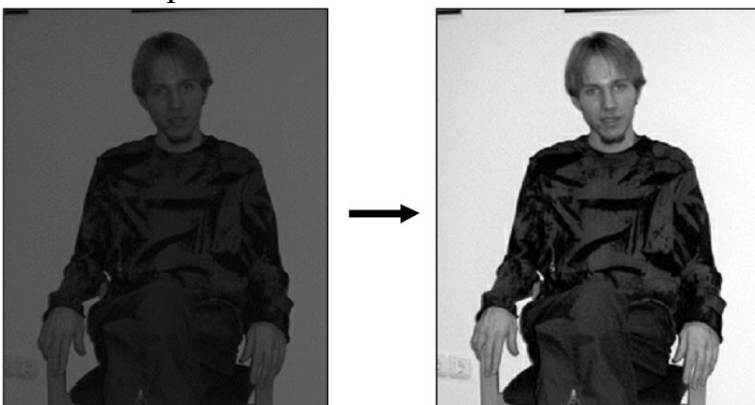
никнути помітні хибні контури [Л8]. Якщо вихідне зображення квантувати з більшим числом рівнів, чим вхідне, то можна одержати рівномірне розміщення вихідних рівнів і завдяки цьому зменшити ефект появи помилкових контурів.

Діапазон яскравостей вихідного зображення, підданого цифровій обробці, може відрізнитися від діапазону яскравостей вхідного зображення. Більш того, у діапазоні числових значень яскравості обробленого зображення можуть виявитися негативні значення, що не можуть мати прямої відповідності з фізичними яскравостями.

Рис. 2.7 ілюструє два можливих способи приведення діапазону яскравостей вихідного зображення у відповідність з діапазоном яскравостей вхідного зображення. Відповідно до першого з них, оброблене зображення лінійно відображається таким чином, щоб цілком охопити відведений йому діапазон яскравостей.

При наявності кольорового монітора з'являється можливість відображення в якомусь яскравому кольорі (наприклад, у червоному) ділянок одноколірного зображення, яскравість яких знаходиться в довільному вузькому інтервалі; інші ділянки цього зображення відображаються звичайним способом. Кожне зображення має два значення яскравості, що відповідають даному розряду кодового слова, що представляє яскравість. Метод яскравісних зрізів виявляється винятково ефективним при візуальному аналізі зображень у тому випадку, коли оператор користується дисплеєм у діалоговому режимі.

#### Лінійна корекція.



## Нелінійна корекція

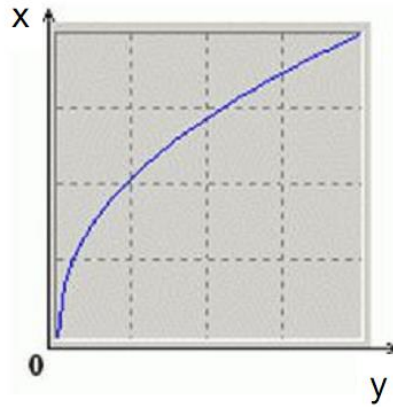
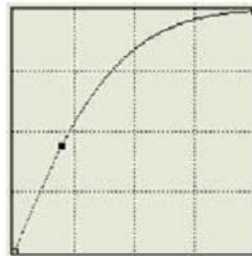
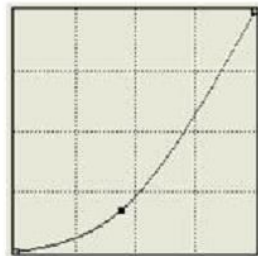


График функции  $f^{-1}(y)$



## Гама-корекція

Неправильный баланс



Корекція кольору  
Правильный баланс



## Корекція по шаблону Професійна корекція



Source: The dark knight



Source: <http://x-rite.com>

Используем цветной шаблон с многими цветами

Какое преобразование в камере?

## 4.2. Цифрові фільтри

### ШУМОПОДАВЛЕННЯ



Шум фотоаппарата



Сильное сжатие JPEG



Original



Salt and pepper noise

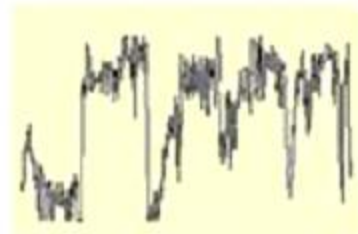
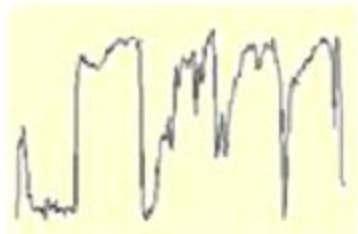


Impulse noise



Gaussian noise

Image  
Noise



Одним з найважливіших у обробці сигналів є клас лінійних перетворень, для яких аргументи сигналу до і після перетворення збігаються. Такі перетворення називаються лінійною фільтрацією, а їхнє цифрове подання називається цифровими фільтрами.

Найпоширенішими і найбільш простими в реалізації є фільтри інваріантні до зсування. Їхня імпульсна реакція залежить тільки від різниці аргументів.

### Загальні відомості про лінійну фільтрацію

Дія фільтра на вхідне зображення  $F(x,y)$  описується згорткою вхідного сигналу й імпульсної реакції фільтра  $H(x,y)$ , яка цілком описує властивості фільтра.

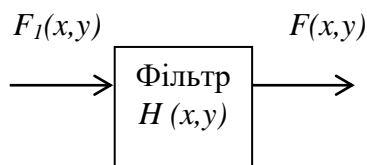


Рис.2.10. Фільтр двовимірних сигналів (зображень)

Для двовимірного сигналу (зображення) дія двовимірного фільтра (оптичного пристрою) описується двовимірною згорткою:

$$F_1(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(z, s) \cdot H(x - z, y - s) dz ds.$$

Тут функція  $F(x,y)$  – описує вхідний двовимірний сигнал (зображення),  $F_1(x,y)$  – описує вихідний двовимірний сигнал (зображення),  $H(x,y)$  – описує імпульсну реакцію (властивості) двовимірного фільтра (оптичного пристрою).

**Імпульсна реакція (характеристика) фільтра** (лінійної просторово інваріантної системи) визначається як відгук фільтра на дельта-функцію (рис.2.11). Для оптичних систем роль дельта-функції може виконувати зображення нескінченне яскравої точки на абсолютно чорному фоні.

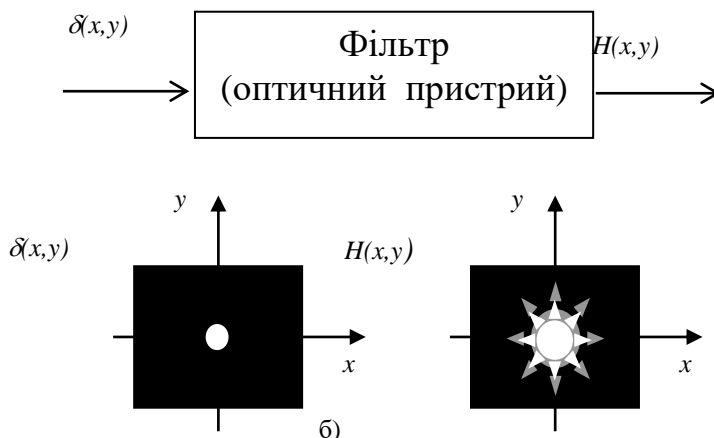


Рис.2.11. Спосіб одержання імпульсної реакції фільтрів для двовимірних фільтрів

Крім імпульсної реакції властивості лінійного фільтра можуть бути описані його **частотною характеристикою**  $H(w_x, w_y)$  у двовимірному випадку, які зв'язані з імпульсними реакціями парою перетворень Фур'є:

$$H(w_x, w_y) = FT \{H(x, y)\}, \quad H(x, y) = FT^{-1} \{H(w_x, w_y)\} -$$

для двовимірних фільтрів.

Тут  $FT\{*\}$  і  $FT^{-1}\{*\}$  – оператори відповідно прямого і зворотного перетворень Фур'є;  $w_x$  і  $w_y$  – горизонтальна і вертикальна просторові частоти.

*Стосовно до напівтонових зображень просторова частота - це кількість циклів зміни яскравості елементарних точок зображення від самого яскравого (білого) до самого темного (чорного) і назад уздовж одного з напрямків (координатних осей) за одиницю виміру координат. Одиниця виміру просторової частоти - 1/метр.*

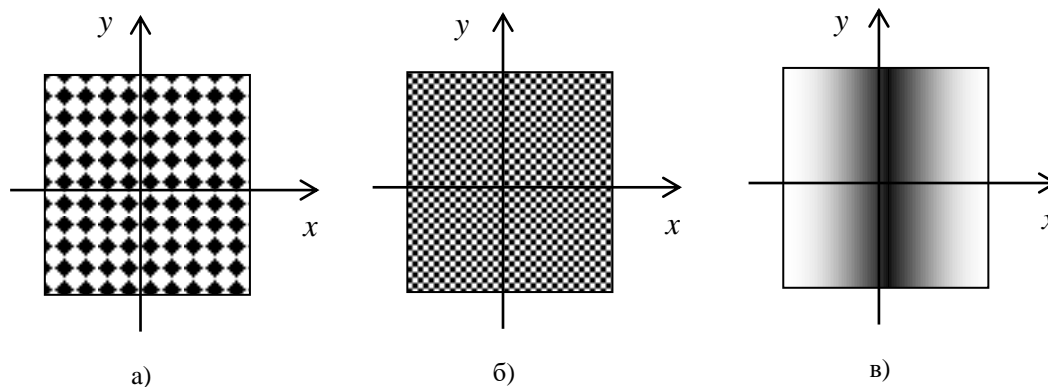


Рис.2.12. Приклади зображень просторових гармонік

На рис.2.12. приведено приклади зображень, що заповнені гармоніками з різними частотами. На рис.3а і рис.3б  $w_x = w_y$ , причому зображення на рис.3б заповнено більш високочастотним сигналом, чим зображення на рис.3а. На рис.3в показане зображення заповнене просторовою гармонікою з частотами  $w_x = 1, w_y = 0$ .

### Дискретні частотні й імпульсні характеристики фільтрів

Дискретна частотна характеристика двовимірного цифрового фільтра зв'язана з відліками його імпульсної реакції дискретним перетворенням Фур'є:

$$H(w_x, w_y) = \sum_{l=0}^{L-1} \sum_{n=0}^{N-1} H(l, n) \exp\{i2\pi(w_x \cdot l \cdot \Delta x + w_y \cdot n \cdot \Delta y)\}.$$

### Принципи реалізації двовимірної локальної фільтрації зображень

Нижче будуть розглянуті алгоритми реалізуючі цифрову фільтрацію зображень і об'єднані ідеєю фільтрації ковзним вікном чи апертурою. Термін "локальна" підкреслює факт, що розміри ковзного вікна (апертури) по обох осях менше розмірів зображення, що фільтрується.

Мета локальної фільтрації звичайно є в поліпшенні якості зображення; найчастіше це усунення перешкод, підвищення різкості, чи підкреслення контурів. Локальна фільтрація представляє для цього досить широкий арсенал засобів обробки.

Звичайний двовимірний фільтр улаштований у такий спосіб. Береться невелика, звичайно прямокутна, ділянка площини цифрового зображення (масиву чисел) і на ньому визначається деяка функція, яка визначає його імпульсну реакцію. Ця ділянка називається **апертурою**, чи вікном, а задана на ньому функція - ваговою функцією, чи функцією вікна. Таким чином, кожному елементу апертури (масиву) відповідає визначене число, назване **ваговим множником**. Сукупність усіх вагових множників і складає **вагову функцію**. Апертура разом із заданою на ній **ваговою функцією** часто називається **маскою фільтра**.

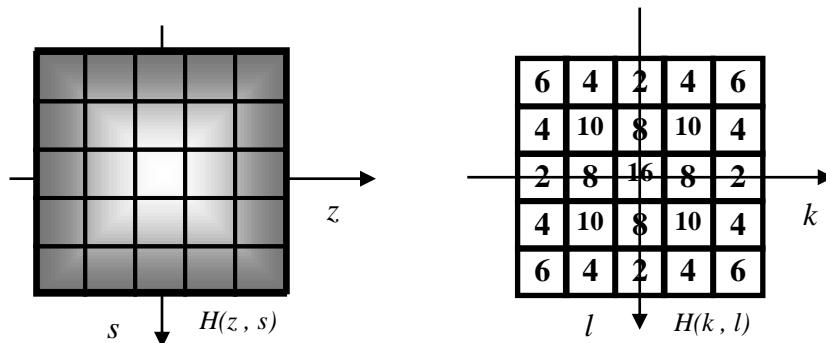


Рис.2.15. Імпульсна реакція двовимірного фільтра  $H(z,s)$  і відповідна маска цифрового фільтра  $H(k,l)$

Апертура звичайно має невеликий розмір, наприклад  $3 \times 3$  чи  $5 \times 5$  елементів. Лінійні розміри апертури звичайно беруться непарними, щоб можна було однозначно вказати її центральний елемент.

Цифровий масив відліків зображення  $F(i,j)$

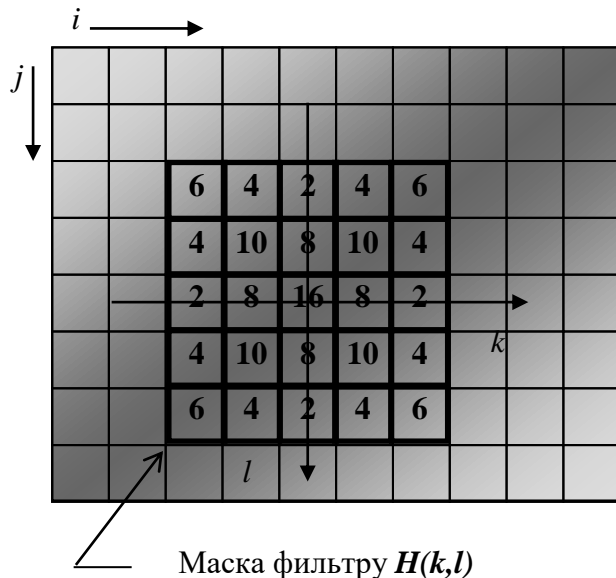


Рис.2.16. Цифровий масив відліків зображення  $F(i,j)$  з розташованою на ньому маскою фільтра  $H(k,l)$

Іноді застосовують "одномірну" апертуру, розмір якої по одному з напрямків дорівнює 1. "Одномірна" апертура, розташування якої не відповідає осям координат, може бути сформована завданням відповідної функції на звичайній, прямокутній апертурі.

Фільтрація здійснюється переміщенням апертури фільтра по зображенню. Переміщення здійснюється щораз на один відлік.

Характерним прикладом служить алгоритм лінійної фільтрації, що складає загалом у наступному. При кожному положенні апертури вагова функція поелементно помножується на значення відповідних елементів вихідного зображення; добутки додаються. Сума поділяється на нормуючий коефіцієнт і отримана величина, що є відгуком фільтра, присвоюється тому елементу нового зображення, що відповідає положенню центра апертури.

Коефіцієнт, що нормує, звичайно береться рівним сумі всіх елементів вагової функції (вагових множників). Якщо сума вагових множників дорівнює нулю, то коефіцієнт, що нормує, приймається рівним одиниці.

Введемо необхідні позначки. Нехай апертура має розмір  $(K \times L)$  елементів; поточний



елемент апертури позначимо через  $(k,l)$ , де  $l$  - поточний рядок;  $k$  - поточний стовпець.

Визначимо спосіб, за допомогою якого вказується положення апертури на зображенні. Виділяється деяка опорна точка апертури (звичайно це центр). Тепер досить задати положення цієї опорної точки в системі координат зображення, щоб тим самим визначити положення всієї апертури. Цю опорну точку називають умовним центром апертури. Координати умовного центра позначимо  $(k_0, l_0)$  (у системі координат апертури!).

Умовний центр може (але не зобов'язаний) збігатися з дійсним геометричним центром апертури. Узагалі, за умовний центр можна взяти будь-яку точку апертури; більш того, умовний центр не зобов'язаний навіть знаходитися у її середині. Однак, визначимо умовний центр так, щоб при непарних розмірах апертури він збігався з її центральним елементом:  $k_0 = \lfloor (K+1)/2 \rfloor$ ;  $l_0 = \lfloor (L+1)/2 \rfloor$ , де квадратні дужки позначають цілу частину числа.

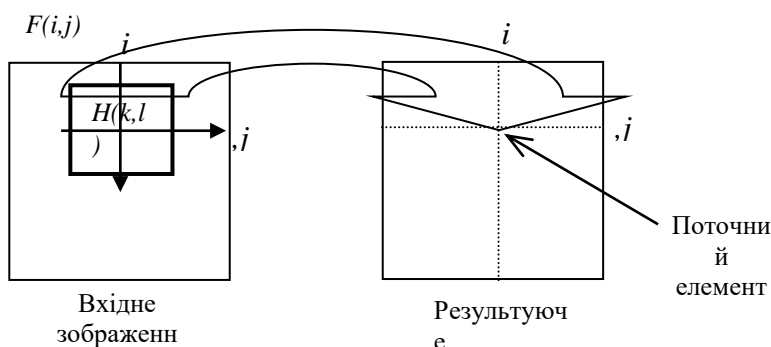


Рис.2.17. Робота нерекурсивного фільтра

Поточне положення умовного центра на вихідному зображенні  $F(i,j)$  позначимо через  $(i,j)$ . Відгук фільтра привласнюється тій же точці  $(i,j)$  нового, профільтованого зображення  $F_1(i,j)$ .

Позначимо тепер через  $h(k,l)$  функцію вікна. Масив вихідного зображення формується шляхом дискретної згортки вхідного поля  $F(i,j)$  і функції вікна  $h(k,l)$ :

$$F_1(i, j) = \frac{1}{K \cdot L} \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} h(k, l) \cdot F(i - k_0 + k, j - l_0 + l). \quad (2.7)$$

Точно говорячи, формула справедлива лише за умови, що функція вікна не виходить за межі зображення.

Поряд з лінійними фільтрами існують і нелінійні. Характер фільтра залежить від операцій, виконуваних у кожному положенні вікна. У лінійних фільтрах відгук є лінійною функцією багатьох перемінних, роль яких грають елементи вихідного зображення, що потрапили у вікно. Вагові множники - це коефіцієнти лінійної функції. Фільтри, у яких відгук не може бути виражений лінійною функцією від значень елементів зображення, є по визначенню нелінійним.

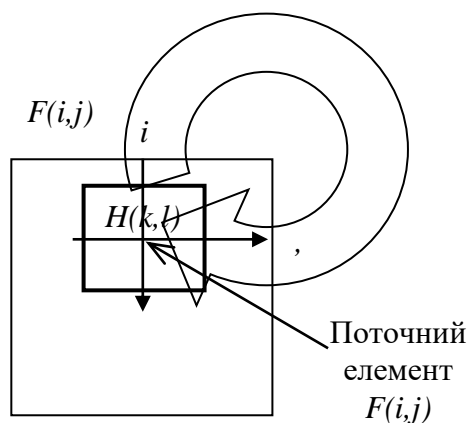


Рис.2.18. Робота рекурсивного фільтра

У залежності від того, куди (у яке поле пам'яті) записується відгук фільтра, розрізняють **прості (нерекурсивні)** і **рекурсивні** фільтри. Якщо в простих (нерекурсивних) фільтрах відгук записується у вихідне зображення  $F_1(i,j)$ , то в рекурсивних він записується назад у початкове зображення  $F(i,j)$ , змінюючи значення елементів початкового зображення безпосередньо в процесі фільтрації. Тому в рекурсивних фільтрах вже оброблені елементи зображення впливають на результат фільтрації наступних.

Іноді виявляється корисним багаторазове повторення процедури нерекурсивної фільтрації; у цих випадках говорять про ітеративне застосування фільтрів (не плутати з рекурсивною фільтрацією).

### Застосування локальної лінійної фільтрації для згладжування і підкреслення границь

Одне з найбільш розповсюджених застосувань лінійних фільтрів - згладжування шуму. Для цього застосовуються вагові функції наступного виду:

$$H(m,n) = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (\text{а}) \qquad H(m,n) = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (\text{б})$$

При частотній інтерпретації процесів фільтрації такий фільтр називається фільтром нижніх частот.

Для підкреслення ліній визначеного напрямку використовуються вагові функції виду

$$H(m,n) = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad (\text{в}) \qquad H(m,n) = \frac{1}{16} \begin{bmatrix} 2 & 1 & 2 \\ 1 & 4 & 1 \\ 2 & 1 & 2 \end{bmatrix} \quad (\text{г})$$

Вагова функція (в) підкреслює горизонтальні і вертикальні лінії вихідного зображення, (г) - діагональні лінії.

Фільтри, що виділяють границі (у частотній інтерпретації - це високочастотні фільтри), використовують наступні три вагові функції:

$$H(m,n) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} \quad (\text{д}) \qquad H(m,n) = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix} \quad (\text{е})$$

$$H(m, n) = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 5 & -2 \\ 1 & -2 & 1 \end{bmatrix} \text{ (ж)}$$

Слід звернути увагу на те, що сума вагових множників тут дорівнює одиниці. Тому на результуючих зображеннях зберігаються фонові області з постійною яскравістю. Границі виділяються незалежно від їхнього напрямку.

Для виділення перепадів яскравості визначеної орієнтації використовуються в залежності від необхідного напрямку вагові функції, називані курсовими градієнтними масками:

$$H(m, n) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & 1 \\ -1 & -1 & -1 \end{bmatrix}$$

(з) - "північ"

$$H(m, n) = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & -1 & 1 \end{bmatrix}$$

(и) - "північний схід"

$$H(m, n) = \begin{bmatrix} -1 & 1 & 1 \\ -1 & -2 & 1 \\ -1 & 1 & 1 \end{bmatrix}$$

(к) - "схід"

$$H(m, n) = \begin{bmatrix} -1 & -1 & 1 \\ -1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(л) - "південний схід"

$$H(m, n) = \begin{bmatrix} -1 & -1 & -1 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(м) - "південь"

$$H(m, n) = \begin{bmatrix} 1 & -1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & 1 \end{bmatrix}$$

(н) - "південний захід"

$$H(m, n) = \begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & -1 \\ 1 & 1 & -1 \end{bmatrix}$$

(о) - "захід"

$$H(m, n) = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -2 & -1 \\ 1 & -1 & -1 \end{bmatrix}$$

(п) - "північний захід"

Назва курсу говорить про напрямок перепаду яскравості, що викликає максимальний відгук фільтра.

Для виділення перепадів без указівки їхньої орієнтації використовуються наступні три види вагових функцій (оператори Лапласа):

$$H(m, n) = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

$$H(m, n) = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

р)

с)

$$H(m, n) = \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix}$$

т)

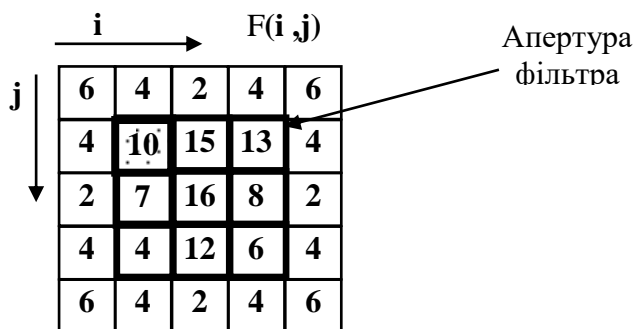
Ці оператори відрізняються від масок виду (д, е, ж) тим, що сума вагових множників у них дорівнює нулю. Це говорить про те, що на результуючому зображенні залишаться тільки виділені границі без тла.

### Нелінійна фільтрація

Основні поняття теорії локальної фільтрації, розглянуті раніше, справедливі і для локальної нелінійної фільтрації. Маються на увазі поняття процесу фільтрації, апертури, способи й алгоритми переміщення апертури по зображенню; вагова функція застосовується не завжди. Головна відмінність полягає в тому, що вихід нелінійного фільтра формується нелінійним образом від даних вихідного зображення.

### Медіанні фільтри

Медіанна фільтрація - метод нелінійної обробки сигналів. Медіанний фільтр являє собою ковзне вікно, що охоплює непарне число елементів зображення по рядку і стовпцю. Відгук медіанного фільтра дорівнює медіані даних, що знаходяться в апертурі. Медіана являє собою центральний елемент у варіаційному ряді, побудованому з даних, що знаходяться в межах апертури (рис 2.21). Ранг, значення і координати в апертурі медіанного елемента позначені штрихуванням.



Ранг R	0	1	2	3	4	5	6	7	8
$F(i+k, j+l)$	4	6	7	8	10	12	13	15	16
$(k, l)$	-1, 1	1, 1	-1, 0	1, 0	-1, -1	0, 1	1, -1	0, -1	0, 0

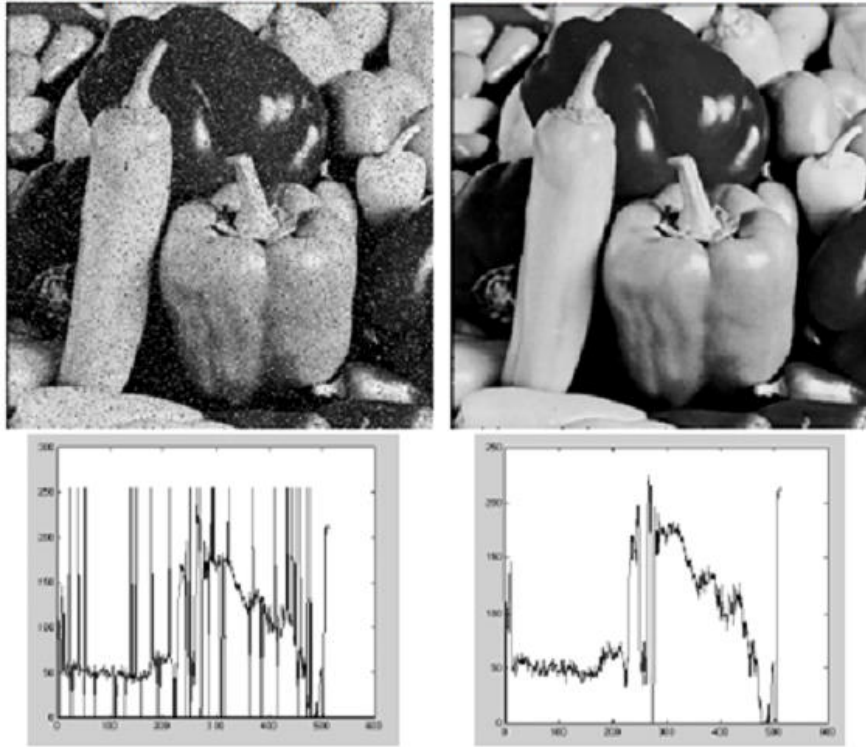
Рис.2.21. Варіаційний ряд елементів зображення, що потрапили в апертуру фільтра.

Для дискретної послідовності  $a_1, a_2, \dots, a_N$  і непарної довжини варіаційного ряду  $N$  медіаною є той її елемент, для якого існують  $(N-1)/2$  елементів, великих чи рівних йому по величині і  $(N-1)/2$  елементів, менших чи рівних йому по величині. У силу того, що для операції знаходження медіани не виконується одна з аксіом лінійності, медіанний фільтр є нелінійним.

Медіанні фільтри застосовуються для згладжування зображень і для придушення шуму. Раніше розглянуті лінійні низькочастотні фільтри, застосовувані для тих же цілей. Медіанні фільтри по своїх властивостях відрізняються від них. По-перше, медіанні фільтри зберігають різкі перепади, тоді як лінійні низькочастотні фільтри їх змазують. По-друге, медіанні фільтри дуже ефективні при згладжуванні імпульсного шуму, але можуть приводити до повного зникнення дрібних деталей зображення при неадекватному виборі параметрів фільтра.

Медіанні фільтри використовуються також для виявлення границь і виділення об'єктів.

Шум «соль и перец» Медианная фильтрация



### Нелінійні методи контрастування меж

Раніше були розглянуті лінійні методи підкреслення перепадів яскравості і виділення контурів. Відомі і нелінійні методи контрастування.

Один з таких методів запропонований Робертсом і складається у використанні операції двовимірного дискретного диференціювання

$$G(i, j) = \sqrt{U^2 + V^2}, \quad (2.10)$$

де

$$U = F(i, j) - F(i+1, j+1),$$

$$V = F(i, j+1) - F(i+1, j).$$

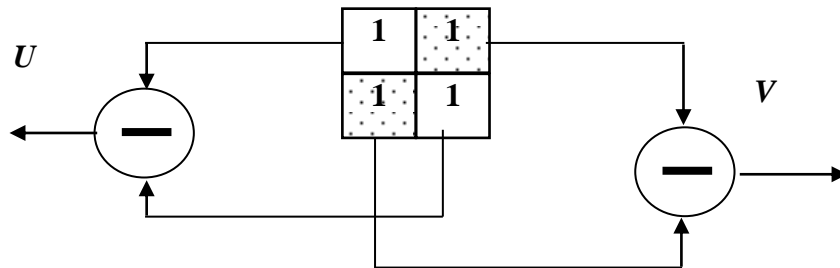


Рис.2.23. Обробка елементів зображення у фільтрі Робертса.

Інший нелінійний оператор контрастування, що використовує апертуру розміром 3x3, запропонований Собелом:

$$G(i, j) = \sqrt{X^2 + Y^2},$$

де

$$X = [F(i-1, j+1) + 2F(i, j+1) + F(i+1, j+1)] - [F(i-1, j-1) + 2F(i, j-1) + F(i+1, j-1)],$$

$$Y = [F(i-1, j-1) + 2F(i-1, j) + F(i-1, j+1)] - [F(i+1, j-1) + 2F(i+1, j) + F(i+1, j+1)].$$

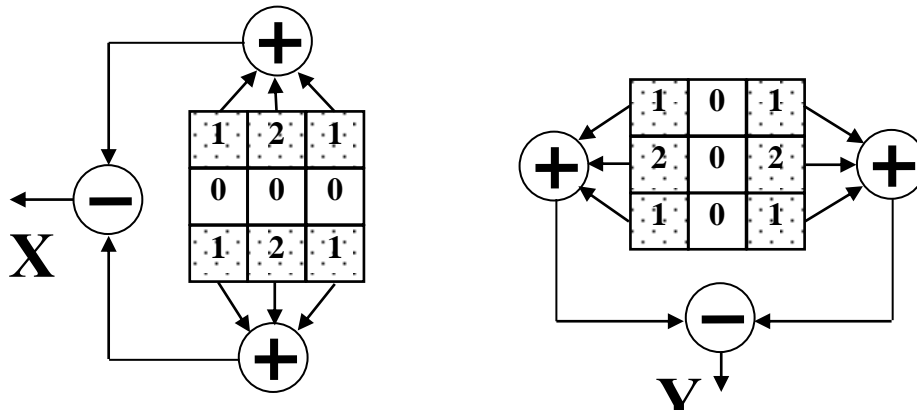
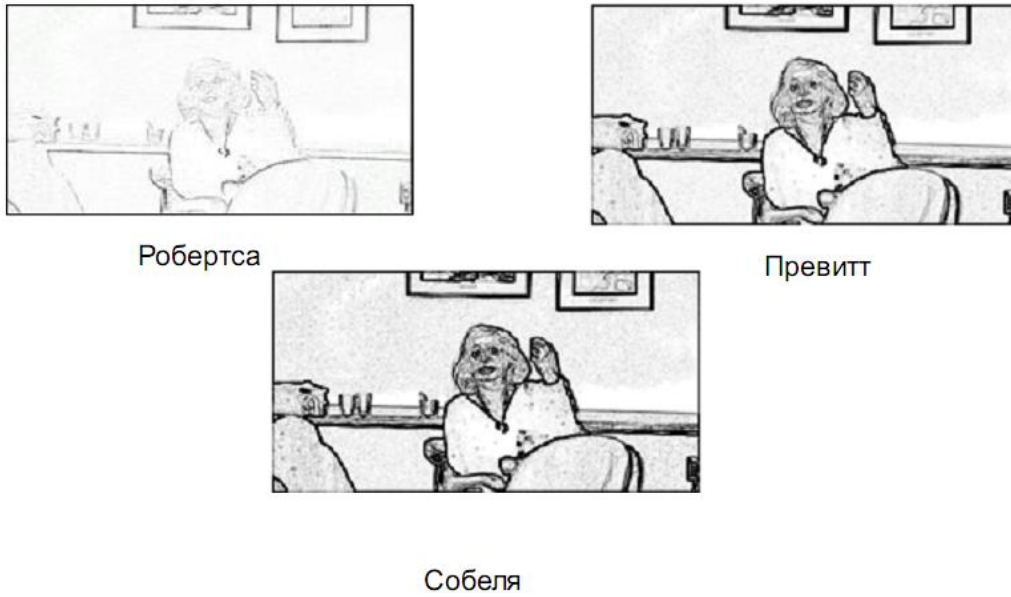


Рис.2.24. Обробка елементів зображення у фільтрі Собела.

Методи нелінійної обробки надзвичайно різноманітні. Ще за один приклад приведемо алгоритм, що виявляє на зображенні локальні максимуми (вершини).



Таким чином, Попередня обробка призначена для покращення якості зображень. Основними операціями попередньої обробки є радіометрична корекція зображень, усунення шумів та підкреслення меж.

ОЧИСТКА ЗОБРАЖЕННЯ ТА ПОШУК КРАЇВ – ЦЕ ЛИШЕ ПОЧАТОК!

#### 4. Виділення об'єктів на зображенні

Постановка завдання Виділення об'єктів на зображенні



Необхідно визначити чи є об'єкти на зображенні та, якщо є, то визначити їх місцеположення.

## ПРОБЛЕМА: МІНЛИВІСТЬ ЗОБРАЖЕНЬ



### Зовнішні чинники:

- Положення камери.
- Внутрішні властивості камери.
- Освітлення.

### Внутрішні чинники:

- Деформації об'єктів.
- внутрішньокласова мінливість.

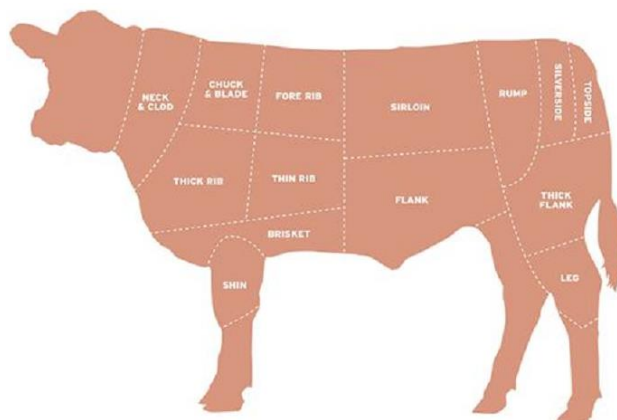
ТОМУ

- Поки доводиться завдання спрощувати, вводячи обмеження на ракурс зйомки, умови освітлення, типи об'єктів.
- Ми будемо розглядати прості випадки, коли всі чинники змінюються незначно.

З чого складається зображення?



Із «шматків» - окремих об'єктів.



## Ознаки об'єктів на зображенні

Ознакою зображення об'єкта називається його найпростіша відмінна характеристика або властивість. Деякі ознаки є природними в тому розумінні, що вони встановлюються візуальним аналізом зображення, тоді як інші, так звані штучні ознаки, виходять у результаті його спеціальної обробки або вимірів. До природних ознак відносяться світлота (яскравість) і текстура різних областей зображення, форма контурів об'єктів. Гістограми розподілу яскравості й спектри просторових частот дають приклади штучних ознак.

### Яскравісні ознаки

Найбільш важливою ознакою об'єкта на зображенні є розподіл по простору світлоти або пов'язаних з нею величин: яскравості, спектральної інтенсивності випромінювання, координат кольору й т.д. Ці величини називаються яскравісними ознаками. Вимір яскравісних ознак можна проводити або в окремих точках зображення, або на ділянках простору зображення.

При використанні значень яскравості точок у деякій області, де приблизно перебуває об'єкт мова найчастіше йде про порівняння безпосередньо зображення ділянки аналізу й зображення еталонного об'єкта. У такому випадку для виділення характерної ознаки використовують лінійний фільтр, апертура якого більше або дорівнює розміру об'єкта виявлення. Вихідний ефект такого фільтра пропорційний схожості ділянки зображення під вікном фільтра й еталонного зображення об'єкта. Він буде максимальним тоді, коли зображення об'єкта у вікні в точності збіжиться із зображенням еталона.

Оскільки інформацію про положення шуканих об'єктів несуть максимуми на результаті фільтрації, те їх можна виділити в масиві значень  $P(i,j)$  яким-небудь алгоритмом виявлення (наприклад ранговим голосуванням або нелінійним алгоритмом виділення максимумів).

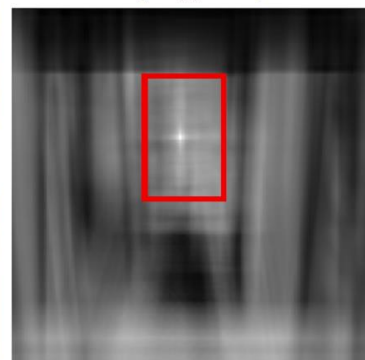
Істотний недолік методу безпосереднього зіставлення зображення об'єкта із зображенням еталона полягає в залежності виходу фільтра від енергії конкретної ділянки зображення. Він частково усувається шляхом переходу до аналізу контурних препаратів зображень. Еталон у цьому випадку також представляється контурним препаратом. При такому підході говорити про яскравісні ознаки не доводиться. Цей підхід ближче до ознак форми.

Головний недолік порівняння безпосередньо зображень складається необхідності використання величезної кількості еталонів для врахування змін об'єктів, що виникають при їхніх геометричних трансформаціях (повороті й зміні розмірів). Із цієї причини бажано обмежитися ознаками, які менше залежать від змін розміру й форми об'єкта при геометричних викривленнях.

### Найпростіший класифікатор (приклад)

Найдём стул в изображении

Выход корреляции



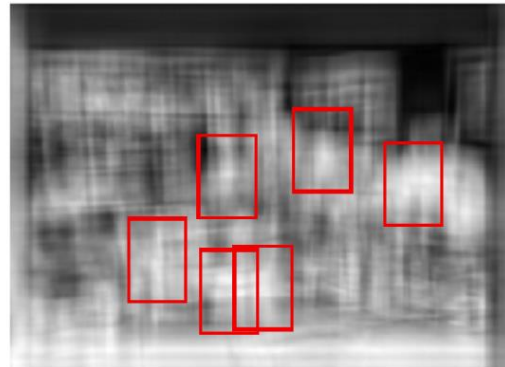
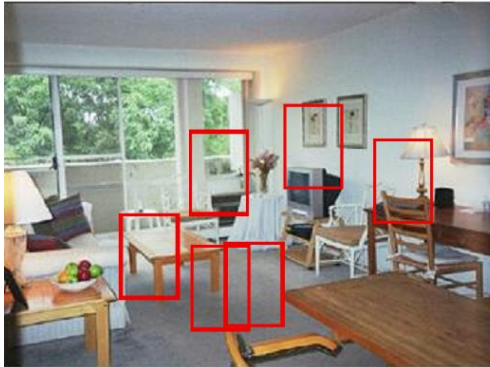


Зіставлення шаблонів («Pattern matching»):

- Фіксуємо зображення об'єкта («шаблон»)
- Класифікатор - порівняння шаблону і поточного вікна (Фрагмента)
- Підсумком ковзного вікна буде карта відгуків (score)



Найдём стул в этом изображении



«Класифікатор» у вигляді порівняння з шаблоном не може врахувати всю множину факторів мінливості.

### **Контурні ознаки**

Різкі зміни (розриви) яскравості або параметрів текстури, є важливими найпростішими ознаками, оскільки вони часто визначають форму зображених об'єктів. Локальні розриви значень яскравості називають яскравістними перепадами, або контурами. Вихідне зображення, представлене масивом чисел  $F(i,j)$ , піддається лінійній або нелінійній обробці для того, щоб підсилити перепади яскравості. У результаті утвориться масив чисел, що описує зображення з підкресленими змінами яскравостей. Для позначення положення перепадів на зображенні часто формують контурний перепад. Ідеальний детектор перепаду повинен указувати на наявність перепаду в єдиній крапці, розташованій в центрі переходу яскравості.

*Для кожної області обмеженої контуром можна підрахувати якусь набір найпростіших числових характеристик (і це вже будуть штучні ознаки):*

- Площа
- Центр мас
- Периметр
- Компактність
- Орієнтацію головної осі інерції
- Подовження (ексцентриситет)



**Ознаки плям**

Плямою називається відносно маленька область зображення, яскравість якого значно відрізняється від яскравості її околиці [Л8]. Для виявлення плями спочатку здійснюється згладжування зображення фільтром низьких частот з маскою розміром  $N \times N$  елементів. Потім рівень кожного елемента усередненого зображення порівнюється із середнім значенням рівнів його чотирьох сусідів (праворуч, ліворуч, зверху й знизу), які віддалені від центрального елемента на відстань в  $N$  елементів. Пляма виявляється, якщо ця різниця досить велика. Інший подібний підхід передбачає формування різниці між середньою яскравістю у вікні розміром  $N \times N$  і середньою яскравістю навколишньої кільцевої області шириною  $N$ .

### Ознаки ліній.

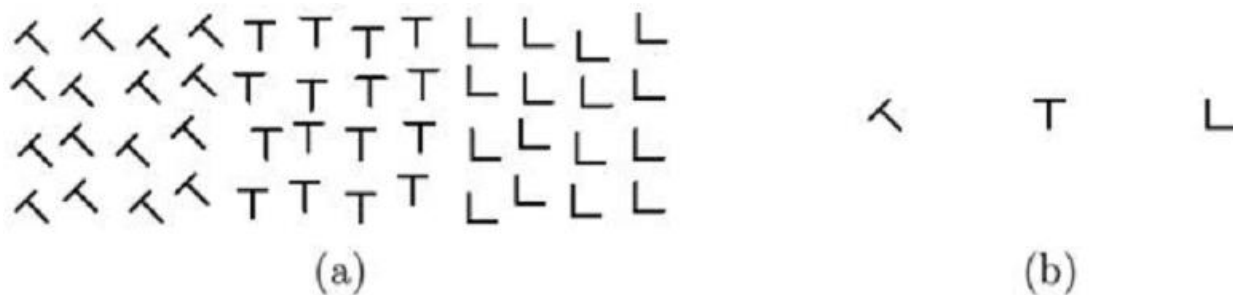
Можна припустити, що лінії й смуги є локальними фрагментами більших структур. Один з підходів до виділення лінії полягає в тому, що виконується просторове об'єднання відгуків детектора перепаду усередині локальної околиці. Лінії одиничної ширини можна виявити шляхом згортки зображення з набором лінійних масок за допомогою яких виконується контрастування вертикальних, горизонтальних і діагональних відрізків лінії [Л8].

$$\begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & 2 \\ -1 & 2 & -1 \\ 2 & -1 & -1 \end{bmatrix} \quad \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

Рис.2.44. Маски фільтрів для виділення ознаки ліній

### Текстурні ознаки

**Текстурою** в теорії обробки зображень називають структуру, яка характеризується наявністю «малюнка», що повторюється, складається з деяких однорідних ділянок приблизно однакових розмірів.



- (b) - типові приклади шаблонів (текстурних елементів) для досліджень психофізіологічного сприйняття зображень
- Людина явно враховує властивості текстури фрагмента при розпізнаванні зображень і виділенні областей

Більша частина природних сцен позбавлена істотних деталей на більших просторах. У цих областях сцену часто можна охарактеризувати як прояв повторюваної структури (малюнка). Текстуру можна розділити на штучну й природну. Штучні текстури - це структури із графічних знаків, розташованих на нейтральному фоні. Природні текстури - це зображення природних сцен, що містять майже періодичні структури. Прикладами можуть служити зображення черепиці дахів, піску, трави й т.д.

Потрібно враховувати, що текстура - це властивість околиці точки зображення. Отже, текстурні ознаки по своїй природі залежать від розміру вікна аналізу.

### Гістограмні ознаки

Форма гістограми розподілу частот дає багато відомостей про властивості зображення. Наприклад, вузька гістограма вказує на низький контраст, а бімодальна гістограма припускає наявність областей різної яскравості.

### Сегментація

**Сегментація** - це спосіб поділу сцени на «Шматки», з якими простіше працювати

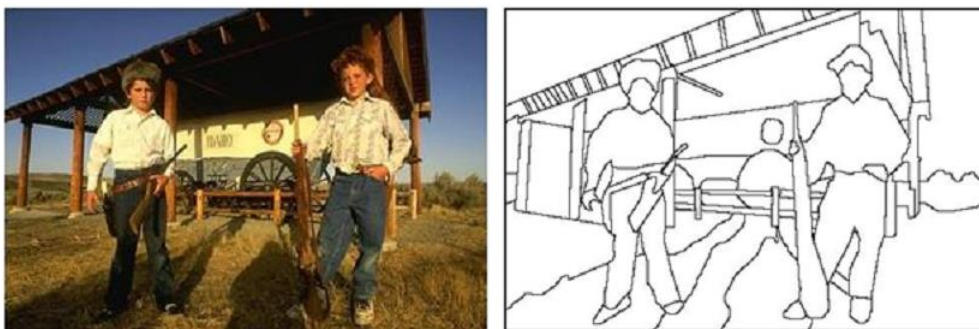
**Метою сегментації** є виділення областей (сегментів), однорідних в якомусь певному заданому сенсі. Однорідність є ознакою приналежності області до певного класу.

Дуже часто сегментація використовується для виділення областей приблизно однакового тону і/або кольору, а також областей, однорідних в сенсі деякої складнішої властивості (наприклад, типу текстури). Такі області прийнято називати **кластерами**.

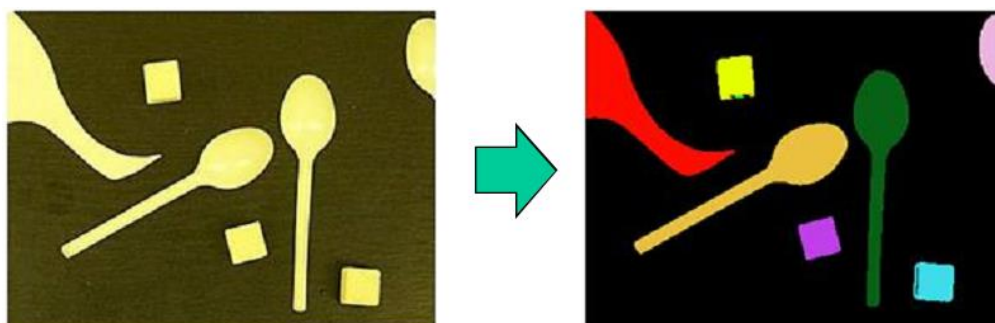


Вимоги до сегментації

- Сегментація - це спосіб поділу сцени на «Шматки», з якими простіше працювати
- Межі сегментів повинні відповідати границям об'єктів



Результат сегментації



Як можна записувати результат сегментації?

- Зробимо карту розмітки - зображення, в кожному пікселі якого номер сегмента, якому належить цей піксель
- Візуалізувати зручно кожен сегмент своїм кольором

## Сегментація, що заснована на методах класифікації

*Алгоритми сегментації зображень, що засновані на методах класифікації, можна розбити на такі основні групи:*

**методи неконтрольованої класифікації** (припускають мінімізацію або максимізацію якого-небудь критерію якості кластеризації);

**методи контрольованої класифікації** (припускають наявність сформованих еталонних образів)

**методи аналізу гістограмних ознак** (як правило, у цих алгоритмах застосовують різні методи виділення мод гістограми);

**методи, засновані на теорії графів** припускають побудови, згортання й розрізування графів, у яких вершини відповідають кластерам, а дуги зв'язкам цих кластерів.

### **Методи сегментації зображень:**

#### 1. Сегментація

##### 1.1. Неконтрольована класифікація

- K – внутрішньо групових середніх;
- ISODATA;
- динамічні згущення.

##### 1.2. Контрольована класифікація

- спосіб спектрального кута;
- спосіб мінімальної відстані;
- спосіб паралелепіпедів;
- спосіб максимальної правдоподібності;
- спосіб дистанції Махаланобіса;
- спосіб бінарного кодування.

##### 1.3. Аналіз гістограмних ознак

- по одновимірним гістограмам;
- по багатовимірним гістограмам.

##### 1.4. Методи теорії графів

#### 2. Контурна сегментація

#### 3. Нарощування ланок

## Приклад кластеризації



Исходное изображение

Кластеры по яркости

Кластеры по цвету

Вектор признак – {яркость}

Вектор признак – {R, G, B}

### Особливості методу контрольованої класифікації з використанням еталонних ділянок

Особливістю даного підходу є те, що за еталонні образи можуть використовуватися підмножини елементів вхідного зображення – еталонні ділянки. Вибір еталонних ділянок має на увазі вибір фрагментів знімка, однорідних по яскравості й розташуванню (рис. 2.49).

Належність еталонних ділянок до тих або інших класів досліджуваних об'єктів визначається заздалегідь:

на підставі попереднього візуального дешифрування із залученням додаткових дешифрувальних ознак;

за результатами польового дешифрування;

за результатами попередніх досліджень, проведених на даній ділянці, або на підставі іншої доступної інформації.

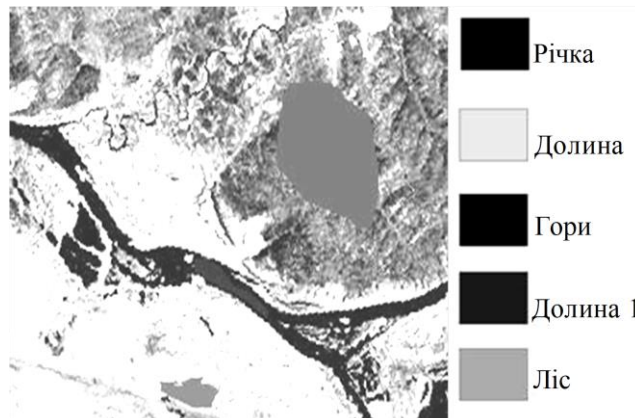
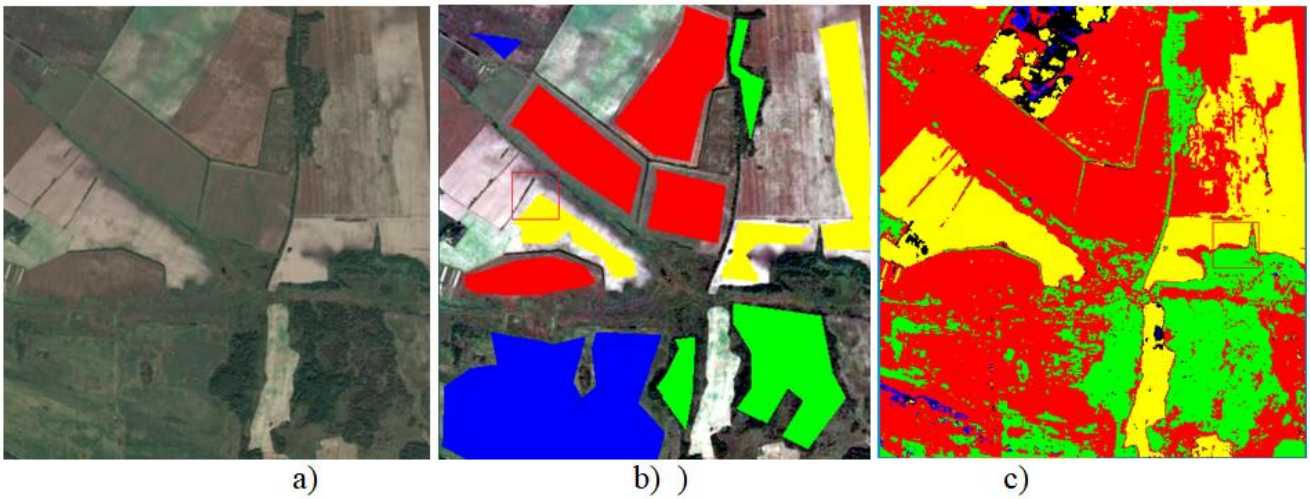
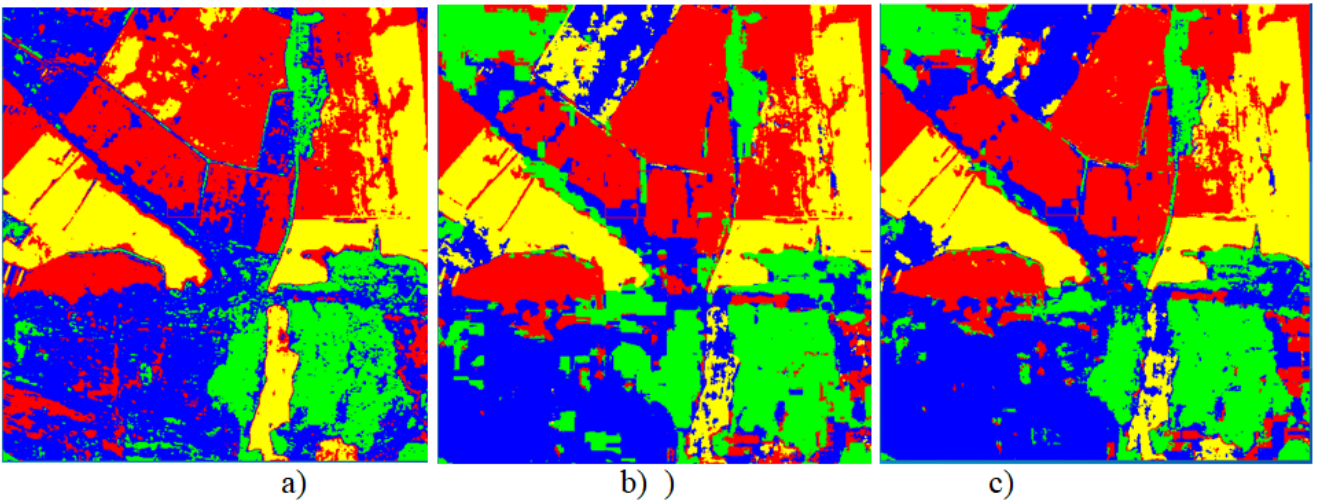


Рис. 2.49. Ілюстрація класифікації методом максимальної правдоподібності



*Figure 1 Image preparation and classification: a) initial image; b) selection of ROI; c) the image is classified by the Parallelepiped method*



*Figure 2 Classification results: a) classification by the Minimum Distance method; b) classification by the Mahalanobis Distance method; c) classification by the Maximum Likelihood method*

## 6. Бібліотека комп'ютерного зору OpenCV

Офіційна сторінка: <http://opencv.org>

Фактично найпопулярніша бібліотека комп'ютерного зору.

- Написана на C/C++, вихідний код відкритий, містить понад 1000 функцій та алгоритмів.

- Ліцензія BSD (дозволяється безкоштовне використання вдома, для навчання, на роботі)

- Розробляється з 1998 року, спочатку в Інтел, тепер у компанії Itseez за активної участі спільноти.

- >5000000 завантажень (без урахування svn/git трафіку)

- Використовується багатьма компаніями, організаціями, ВНЗ, наприклад NVidia, Willow Garage, Intel, Google, Stanford ...

*Деякі приклади використання:*

- Система зору робота PR2, зробленого WillowGarage

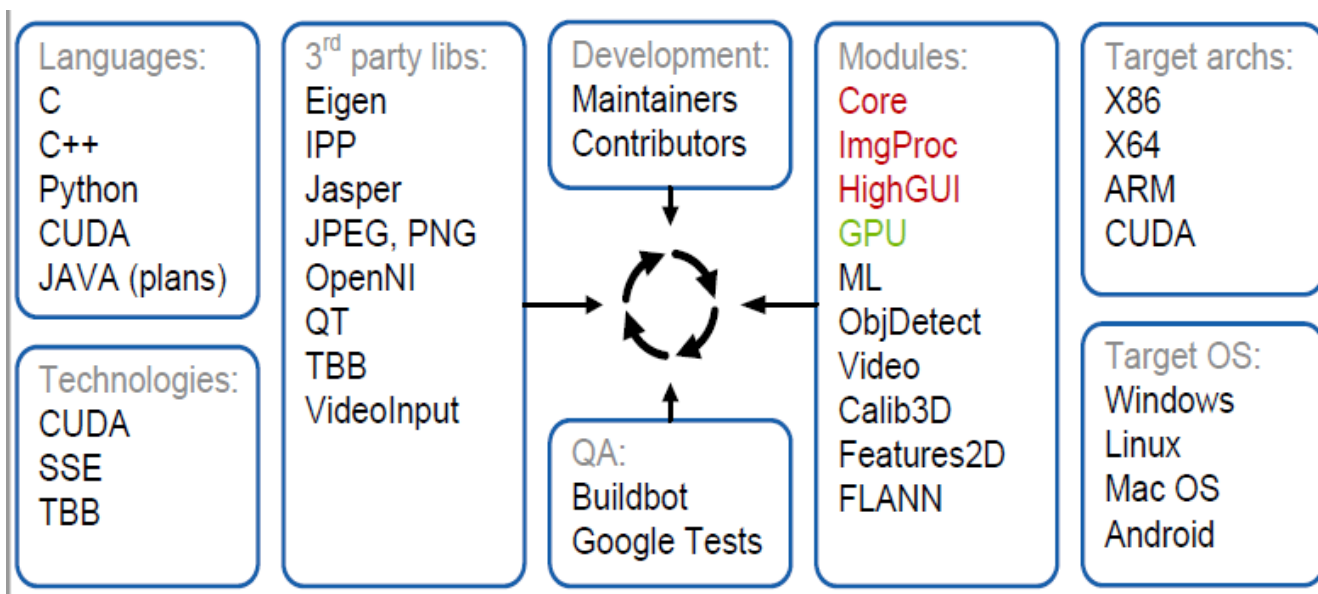
- Аудіо-візуальна інсталяція в Музеї Сучасного Мистецтва (Сан-Франциско)

- Контроль якості монет, що виготовляються Центробанком Китаю.

- Курси комп'ютерного зору в Стенфорді

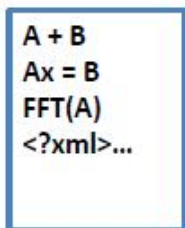
• Панорами вулиць у картах Google

## Архітектура та розробка OpenCV



## Функціональність

### Базовая функциональность



### Обработка изображений



Фильтрация



Трансформации



Ребра,  
контурный  
анализ



Особые точки



Сегментация

### Видео, Стерео, 3D



Калибрация  
камер



Вычисление  
положения в  
пространстве



Оптический  
поток

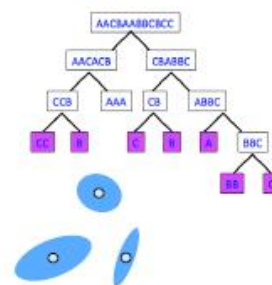


Построение  
карты глубины



Нахождение  
объектов

### Машинное обучение

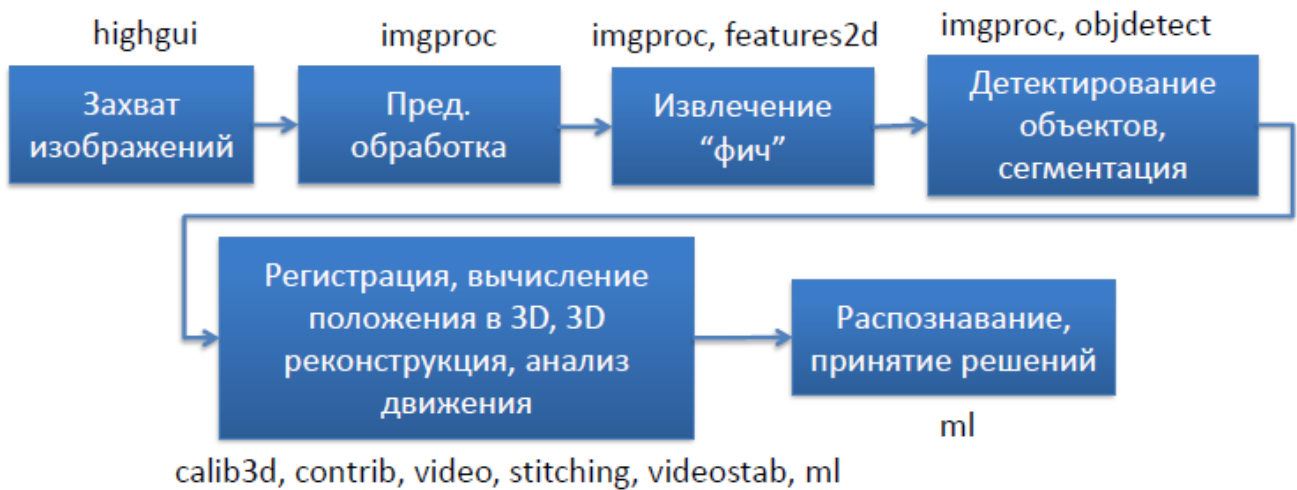


## OpenCV у додатках комп'ютерного зору

OpenCV - базова, в цілому низькорівнева бібліотека.

Вона надає будівельні блоки для додатків. Самі додатки доведеться побудувати вам, користувачам.

## Загальна схема типового додатка CV



Далі справа за вами!

## ВИСНОВКИ

Сучасний комп'ютерний зір застосовується у багатьох областях життєдіяльності людини де є зображення. Він тісно пов'язаний з іншими розділами штучного інтелекту (машинне навчання, нейронні мережі та інші).

Методів і підходів у комп'ютерному зорі дуже багато тому ця область потребує додаткового самостійного вивчення.