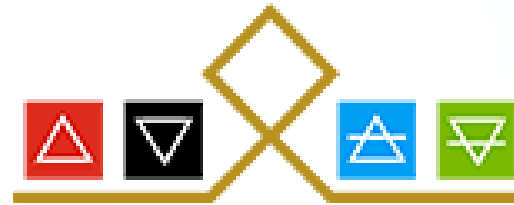




# НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ



## Теорія розпізнавання образів та класифікації в системах штучного інтелекту

*Тема №4. Розпізнавання зображень,  
комп'ютерний зір*

Київ - 2024

# Зміст

- 1. Підходи до розпізнавання зображень.*
- 2. Локалізація об'єктів на зображеннях.*
- 3. Ознаки об'єктів на зображеннях.*
- 4. Препарування зображень.*
- 5. Комп'ютерний зір.*

# Мета лекції

*Вивчення та дослідження основних підходів до розпізнавання зорових образів*

# 1. Підходи до розпізнавання зображень

В загальній теорії розпізнавання образів існує кілька різних підходів до розпізнавання, що відрізняються структурою опису класів ознак:

- детермінований підхід;
- безперервно-груповий підхід;
- структурно-лінгвістичний (синтаксичний) підхід;
- статистичний підхід;
- нейрокомп'ютерний підхід.

## Детермінований підхід

Припускає, що в будь-якій точці простору ознак, використовуваних для розпізнавання, з ненульовою апіорною ймовірністю можуть з'явитися реалізації тільки одного класу. При цьому необхідно із усієї сукупності ознак, що характеризують розпізнавані об'єкти, виділити систему детермінованих властивостей, тобто таких ознак, визначення яких не пов'язане з тими або іншими принципово випадковими механізмами (нормовані інтегральні геометричні параметри, компоненти розкладання зображення по якій-небудь системі функцій, моменти зображення й т.п.).

Один з розповсюджених способів - розпізнавання букв, що заснований на кореляційно-екстремальному алгоритмі (в пам'яті обчислювальної системи повинні зберігатися еталони всіх символів).

## Безперервно-груповий підхід

Підхід є розвитком детермінованого підходу. Вважається, що кожний розглянутий клас породжується в результаті апріорі невідомих топологічних описуваних безперервними групами перетворень деякої еталонної реалізації (наприклад, для обробки геометрично перекручених зображень, одержуваних при дистанційних дослідженнях Землі з космосу, еталонною реалізацією служить географічна карта).

Кожний клас у просторі ознак є орбітою своєї еталонної реалізації, що утворена групою припустимих перетворень просторів ознак.

При безперервно-груповому підході в якості ознак виступають інваріанти діючої в просторі ознак групи. У рамках цього підходу можна безпосередньо визначати параметри положення еталонного об'єкта в пропонованій реалізації.

Через невизначеність положення той самий об'єкт (наприклад, озеро) може всіляким образом представлятися на зображенні. Для математичного опису перетворень зображень використовують **математичний апарат теорії безперервних груп**. Елементами групи є різні перетворення зображень, а в якості групової операції виступає композиція цих перетворень.

При постановці завдання розпізнавання зображень у рамках безперервно-групового підходу насамперед необхідно встановити, яка із груп діє на розпізнаванім зображенні.

Допустимо, що є деяка група  $G$ . Розглянемо деякий еталонний об'єкт, наявність аналога якого на зображенні ми припускаємо. На реальному зображенні цей аналог буде виглядати як  $gs$ , де  $g$  — апріорно невідомий елемент діючої на зображенні групи  $G$ .

При розпізнаванні зображень у цих умовах звичайно розрізняють два типи (етапи) завдань:

ідентифікація об'єкта, виділеного на зображенні, з еталоном  $S$ ;

визначення параметрів положення об'єкта щодо датчика зображення, тобто визначення параметрів (ідентифікація) перетворення  $g$ .

У цьому випадку існують два основні підходи до розв'язку завдання розпізнавання:

- за допомогою нормалізації;
- за допомогою інваріантних ознак.

Нормалізацією зображення, підданого деякому апріорно невідомому перетворенню  $g$  групи  $G$ , називають приведення зображення до так званого нормалізованого виду, тобто до виду, що збігається з відповідним еталоном. Інакше кажучи, нормалізація — це усунення викривлень, викликаних дією перетворення  $g$ . Потім нормалізоване зображення ототожнюють із одним з наявних еталонів. Після цього, знаючи етalon  $S$  і його вихідне зображення  $gs$ , обчислюють параметри перетворення  $g$ , тобто вирішують завдання визначення параметрів положення.

При розпізнаванні зображень за допомогою нормалізації важливе місце займає поняття еквівалентності. Говорять, що зображення  $f_1$  і  $f_2$  еквівалентні щодо дії групи перетворень  $G$ , якщо знайдеться таке перетворення  $g$ , при якому  $gf_1 = f_2$ .

Множину усіх зображень, еквівалентних (щодо дії групи перетворень  $G$ ) заданому еталону  $S$ , називають орбітою  $S$ , а множину еквівалентних зображень — класом еквівалентності.

Найпростішим прикладом використання нормалізації може служити випадок, коли  $G = G_c$  — група зсувів. У цьому випадку класом еквівалентності еталона  $S$  є множину зображень, отриманих з  $S$  шляхом різних зрушень. Нехай на еталонах присутній мітка, що вказує на (умовний) центр, при цьому кожному класу еквівалентності відповідає один, і тільки один еталон. На вихідному, пропонованому для розпізнавання зображенні  $(x, y)$  визначають центр  $(x_c, y_c)$ . Потім  $(x, y)$  нормалізують (центрують), тобто піддають перетворенню  $g$ .

$$x' = x - (x_c - x_0), y' = y - (y_c - y_0),$$

де  $(x_0, y_0)$  — геометричний центр зображення. Отримане центроване зображення  $gf(x, y) = f(x', y')$  ототожнюють із одним з наявних еталонів (які також центровані).

Параметрами положення в даному прикладі є координати центру щодо геометричного центру зображення, тобто  $x_c - x_0$  і  $y_c - y_0$ .

Розпізнавання зображень за допомогою інваріантних ознак припускає перехід із простору зображень у простір ознак. Вибір (синтез) ознак здійснюється так, щоб еквівалентним зображенням відповідали однакові ознаки й однакові вектори ознак. В умовах дії деякої групи перетворень зображень ознаки зображень повинні бути інваріантні щодо перетворень даної діючої групи. Загальне теоретичне приймання синтезу системи інваріантних ознак і нормалізаторів (перетворень нормалізації) будується на основі теорії диференціальних рівнянь у часткових похідних першого порядку.

Типовим прикладом завдання розпізнавання зображень за допомогою інваріантних ознак може служити завдання орієнтації космічного апарата по зображеннях ділянок зоряного неба. При наявності в кадрі  $k$  зірок з координатами  $(x_i, y_i), i = \overline{1, k}$ , повна система елементарних інваріантних ознак для цього випадку (в умовах дії групи  $G_{СП}$  — зсуву й повороту):  $r_i^2, i = \overline{1, k-1}; r_i r_{i+1}, i = \overline{1, k-2}$ , де  $r_i$  — двовимірний вектор, проведений на зображенні від  $i-1$  точки до  $i$ -1.

Найпростішим прикладом використання нормалізації може служити випадок, коли  $G = G_c$  — група зсувів. У цьому випадку класом еквівалентності еталона  $S$  є множину зображень, отриманих з  $S$  шляхом різних зрушень. Нехай на еталонах присутній мітка, що вказує на (умовний) центр, при цьому кожному класу еквівалентності відповідає один, і тільки один еталон. На вихідному, пропонованому для розпізнавання зображенні  $(x, y)$  визначають центр  $(x_c, y_c)$ . Потім  $(x, y)$  нормалізують (центрують), тобто піддають перетворенню  $g$ .

$$x' = x - (x_c - x_0), y' = y - (y_c - y_0),$$

де  $(x_0, y_0)$  — геометричний центр зображення. Отримане центроване зображення  $gf(x, y) = f(x', y')$  ототожнюють із одним з наявних еталонів (які також центровані).

Параметрами положення в даному прикладі є координати центру щодо геометричного центру зображення, тобто  $x_c - x_0$  і  $y_c - y_0$ .

Розпізнавання зображень за допомогою інваріантних ознак припускає перехід із простору зображень у простір ознак. Вибір (синтез) ознак здійснюється так, щоб еквівалентним зображенням відповідали однакові ознаки й однакові вектори ознак. В умовах дії деякої групи перетворень зображень ознаки зображень повинні бути інваріантні щодо перетворень даної діючої групи. Загальне теоретичне приймання синтезу системи інваріантних ознак і нормалізаторів (перетворень нормалізації) будується на основі теорії диференціальних рівнянь у часткових похідних першого порядку.

Типовим прикладом завдання розпізнавання зображень за допомогою інваріантних ознак може служити завдання орієнтації космічного апарата по зображеннях ділянок зоряного неба. При наявності в кадрі  $k$  зірок з координатами  $(x_i, y_i), i = \overline{1, k}$ , повна система елементарних інваріантних ознак для цього випадку (в умовах дії групи  $G_{СП}$  — зсуву й повороту):  $r_i^2, i = \overline{1, k-1}; r_i r_{i+1}, i = \overline{1, k-2}$ , де  $r_i$  — двовимірний вектор, проведений на зображенні від  $i-1$  крапки до  $i$ .

## Структурно-лінгвістичний (синтаксичний) підхід

Підхід заснований на описі геометричної структури класу за допомогою спеціальної формальної мови, алфавіт якого становлять так звані не виробничі елементи структури й операції їх композиції, що задаються граматиною цієї мови.

Підхід припускає, що структура й характер зв'язків між елементами всіх підлягаючих розпізнаванню об'єктів підкоряються цілком певним, заздалегідь відомим правилам (системи розпізнавання людської мови, автоматична класифікація деяких мікрооб'єктів (наприклад, хромосом), розпізнавання геологічних утворів на космічних зображеннях Землі, тощо).

Методи розпізнавання, що засновані на детермінованому статистичному й безперервно-груповому підходах, припускають поділ об'єктів у просторі ознак.

При синтаксичному (іноді називаному граматичним) підході шукають і використовують правила, яким підкоряється структура розглянутих образів.

Процедури класифікації й ідентифікації складаються із двох етапів і є досить простим і природнім наслідком із припущення, що конструкція, або структура, усіх підлягаючих розпізнаванню об'єктів підкоряється цілком певним, заздалегідь відомим правилам:

- на першому етапі розробляються ті правила, яким підкоряються конструкції; інакше кажучи, створюється свого роду граMATика в тому розумінні, у якому цей термін використовується в лінгвістиці;
- на другому етапі (коли граMATика визначена) визначається рішення про те, чи належить розглянутий об'єкт множині всіх об'єктів, породжуваних цією граMATикою.

*В основі структурно-лінгвістичного підходу до розпізнавання лежать основні поняття математичної лінгвістики, яка, у свою чергу, базується на самих загальних принципах побудови різних інструкцій у природних мовах.*

Синтаксичне розпізнавання має на увазі застосування до одномірних структур.

Термін «структурне розпізнавання» використовується для двовимірних (2 D-Даних) вихідних даних, зокрема зображень.

Для тривимірних структур (3 D-Структур) частіше застосовують терміни «аналіз сцен» і «розпізнавання сцен».

### ***Приклад структурного підходу до подання й опису***

Нехай алфавіт, тобто множину елементарних примітивних об'єктів (примітивів), з яких можуть полягати реальні об'єкти на зображенні, містить 4 елемента:  $\{a, b, c, d\}$ , де  $a$  — стрілка нагору  $\uparrow$ ,  $b$  — стрілка вправо  $\rightarrow$ ,  $c$  - стрілка вниз  $\downarrow$ ,  $d$  — стрілка вліво  $\leftarrow$ . Із цих примітивів можна скласти різні фігури, яким відповідають слова й пропозиції з використанням алфавіту. Слово  $b^m c^m d^m c^m$  означає, що мова йде про прямокутник, горизонтальні сторони якого містять  $m$  однакових відрізків, а вертикальні —  $n$  відрізків.

*До лінійних зображень можна застосувати різноманітні процедури кодування контурів, у тому числі так зване ланцюжкове кодування по методу Фримена, коли алфавіт утворений досить більшим числом елементів, подібних згаданам, але спрямованих під різними кутами.*

*При структурному підході під розпізнаванням розуміється встановлення відповідності між пропонованим об'єктом (точніше, його описом) і одним зі слів мови, породженого використовуваною граматиною.*

### Статистичний підхід

Виходить з припущення, що в одній точці простору ознак з відмінної від нуля апріорною ймовірністю можуть з'являтися реалізації, що ставляться до різних класів. Це може бути пов'язане з неминучими випадковими помилками, що присутні в детермінованих ознаках, а також з використанням ознак, що є принципово випадковими величинами, тобто імовірнісних властивостей.

**Методи статистичного розпізнавання образів** найбільше часто застосовують у завданнях дистанційного зондування:

- внаслідок випадкового характеру протікання природних явищ результати дистанційного зондування піддаються випадковим змінам, що маскують характерні відмінності між класами (статистичний аналіз дозволяє врахувати ці зміни, потенційно зменшити їхній негативний вплив на точність класифікації);
- досліджувані класи можуть перекриватися в просторі вимірів (деякі виміри одного класу можуть не відрізнятися від вимірів інших класів), що дозволяє здійснювати більш чітку класифікацію.

Методи статистичного розпізнавання образів використовують функції розподілу ймовірностей, пов'язані із класами образів. Найчастіше ці функції невідомі і повинні оцінюватися по множині навчальних образів.

Якщо форма функції розподілу ймовірностей вважається відомою (наприклад, нормальною) і по навчальних образах необхідно оцінити тільки окремі параметри, пов'язані із цими функціями (математичні сподівання, дисперсії, функції кореляції) - параметричний метод.

Параметричні методи звичайно легше реалізуються, але вимагають більшого обсягу апріорної інформації або фундаментальних припущень щодо природи образів.

Якщо форма функцій розподілу ймовірностей невідома заздалегідь – непараметричний метод.

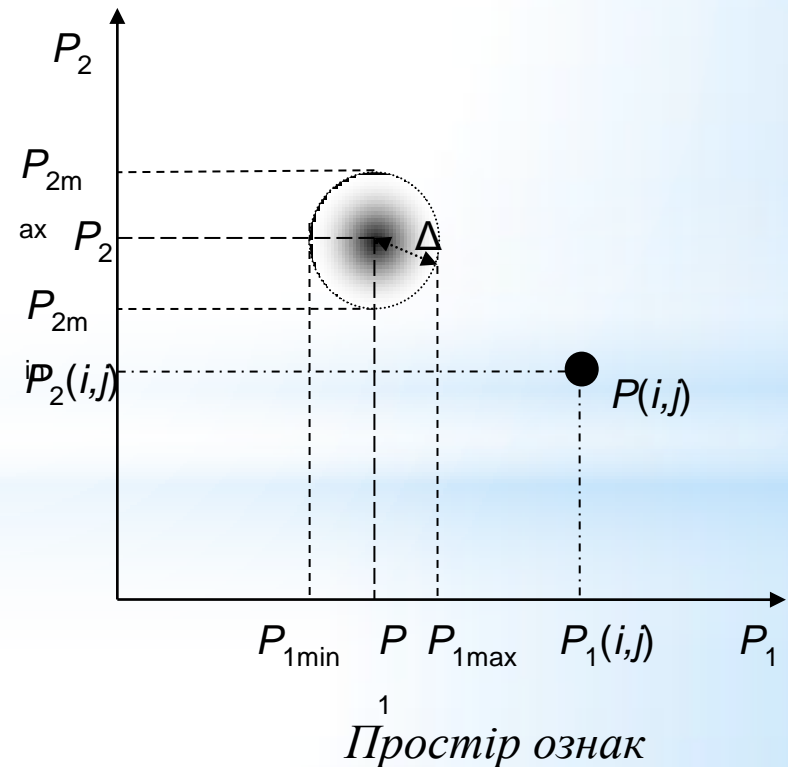
Непараметричні методи мають більші потенційні можливості для точної оцінки функцій розподілу ймовірностей і для точного розпізнавання, але ця перевага зазвичай вимагає складних систем, що розпізнають, великої кількості навчальних образів і головне, більших часових витрат.

## 2. Локалізація об'єктів на зображеннях

Поняття локалізація має на увазі процедуру спільного виявлення об'єктів і виміру їхніх координат на зображенні.

Один з основних способів виявлення об'єктів на зображеннях і визначення їхнього місця розташування складається в зіставленні з еталоном (еталонне зображення об'єкта порівнюється послідовно з усіма ділянками зображення, де може перебувати потрібний об'єкт).

Якщо розмір об'єкта на зображенні відомий, можна використовувати ковзне по зображенню вікно аналізу. У кожному положенні вікна з координатами  $(i,j)$  буде прийматися рішення про наявність або відсутність у цьому місці розшукуваного об'єкта:  $A(i,j)=1$  (у вікні об'єкт є) або  $A(i,j)=0$  (об'єкта у вікні немає). Якщо об'єкт виявлений ( $A(i,j)=1$ ), то по поточному положенню вікна аналізу  $(i,j)$  можна в першому наближенні визначити положення об'єкта.



Для рішення завдання локалізації об'єктів на зображенні необхідно:

*Перше.* Визначити чим саме повинен відрізнятися шуканий об'єкт від інших об'єктів і фону. Іншими словами, потрібно визначити простір характерних відмітних ознак об'єкта  $P_n$ . У найпростішому випадку необхідна єдина відмітна ознака  $P$ . Зрозуміло, що з безлічі ознак необхідно вибрати таку ознаку, що давала би найбільші розходження для потрібного об'єкта й фонових об'єктів. На рисунку представлений графічно двомірний простір ознак  $P_1, P_2$ . Можливі сполучення ознак для шуканого об'єкта показані у вигляді круглої області з радіусом  $\Delta$ .

*Друге.* Після визначення простору ознак зображення  $F(i,j)$  необхідно перетворити зі значень яскравості (кольору) у значення ознак  $P_n(i,j)$ . Це перетворення може бути зроблене відразу для всього зображення або здійснюватися під час руху вікна. Очевидно, що еталон повинен являти собою деяку множину припустимих значень ознак для заданого об'єкта.

*Третє.* Якщо ознака не одна, потрібно визначити спосіб порівняння розбіжностей у різних ознаках. Звичайно для цього визначають один показник, що залежить від розбіжностей у всіх ознаках обраного простору. Такий показник називають метрикою простору ознак. Способи визначення метрики можуть бути різними й залежать від розподілу ймовірностей розбіжностей ознак. Наприклад, часто використовують метрику у вигляді вагової суми квадратів різниць значень ознак у вікні аналізу  $P_n(i,j)$  і еталонних значень ознак  $P_n^e$ :

Приклад використання метрики у вигляді вагової суми квадратів різниць значень ознак у вікні аналізу  $P_n(i,j)$  і еталонних значень ознак  $P_n^e$ :

$$\delta^2 = \sum a_n (P_n^e - P_n(i,j))^2.$$

Вагові коефіцієнти  $a_n$  звичайно вибирають зворотно пропорційними значимості (чутливості) ознаки  $P_n$ . У випадку виявлення по одній ознаці як метрика виступає нормована різниця обмірюваного  $P(i,j)$  і еталонного  $P^e$  значень характерної ознаки. Повний збіг еталона з якою-небудь частиною зображення буває рідко через дію шумів і перекручувань, а також внаслідок відсутності точної апріорної інформації щодо ознак об'єкта, який потрібно виявити.

Якщо подібність між невідомим об'єктом і еталоном досить велика, то цей об'єкт позначається як відповідний еталонному. Звичайне значення метрики розбіжностей порівнюють із заданим порогом розбіжностей  $\Delta$ . Якщо значення метрики для даного положення вікна менше значення порога, то приймається рішення про наявність зображення об'єкта у вікні аналізу  $A(i,j)=1$ , і навпаки. Як правило, поріг вибирається постійним для всіх точок зображення.

Вибір порога є одним із ключових питань виявлення. При занадто високому рівні порога не будуть виявлені об'єкти з низьким контрастом у відмітних ознаках. Навпаки, занадто низький рівень порога з'явиться причиною того, що шум або фон буде хибно прийнятий за об'єкт. Для оптимального вибору порога використовують досягнення статистичної теорії перевірки гіпотез на основі апріорних припущень про характер і параметри імовірнісних розподілів значень ознак для ділянок зображень із об'єктом і без нього.

### 3. *Ознаки об'єктів на зображеннях*

Ознака зображення об'єкта - його найпростіша відмітна 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 яскравісні ознаки:

- розподіл по простору світлоти;
- величин яскравості;
- величин спектральної інтенсивності випромінювання;
- координат кольору;
- тощо.

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.

Інформацію про положення шуканих об'єктів несуть максимуми на результаті фільтрації. Їх можна виділити в масиві значень  $P(i,j)$  яким-небудь алгоритмом виявлення (ранговим голосуванням або нелінійним алгоритмом виділення максимумів). Звичайно, для прийняття рішення  $A(i,j)$  отриманий з виходу фільтра сигнал (значення ознаки  $P(i,j)$ ) порівнюється із заданим порогом  $\Delta$ .

Частотна характеристика такого фільтра визначається зображенням еталонного об'єкта й статистичними параметрами шуму й фону. Якщо задатися метою мінімізувати помилки хибного виявлення, оптимальним буде фільтр, що має частотну характеристику

$$H(w_x, w_y) = \frac{Z_0^*(w_x, w_y)}{\langle |Z_{\Phi}(w_x, w_y)|^2 \rangle}$$

$Z_0(w_x, w_y)$  – спектр еталонного зображення об'єкта, \* - позначка комплексного спряження. Знаменник - це усереднений по невідомим параметрах енергетичний спектр тієї частини зображення, що не містить об'єкт і виступає в якості "шуму" (енергетичний спектр фону).

## *Контурні ознаки*

Різкі зміни (розриви) яскравості або параметрів текстури - важливі найпростіші ознаки (визначають форму зображених об'єктів).

Локальні розриви значень яскравості - яскравісні перепади, або контури.

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 формують контурний препарат. Ідеальний детектор перепаду має вказувати на наявність перепаду в єдиній точці, розташованій в центрі переходу яскравості.

Для кожної області можна підрахувати якийсь набір найпростіших числових характеристик:

- площа;
- центр мас;
- периметр;
- компактність;
- орієнтацію головної осі інерції;
- подовження (ексцентриситет).



## *Ознаки плям*

Пляма - відносно маленька область зображення, яскравість якого значно відрізняється від яскравості її околиці.

Для виявлення плями спочатку здійснюється згладжування зображення фільтром низьких частот з маскою розміром  $N \times N$  елементів. Далі рівень кожного елемента усередненого зображення порівнюється із середнім значенням рівнів його чотирьох сусідів (праворуч, ліворуч, зверху й знизу), що 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  $N$ .

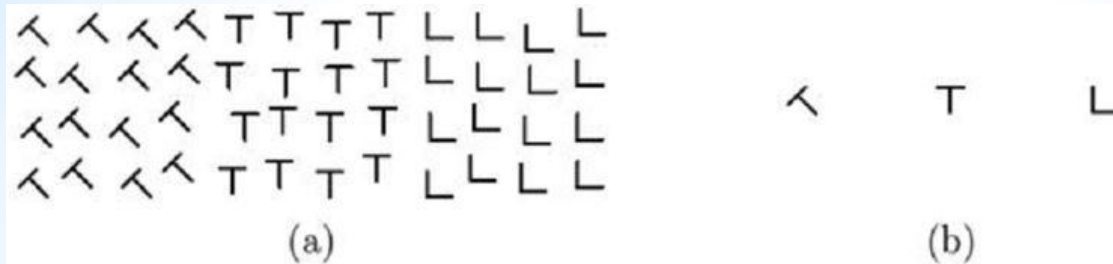
## *Ознаки ліній*

Можна припустити, що лінії і смуги є локальними фрагментами більших структур. Метод 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

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

*Маски фільтрів для виділення ознаки ліній*

## Текстурні плям



- (a) текстура - переважна орієнтація елементів, складових матеріалу (одне з визначень)
- (b) - типові приклади шаблонів (текстурних елементів) для досліджень психофізіологічного сприйняття зображень
- Людина явно враховує властивості текстури фрагмента при розпізнаванні зображень і виділення областей.

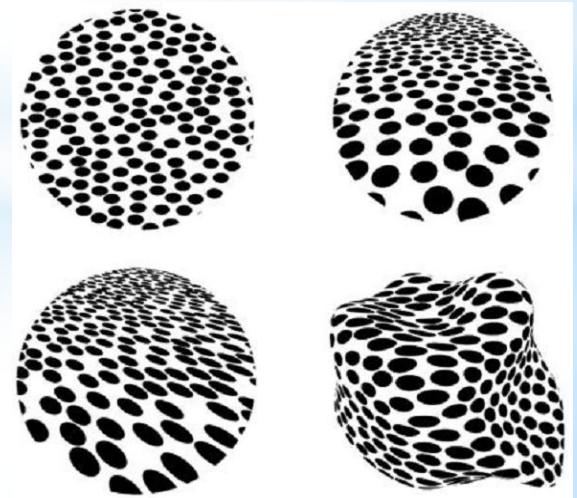
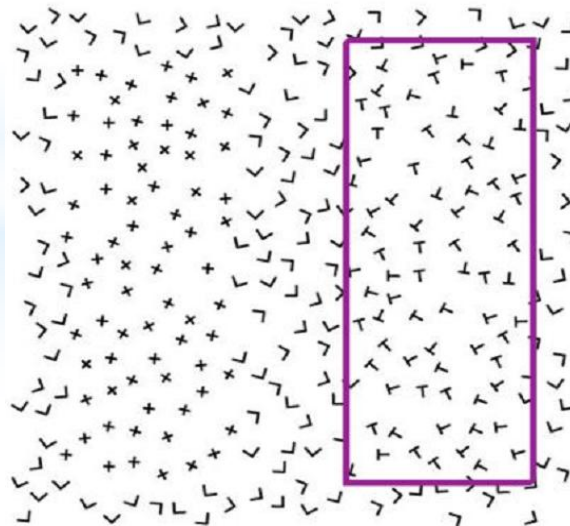
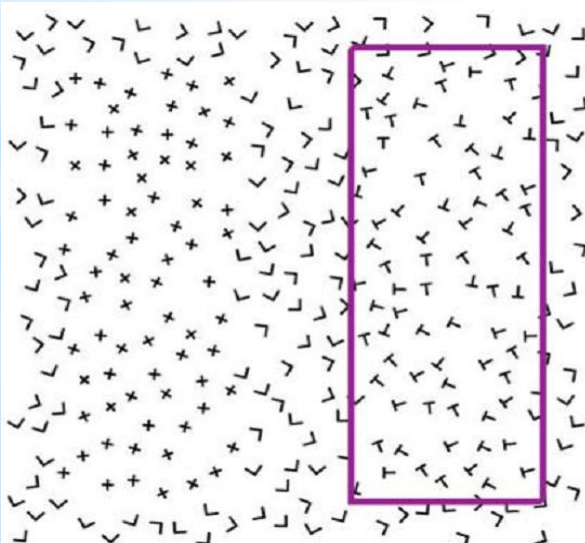
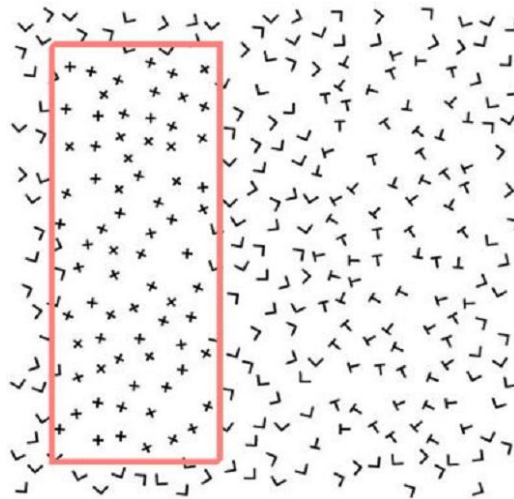
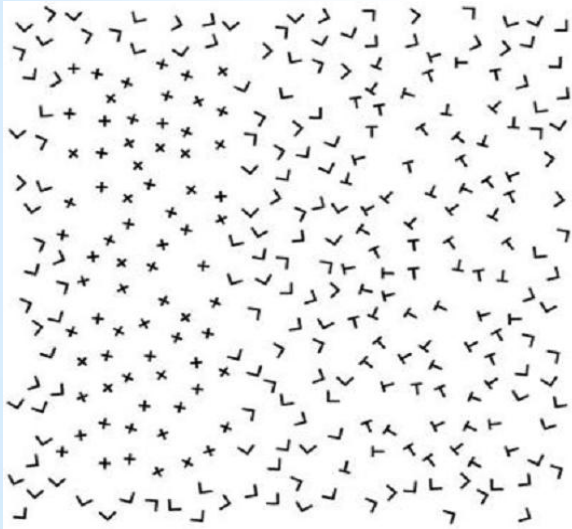
Текстура:

- штучна (структури із графічних знаків, розташованих на нейтральному фоні);
- природна (зображення природних сцен, що містять майже періодичні структури).

При автоматичній обробці знімків необхідно визначати границі текстурних областей і розмір зерен текстури усередині кожної області. Розмір зерна пов'язаний з періодом просторової повторюваності локальної структури. Великий період відповідає великій текстурі, а малий - дрібній. Розмір зерна недостатній для кількісного виміру текстури можна використовувати для оцінки напрямку зміни текстурних ознак.

***Текстура - властивість околиці точки зображення. Текстурні ознаки по своїй природі залежать від розміру вікна аналізу.***

За основну характеристику текстури можна використовувати просторову автокореляційну функцію. Розмір зерна текстури пропорційний ширині автокореляційної функції. Однієї з можливих мір ширини автокореляції є другий момент.



## 4. Препарування зображень

### 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.

### Препарування методом адаптивного квантування мод гістограми

Один з напрямків препарування зображень - застосування *вирішальних алгоритмів*. Зображення, що виходять з допомогою таких алгоритмів, являють собою *поля рішень* по деяких ознаках. Вирішальні алгоритми використовуються коли на зображенні необхідно виділити ділянки з визначеними властивостями і, потім, візуально проаналізувати форму цих ділянок, їхнє розташування, знайти розміри, площу і т.д.

Факт групування елементів зображення по ознакам виявляється в тім, що в гістограмі розподілу елементів по величині ознаки навколо деяких значень утворюються згущення, чи моди (їх ще називають кластерами) (рис.1). Виділивши ці моди, можна знайти поділяючі їхні границі, тобто граничні значення ознак по яким можна зробити поділ зображення на ділянки (рис.2).

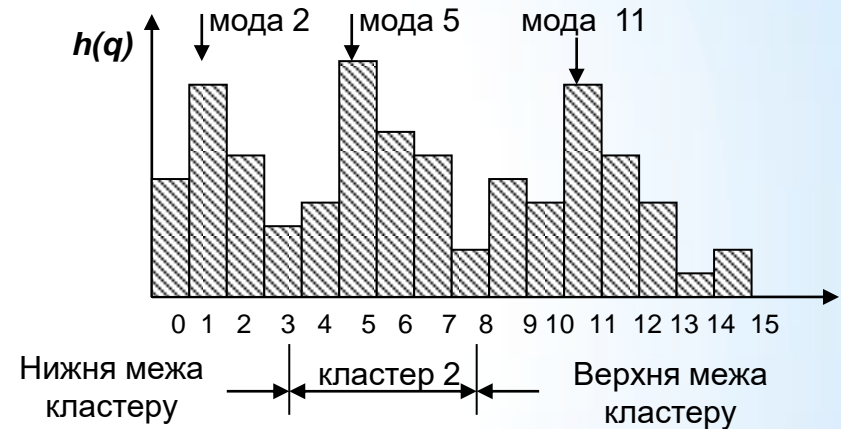


Рис. 1. Гістограма розподілу ознаки  $q$  на вихідному зображенні з окремими модами

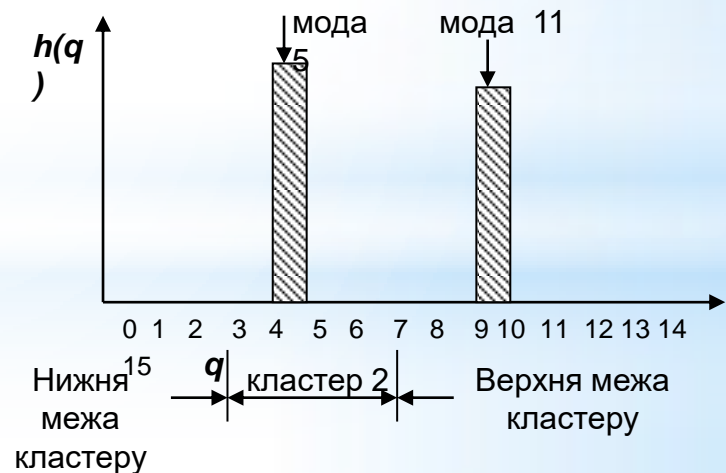


Рис.2. Гістограма розподілу ознаки  $q$  на зображенні обробленому з метою виділення ділянок з ознаками «( $q=5$ ) чи ( $q=11$ )»

Поділ мод у розподілі значень ознак можна назвати адаптивним квантуванням мод. Якість адаптивного квантування мод залежить від того, наскільки добре розділяються моди. Поділ мод залежить від того наскільки близько розташовані моди в гістограмі і наскільки вони розмиті.

Кількість мод у гістограмі звичайно залежить від ступеня детальності зображення. Варто обирати розмір аналізованого фрагмента так, щоб кількість мод у локальній гістограмі було не великим (зростає ризик зливання близько розташованих мод за рахунок їхнього розмивання).

Розмивання мод у гістограмі зв'язують із впливом операторів лінійного згладжування на зображення і наявністю на зображенні адитивних і мультиплікативних шумів. Створити адекватну реальним зображенням модель розмивання мод не завжди представляється можливим.

### ***Приєм створення моделі розмивання мод***

1. *Обмеження розмірів аналізованих зображень* (зображення розбивається на фрагменти, у межах яких моди розмиваються не значно; обробка ведеться по фрагментах).

2. *Поділ по типах розмивання мод* (обробка фрагментів з розмиванням мод за рахунок зашумлення ведеться іншими алгоритмами, чим фрагментів, де розмивання мод відбулося через згладжування лінійними операторами).

3. *Відбраковування мод по "заселеності"* (ігнорують моди з малим числом елементів, поріг вибирається з розумінь мінімально припустимої площі виділюваних ділянок зображення).

4. *Відбраковування дрібних деталей* (відфільтровувати шляхом приєднання до навколишнього моді елементи зображення, що належать до іншої моди, чим їхнє оточення, і утворюючі ділянки з площею, меншого порога, що задається).

## Препарування зображень по гістограмним і яскравістним ознакам

Найбільш важливою ознакою зображення є світлота. Світлота виражається через такі величини, як яскравість, спектральна інтенсивність випромінювання, координати кольору і т.д., що будуть називатися яскравістними ознаками. Вимір яскравістних ознак можна проводити або в окремих точках зображення, або в їхніх околицях (наприклад, після згладжування). Виділення об'єктів по яскравістним ознакам може бути здійснене за допомогою поелементних операцій спеціальних перетворень контрасту.

Форма гистограми розподілу частот дає багато відомостей про властивості зображення (вузька гистограма вказує на низький контраст, а бімодальна гистограма припускає наявність областей різної яскравості).

*Характеристики, що описують форму гистограм першого порядку*

Середнє 
$$\bar{f} = \sum_{f=0}^{N-1} f \cdot h(f).$$

Дисперсія 
$$\sigma_f^2 = \sum_{f=0}^{N-1} (f - \bar{f})^2 \cdot h(f).$$

Коефіцієнт асиметрії 
$$f_s = \frac{1}{\sigma_f^3} \sum_{f=0}^{N-1} (f - \bar{f})^3 \cdot h(f).$$

Коефіцієнт ексцесу 
$$f_k = \frac{1}{\sigma_f^4} \sum_{f=0}^{N-1} (f - \bar{f})^4 \cdot h(f) - 3.$$

Енергія 
$$E_f = \sum_{f=0}^{N-1} h^2(f).$$

Ентропія 
$$E_n = - \sum_{f=0}^{N-1} h(f) \cdot \log_2(h(f)).$$

В наведених формулах  $f$  – значення елементів зображень,  $N$  – кількість рівнів квантування сигналу зображення. Доданок - 3, введене у вираження для коефіцієнта ексцесу, нормалізує величину  $f_k$  так, що вона дорівнює нулю для квантованої гаусової щільності імовірності при прагнучому до нуля кроці квантування.

Гістограмні ознаки *другого порядку* засновані на визначенні спільного розподілу ймовірностей пари елементів зображення. Розглянемо два елементи  $F(i,j)$  і  $F(m,n)$  з координатами відповідно  $(i,j)$  і  $(m,n)$ . Тоді спільний розподіл значень яскравості має вид

$P(a,b) = P_r\{F(i,j)=a, F(m,n)=b\}$ , де  $a$  і  $b$  – квантовані значення яскравості.

Розподіл частот, що оцінює розподіл другого порядку, визначається формулою

$h(a,b) = N(a,b)/M$ , де  $M$  - повне число елементів у вікні, а  $N(a,b)$  - число випадків, коли  $F(i,j) = a$  і  $F(m,n) = b$ .

### Характеристики другого порядку

Автокореляція	$B_A = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} a \cdot b \cdot h(a,b).$
Коваріація	$B_C = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} (a - \bar{a}) \cdot (b - \bar{b}) \cdot h(a,b).$
Момент інерції	$B_I = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} (a - b)^2 \cdot h(a,b).$
Середня абсолютна різниця	$B_V = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1}  a - b  \cdot h(a,b).$
"Зворотна різниця"	$B_D = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} \frac{h(a,b)}{[1 + (a - b)^2]}.$
Енергія	$E_{ab} = \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} h^2(a,b)$
Ентропія	$E_n = - \sum_{a=0}^{N-1} \sum_{b=0}^{N-1} h(a,b) \cdot \log_2(h(a,b))$

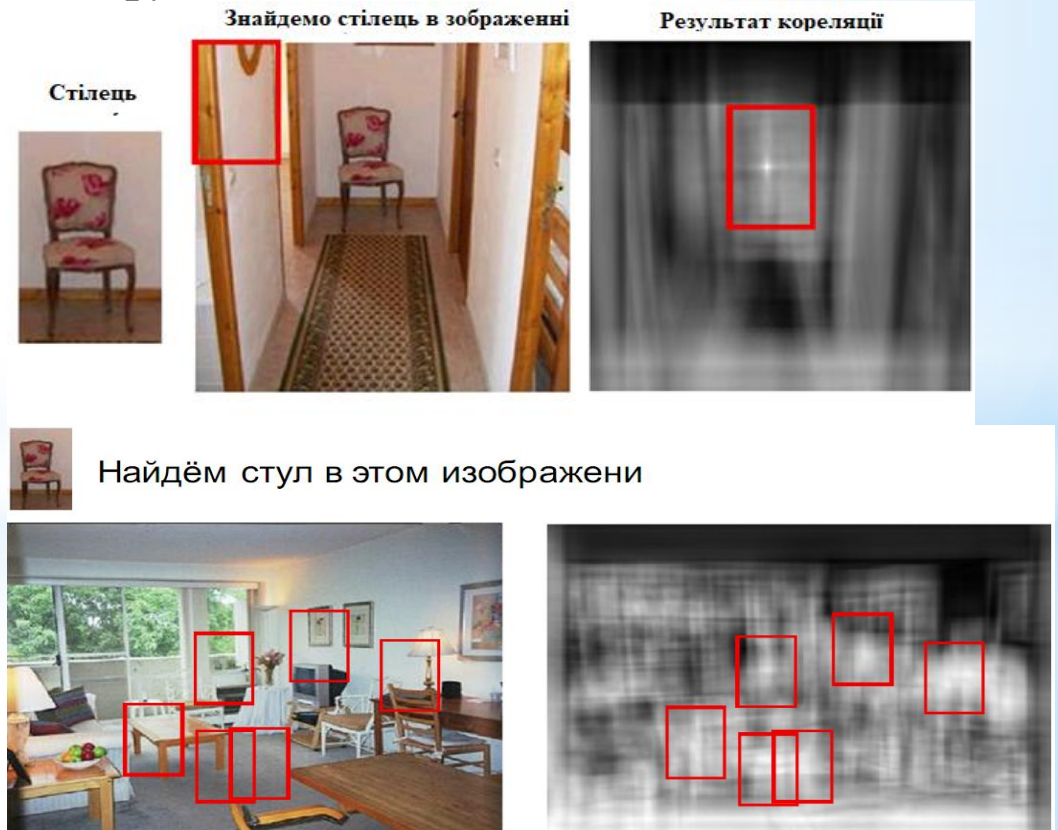
При препаруванні за будь-якою гістограмною ознакою потрібно спочатку перетворити початковий масив відліків зображення  $F(i,j)$  в масив значень відповідної ознаки  $O(i,j)$  з такими самими розмірами. Отриманий масив розподілу ознаки по простору  $O(i,j)$  можна препарувати розглянутими методами. Препарований масив значень ознак  $O_1(i,j)$  після відповідного нормування можна використати, як масив коефіцієнтів контрастування для кожного пікселя початкового зображення  $F(i,j)$ . Контрастоване таким чином зображення буде мати локальні контрасти в залежності від значень ознак, які вже мають чітке розділення по об'єктах.

### Приклад найпростішого класифікатора

Зіставлення шаблонів («Pattern matching»):

- фіксуємо зображення об'єкта («шаблон»);
- класифікатор - порівняння шаблону і поточного вікна (Фрагмента);
- підсумком кожного вікна буде карта відгуків (score).

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

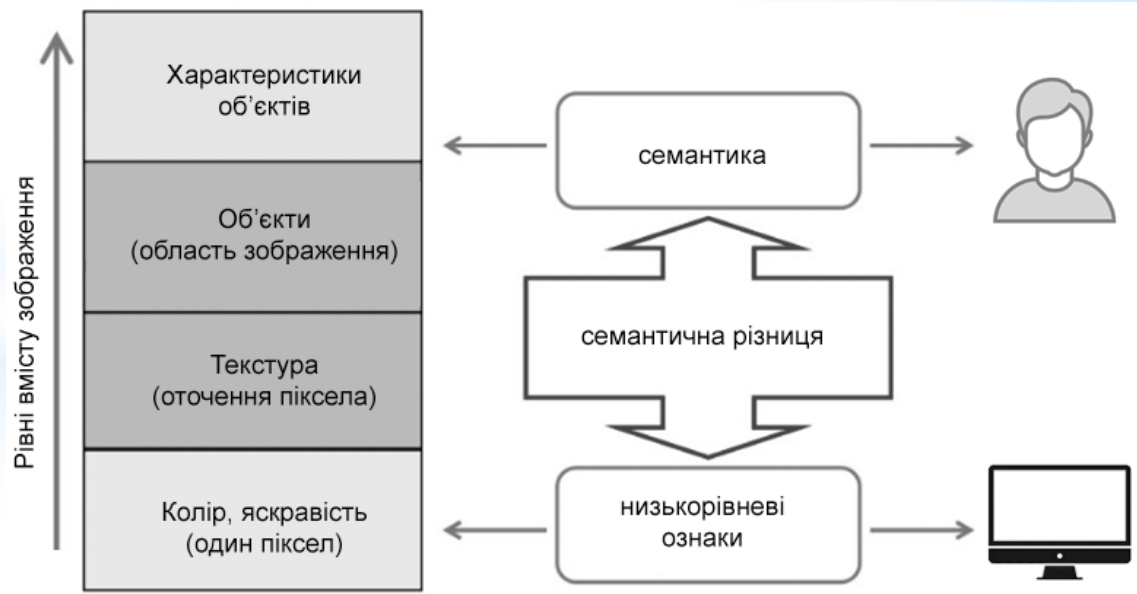


## 5. Комп'ютерний зір

Завдання комп'ютерного зору в загальній постановці (розуміння, що зображено на картинці) вважається ШІ-складною. Для розуміння вмісту зображення людина використовує весь свій життєвий досвід, освіту, а іноді і інтуїцію. На сьогодні, «сильного» штучного інтелекту, що за своїми можливостями співмірний до людини поки не створено, тому, розробники систем комп'ютерного зору вирішують тільки певні задачі.

Зорове сприйняття людини є досконалим надбанням природи. Людський мозок - це унікальна система розуміння і обробки тієї картинки, яку бачить людина. Вона схильна бачити те, що хоче бачити, і цьому не можна навчити комп'ютер. Мозок часто шукає знайомі образи, «добудовує» картинку і додає семантику.

*Особливості сприймання  
візуальної інформації*



Методи комп'ютерного зору використовуються для вирішення завдань, які умовно можна розділити на прості і складні.

Складні завдання відповідають на питання, який об'єкт зображено, до якого класу він належить. Для вирішення цих завдань найчастіше використовуються методи машинного навчання, а комп'ютерний зір називається **високорівневим**.

При вирішенні простих завдань здійснюються маніпуляції безпосередньо з пікселями, використовуються евристики, а методи машинного навчання, як правило, не застосовуються. Такі завдання називаються **низькорівневим** комп'ютерним зором

Завдання низькорівневого комп'ютерного зору часто використовуються як складова частина більш складних завдань розпізнавання. Наприклад, попередня обробка картини дозволяє алгоритмам машинного навчання краще зрозуміти, що на ній зображено.

Головний прогрес відбувається в сфері високорівневого комп'ютерного зору, що пов'язана з розумінням зображень. Тут вирішуються завдання, які легко формулюються. Наприклад, визначити об'єкти на фотографії, їх клас (людина, автомобіль, собака, дерево) та розташування. Або визначити, що роблять люди на відео, спокійна ситуація чи починається щось незвичайне, наприклад бійка. Тому, недостатньо просто встановити багато камер спостереження - потрібно, щоб потокове відео з них розумно оброблялося.