

Всеукраїнський конкурс студентських наукових робіт

Комп'ютерні науки

СТУДЕНТСЬКА НАУКОВА РОБОТА

На тему: «Методи класифікації зображень з використанням просторового
оброблення структурних описів»

Шифр: «Ієрархічні ознаки»

ЗМІСТ

ВСТУП	3
1. Застосування ієрархічних ознак для визначення релевантності зображень	5
1.1 Методи виділення ключових точок зображення.....	5
1.2 Статистичні моделі у просторі ознак	8
1.3 Інтегрована ієрархічна система ознак	12
2 Результати дослідження.....	17
2.1 Опис програмної моделі	17
2.2 Аналіз тестування комп'ютерної моделі	18
2.3 Аналіз результатів дослідження	23
ВИСНОВКИ.....	24
ПЕРЕЛІК ПОСИЛАНЬ	25

ВСТУП

Декілька напрямів науки, техніки і виробництва в значній мірі орієнтуються на розвиток систем, в яких інформація представлена у вигляді зображень. При обробленні таких даних виникає ряд складних наукових, технічних і технологічних проблем. Однією з найскладніших з них на сьогоднішній момент є аналіз та розпізнавання зображень. Мета розпізнавання може бути різною: як виділення окремих елементів на зображенні, так і класифікація зображення в цілому в межах деякої множини класів [1]. Класифікація зображень застосовується при автоматизації, ідентифікації видів, процесах управління (наприклад промисловими роботами), системах розпізнавання тексту, виявленні подій, у задачах взаємодії між людиною та комп'ютером, моделюванні та інших сферах. Тому на цей час комп'ютерний зір, а також пов'язані з ним технології програмної та апаратної обробки, є одним з перспективніших напрямків в світі комп'ютерних технологій. Постає питання створення нових ефективних методів чи алгоритмів, спрямованих на вирішення завдань класифікації, які працюють максимально точно, з мінімумом помилок та з достатньо високою швидкістю.

Автоматизація аналізу та обробки зображень, а також класифікація візуальних об'єктів є важливим завданням при створенні систем штучного інтелекту для сучасних високотехнологічних виробництв. У задачі розпізнавання у рамках об'ємних баз візуальних об'єктів у системах комп'ютерного зору важливими показниками є результативність і час оброблення. Основними ускладнюючими факторами при розпізнаванні зображень є: зовнішнє середовище у вигляді фону і перешкод, а також геометричні перетворення, які приводять до зміни об'єктів та відповідно заважають алгоритмам дати правильні відповіді. Адже такий вплив може призвести до викривлення фрагментів аналізованого об'єкта або сформувати помилкові об'єкти з близькими характеристиками [2].

Процес реалізації дієвих класифікаційних рішень у сучасних системах комп'ютерного зору також потребує надійного аналізу сигналів із багатовимірною природою оброблюваних даних. При застосуванні структурних методів розпізнавання описи об'єктів подаються у вигляді наборів дескрипторів ключових точок (КТ) як скінченної множини числових векторів достатньо високої розмірності [3-5]. Так, бінарні описи, отримані детекторами ORB, BRISK [3; 6], містять 256 та 512 компонентів. У такому випадку перехід до подання даних у вигляді системи їх фрагментів меншої розмірності сприяє суттєвому спрощенню їх прикладного застосування [5]. Основним інструментом при цьому виступає апарат просторово-статистичного аналізу даних, який в аспекті розпізнавання базується на узагальненні даних опису об'єкта шляхом обчислення нових ознак для сукупності фрагментів [2;7].

Викликає також інтерес дослідження властивостей ієрархічної системи ознак, що може бути побудована на підставі обчислених розподілів найпростішого для оброблення виду фрагментів, що містять по 1 біту. Аналогічно пірамідальній структурі даних на базі таких ознак можна спробувати побудувати атрибути більш високого рівня, застосування яких в разі прискорює процес оброблення та прийняття рішень.

Метою роботи є розвинення та дослідження методу структурного розпізнавання зображень на підставі впровадження просторово-статистичного оброблення для кортежу значень блоків опису у складі множини дескрипторів КТ. Важливим представляється опрацювання моделей для обчислення релевантності модифікацій структурного опису, ефективних за швидкодією оброблення даних, аналіз властивостей цих моделей та експериментальне оцінювання результативності запропонованих підходів за наслідками аналізу прикладних баз зображень.

1. Застосування ієрархічних ознак для визначення релевантності зображень

1.1 Методи виділення ключових точок зображення

Складну проблему представляє собою потреба витягнути осмислену та структуровану інформацію із зображення. Вирішенням можуть стати технології, засновані на аналізі кінцевого числа характерних ознак (ХО) зображення в окремих точках [2-7]. Такий підхід має багато переваг: спрощення процедури формування ознак, скорочення простору ознак, універсальна можливість структурного аналізу об'єктів, стійкість до впливу фонових спотворень та помилковості об'єктів.

Ключова точка (КТ), або точкова особливість зображення – це точка, розміщення якої виділяється на фоні будь-якої іншої точки. Формування КТ – це виділення фрагменту з навколишнього фону. Сам процес побудови множини дескрипторів КТ зображення можна поділити на два етапи. Перший етап – це виявлення ключових точок зображення (застосування детектору). Детектором називають метод, що виділяє точки або локальні області, які мають ознаки, які виділяють їх серед множини сусідніх точок.

Другий етап – безпосередньо формування дескриптора. Дескриптором називається функція, що описує КТ за її околицею. Можна вважати його альтернативно поданою КТ, для котрого, до того ж, можна ввести міру схожості. Так, наприклад для опису точки c з координатами $c = (x, y)$ використовується просторова околиця $O(c)$, $c \in O(c)$, де $O(c)$ – результат застосування оператора з побудови множини точок, координати яких знаходяться в безпосередній близькості, тобто $O(c) = [\{c_i = (x_i, y_i)\}, \forall c_i : \rho(c_i, c) \leq \nu]$, де параметр задає розмір околиці, а ρ – деяку відстань між точками на площині [4].

Точкова особливість на зображенні $B(x, y)$, $(x, y) \in D$, $D \subset R \times R$ – це точка c^* , в околиці $O(c^*)$, $O(c^*) \subset D$ якої функція $B(x, y)$ за деякими властивостями відрізняється від множини своїх значень з більшою за розміром околиці $O_1(c^*)$, $O(c^*) \subset O_1(c^*)$ [4,10]. ХО формується з множини точок з околиці $O(c^*)$.

Дослідниками розроблена серія алгоритмів пошуку КТ та побудови матриці дескрипторів, у кожного з яких є свої переваги та недоліки. Усі алгоритми можна поділити на дві основні групи за способом представлення дескриптору:

- дескриптори з рухомою точкою (найбільш популярні алгоритми: SURF, SIFT, CSLBP та ін.);
- бінарні дескриптори (найбільш популярні алгоритми: BRIEF, BRISK[3], AKAZE, ORB[6] та ін.).

У нашій роботі було надано перевагу бінарним дескрипторам. Бітовий склад дескрипторів КТ у просторі B^n бінарних векторів (n – ступінь двійки) дає можливість запровадити оброблення дескриптора як кортежу послідовних елементів, діапазон значень яких відомий. Це дає змогу розглядати наявний структурний опис (множина бінарних векторів) у синтезованому просторі даних та здійснювати статистичний аналіз розпізнаваних зображень з урахуванням властивостей складу еталонних описів. Бінарне подання значно прискорює процес порівняння дескрипторів за рахунок можливості застосування двійкових операцій та відповідно спрощує апаратну реалізацію системи розпізнавання. Крім того, бінарна арифметика дає можливість застосувати ефективний апарат оброблення бінарних даних та синтезувати нові підходи для визначення подібності дескрипторів при побудові правил класифікації.

У роботі використано метод обчислення бінарних дескрипторів, який називається ORB. Метод використовує алгоритм FAST для знаходження ключових точок, який приймає як параметр порогове значення інтенсивності

між центральним пікселем і описаним навколо нього колом. Він ефективний, зручно використовується у системах реального часу, знаходить «розумні» КТ, хоча і має бути доповненим кутовим фільтром Харріса для відхилення країв та отримання розумної оцінки. В ORB радіус кола, яке аналізується навколо кожного пікселю, приймається рівним 9, оскільки саме ця відстань виявилася найбільш ефективною з точки зору продуктивності.

За замовчуванням КТ у цьому алгоритмі не більше 500. Якщо їх більше, то до них застосовується детектор кутів Харріса для виключення найменш значущих: щоб отримати N ключових точок, спочатку використовується низький поріг для того, щоб отримати більше N точок, потім вони упорядковуються за допомогою метрики Харріса і вибираються перші N точок.

Для побудови значень дескриптора для отриманих точок використовується модифікація BRIEF (спрямований дескриптор BRIEF). Дескриптор представляється у вигляді вектора довжиною 256, що складається з результатів бінарних тестів навколо особливої точки. В околиці порівнюються середні значення яскравостей між x та y , де x, y - області 5×5 пікселів:

$$\tau(I; x, y) := \begin{cases} 1: I_x < I_y \\ 0: I_x \geq I_y \end{cases}, \quad (1.1)$$

де I – середня яскравість обраної області.

Всі $n = 256$ наборів x_i та y_i формують матрицю S розмірністю $2 \times n$. Далі S за допомогою матриці повороту R_θ орієнтується відповідно до кута θ : $S_\theta = R_\theta S$. А сам вектор дескриптора записується як:

$$g_n(I, \theta) := f_n(I) | (x_i, y_i) \in S_\theta, \quad (1.2)$$

$$\text{де } f_n(I) := \sum_{i=1}^n 2^{i-1} \tau(I; x_i, y_i).$$

Детектор ORB формує приблизно в два рази більше дескрипторів, ніж SURF, і при цьому числом дескрипторів можна управляти. Але за рахунок більш коротких дескрипторів в 32 елемента час розпізнавання для дескрипторів ORB на порядок менше, ніж для SURF.

1.2 Статистичні моделі у просторі ознак

У структурних методах розпізнавання зображень, що набули поширення в системах комп'ютерного зору, опис візуального об'єкта представляють множиною векторів, закріплених в координатах ключових точок зображення і інваріантних до геометричних перетворень об'єктів в полі зору [12]. Ця множина однотипних ознак відображає конструкцію образу об'єкта у вигляді списку точок із простору векторів.

Аналіз зображень вимагає суттєвих обчислювальних ресурсів. Головною перевагою уявлення сигнатури в вигляді дескриптора є обчислювальна ефективність. Однак, такий підхід не володіє необхідною для подання складної семантики зображення інформацією. Необхідна розробка більш складних структурно-ієрархічних підходів, заснованих на аналізі сукупності частин об'єктів. Тобто можна використовувати конкатенацію дескрипторів окремих областей. Це робить більш здійсненою можливість поділу зображення на окремі об'єкти, зберігаючи інформацію про просторовий розподіл останніх. Тому основне завдання структурного розпізнавання пов'язане з удосконаленням інформаційних технологій класифікації у просторі описів як множин дескрипторів ключових точок зображень.

Статистичне подання є одним із найбільш популярних інструментів у сучасному інтелектуальному аналізі даних. Воно набуло істотного поширення в багатьох завданнях, в тому числі і при прийнятті класифікаційних рішень у системах комп'ютерного зору [12-15]. Ключовою перевагою є використання для класифікації певної узагальненої інформації про властивості класів

розпізнаваних об'єктів, що дає змогу результативніше врахувати особливості об'єктів у просторі ознак.

Один із підходів заснований на використанні таких характеристик, як математичне очікування, дисперсія, оцінки медіани. Але більш інформаційним є використання безпосередньо самих значень розподілів, як це робиться в методах інтелектуального аналізу. Як правило, дані у комп'ютерному зорі мають багатовимірний ймовірнісний розподіл. Але іноді їх можна подати у вигляді множини розподілів одновимірних величин, що значно спрощує класифікацію та знижує обчислювальні затрати [16]. Це дає змогу більш чуттєво врахувати відмінності та особливості значень даних, що відображають властивості розпізнаваних класів.

Одним з перспективних шляхів побудови нових підходів і удосконалення традиційних способів зіставлення описів візуальних даних є локально-просторове уявлення зображення [13,17]. В результаті аналізоване зображення набуває вигляду множини локально заданих функцій або множини локальних ознак, областю визначення яких є підобласті поля зору. Таке уявлення має більшу спільність у порівнянні з цілісним описом, тому що дозволяє аналізувати складові частини об'єкта, зокрема, при їх спотворенні під впливом просторових перешкод, та приводить до результативних рішень і забезпечує потрібну завадостійкість.

Перспективи структурного розпізнавання пов'язані з удосконаленням інформаційних технологій класифікації в просторі описів у вигляді множин дескрипторів ключових точок зображень, а також необхідністю оцінювання дієвості розпізнавання на прикладних зразках.

Розглядаємо простір B^n бінарних векторів. Опис Z розпізнаваного візуального об'єкту визначимо у вигляді скінченної множини $Z = \{z_v\}_{v=1}^s$, $z_v \in B^n$, $Z \subset B^n$ із s бінарних дескрипторів КТ сформованих детектором ORB.

Для фіксованого числа n опис об'єкту можна подати бінарною матрицею $D = \{\{d_{i,j}\}_{i=1}^s\}_{j=1}^n$, сформовану довільною послідовністю отриманих детектором s дескрипторів КТ.

Для розв'язання задачі розпізнавання та класифікації представимо множину дескрипторів у вигляді матриці $s \times n$ [18-20]:

$$D = \begin{pmatrix} d_{11} & \dots & d_{1n} \\ \vdots & \ddots & \vdots \\ d_{s1} & \dots & d_{sn} \end{pmatrix} \quad (1.3)$$

де s – обрана кількість дескрипторів, $s = 1, 2, \dots, 500$, $d_{sn} \in [0; 1]$.

Для спрощення аналізу вважаємо значення s однаковим для всіх еталонів, цього нескладно досягти відбором необхідної кількості дескрипторів.

Кожний рядок матриці D , що є дескриптором КТ, подамо у виді m непересічних фрагментів (блоків), що послідовно слідуєть один за одним. Вони повністю складають початковий розмір n (рис.1.1), де кожний елемент приймає значення певного діапазону. Так, якщо елемент представлений одним бітом, то маємо два значення (0, 1), якщо байтом – множину значень $0, \dots, 255$.

елемент 1	елемент 2	...	елемент m
-----------	-----------	-----	-----------

Рисунок 1.1 – Ланцюжкова структура дескриптора даних

Отримаємо новий простір даних, перевагою якого є можливість аналізувати внутрішню структуру та статистичні властивості опису, на основі чого його можна буде відносити до одного із еталонних класів. Застосуємо мережу блоків як основу для прийняття рішення про клас об'єкту. Блоки дескриптора – суть нова система ознак. Такий спосіб формування ознак надає додаткову гнучкість при практичній реалізації [11].

Після розбиття $\forall z_v \in Z$ на m непересічних блоків для кожного дескриптора z_v отримуємо фіксовану структуру $z_v = z_v^1 \& z_v^2 \& \dots \& z_v^m$ «зчеплених» елементів, де z_v^k – k -й блок дескриптора z_v . Кожний із дескрипторів тепер складається із m однотипних бінарних елементів, значення яких у послідовності можна записати кортежем цілих чисел. У результаті опис Z набуває виду матриці із s рядків по m елементів (число блоків) у рядку.

Зважаючи на впроваджену рівноцінність діапазонів для значень кожного із фрагментів та відображення сукупністю дескрипторів властивостей аналізованого об'єкту, побудуємо розподіли даних на підставі аналізу матриці D [5]. Для цього задамо відображення $\Omega: Z \rightarrow Q$, $Z \subset B^n$ із множини бінарних векторів у множину Q цілочисельних векторів потужності $w < n$, що містять статистичні розподіли значень їх фрагментів. Відображення Ω призначене забезпечити ідентифікацію та розрізнення об'єктів у просторі значень розподілів їх фрагментів.

Визначимо розподіл (рейтинг зустрічальності) $q \in Q$ як вектор $q = \{q_1, \dots, q_w\}$ цілих чисел, де q_i – це число значень для k -го фрагмента, що дорівнюють i , $i = \overline{1, w}$, серед відповідних фрагментів усієї множини Z

$$q_i = \text{card}\{z_v^k \in z_v, z_v \in Z \mid z_v^k = i\}. \quad (1.4)$$

Тоді виконується умова $\sum_i q_i = s$, так як сума значень ланок для розподілу q дорівнює загальному обсягу s дескрипторів опису Z . Зауважимо, що w – це число ланок розподілу, які належать діапазону $0, \dots, w-1$.

Для кожного із m фрагментів побудуємо розподіл $q = \{q_1, \dots, q_w\}$, величина w визначена діапазоном значень даних для фрагмента, що у бінарному поданні визначається його розміром. Наприклад, для дескриптора ORB при розбитті на байти для $n = 256$ маємо $m = 32$, $w = 256$. Загалом

множина Z під дією відображення Ω описується матрицею $Q = \{\{q_{i,v}\}_{i=1}^w\}_{v=1}^m$, що містить розподіли системи фрагментів.

Матриця Q відтворює статистичні властивості опису Z у вигляді кортежу розподілів значень фрагментів його складових. Статистичні розподіли даних узагальнюють знання у відповідності до фундаментальної концепції машинного навчання [2].

1.3 Інтегрована ієрархічна система ознак

На основі одно-бітових розподілів побудуємо інтегровану в межах окремих фрагментів систему ознак. Для матриці D опису обчислимо значення суми вздовж стовпців (1-бітовий фрагмент), отримаємо вектор $t = (t_1, \dots, t_j, \dots, t_n)$, де $t_j = \sum_{i=1}^s d_{i,j}$, $j = \overline{1, n}$ (фактично t_j – це сума одиниць у стовпці). Це і будуть ознаки найнижчого рівня, що у повній мірі зберігають інваріантні властивості множини дескрипторів Z [19].

Тепер на базі ознак t_j обчислимо ознаки більш високого рівня u_k для блоків у вигляді сукупностей стовпців

$$u_k = \sum_{j=k}^{k+b-1} t_j, \quad (1.5)$$

де $b = n / m$ – розмір, а $k = 1, b+1, 2b+1, \dots, n-b+1$ – номер фрагмента.

Ознаки (1.5) реалізують крос-кореляційне оброблення матриці D з прямокутною маскою розміром $b \times s$ [7]. У результаті обчислення (1.5) отримуємо цілочисельний вектор u_k розмірністю m . Параметр m є характеристикою системи фрагментів, він зменшується від n до 1 зі збільшенням розміру фрагмента від 1 до n .

Значення вектора $u = (u_1, \dots, u_k, \dots, u_m)$ можуть бути використані як самостійні структурні ознаки статистичного виду, вони обчислюються на підставі побітового аналізу стовпців матриці D . Опрацювання (1.5) залежить від параметра b та реалізує просторове оброблення (інтегрування даних) на множині дескрипторів. Зважаючи на таку просту модель обчислення функцій (1.5), всі вони для довільного розміру фрагмента визначаються нескладно (логічно або додаванням цілих чисел), базуючись на 1-бітових значеннях статистичного розподілу даних.

На основі подання (1.5) може бути застосовано також ієрархічний спосіб розпізнавання, що використовує для зіставлення з еталонами систему ознак u_k з різним ступенем інтегрування даних, що визначається розміром блока.

Діапазон цілих значень для ознак u_k безпосередньо можна визначити за розміром фрагмента як $u_k \in \{0, \dots, sb\}$. Загалом вектор u є інтегрованою характеристикою опису, його можна окремо нормувати за кількістю дескрипторів або розміром фрагменту.

Обсяг значень для ознак (1.5) змінюється від одного числа до вектора розміру n (розмір дескриптора) і зменшується зі збільшенням розміру фрагмента. При $m=1$ маємо $b=n$, тобто кожний дескриптор містить один фрагмент, а величина u_1 визначається одним значенням і дорівнює сумі бітів усього опису. При $m=n$ маємо іншу граничну ситуацію – n 1-бітових фрагментів, а u_k обчислюється як сума стовпця матриці D .

Модель (1.5) реалізує процедуру зменшення інформаційної надмірності просторового сигналу за рахунок допустимого зниження (з точки зору якості розрізнення) ступеня розподільної здатності системи ознак подання опису.

Оброблення виду (1.5) є різновидом просторового аналізу значень для множини дескрипторів вздовж координат їх компонентів. Іншими варіантами такого оброблення є фільтрація чи розклад вектору дескриптора за системою ортогональних функцій [7].

Важливим завданням є також оцінювання еквівалентності об'єктів по їх структурним описами. Зауважимо, що наявність геометричних перетворень об'єктів значно ускладнює аналіз відповідностей КТ. Доводиться вирішувати проблему інваріантності, яка актуальна як на локальному рівні формування КТ, так і на етапі зіставлення описів.

Таким чином, встановлення відповідності описів об'єктів, представлених у вигляді множин ХО, в загальному плані зводиться до оцінки ступеня еквівалентності множин. У спрощеному випадку завдання можна вирішити на основі еквівалентності векторів.

З іншого боку, інформації від розподілів окремих бітів чи їх пар може не вистачити для забезпечення необхідного рівня розрізненості описів, і тоді прийдеться йти на збільшення обсягу обчислень. Одним із шляхів спрощення аналізу та оброблення ланцюжкових даних є використання значень окремих параметрів побудованих розподілів.

Визначення релевантності для цього способу виконується шляхом обчислення відстані між векторами або з врахуванням по-елементного аналізу. Аналіз виду (1.5) відповідає принципу максимуму апостеріорної ймовірності. Для суттєво різних описів такий аналіз приносить необхідний результат [19].

Подання та аналіз даних у вигляді множини фрагментів отримали значне поширення у теорії комп'ютерного зору. Воно забезпечує можливість прийняття рішень за окремими елементами візуальних об'єктів, але визначення порогів на кількість та подібність фрагментів, виконується користувачем. Зрозуміло, що встановлення оптимального розміру елемента, порогу для визначення еквівалентних за описом елементів, порогу для встановлення значущості опису елемента повинно бути виконано, виходячи із заданої бази зображень, в межах якої виконується розпізнавання.

Основною моделлю зіставлення ХО в більшості систем структурного розпізнавання є порівняння числових векторів в лінійному нормованому просторі. Так як метрики в цілому забезпечують більш прийнятний діапазон властивостей, ніж функції подібності, то основне застосування в роботі при

зіставленні на локальному рівні отримали метричні характеристики і відповідні метричні підходи.

Метрики – важливий інструмент вирішення багатьох завдань розпізнавання образів та інтелектуального аналізу даних. Значення метрики пов'язане з ймовірними характеристиками віднесення елемента до класу [8]. Обчислення метрики $\rho_{\Lambda}(\Lambda^1, \Lambda^2)$ для множин Λ^1, Λ^2 реалізується двоетапною процедурою: спочатку обчислюються відстані $\rho(\lambda^1, \lambda^2)$, $\lambda^1 \in \Lambda^1, \lambda^2 \in \Lambda^2$ між елементами, а потім на основі множини $\rho(\lambda^1, \lambda^2)$ визначається значення метрики для множин у вигляді функції $\rho_{\Lambda}(\Lambda^1, \Lambda^2) = f(\{\rho(\lambda^1, \lambda^2)\})$ [7, 20].

У роботі було прийнято рішення використовувати Манхеттенську відстань. Манхеттенська відстань обчислюється за формулою: релевантність r описів a та b визначимо на підставі зіставлення розподілів за допомогою відстані між матрицями $u(a), u(b)$:

$$r[u(a), u(b)] = \sum_{i=1}^m |u_i(a) - u_i(b)|, \quad (1.6)$$

Як правило, прийняття рішення ґрунтується на нормуванні значення (1.6):

$$r^* = \frac{r}{r_1}, \quad (1.7)$$

де r_1 – максимальна відстань для обраної кількості дескрипторів.

Для розпізнавання візуальних об'єктів за їх структурним описом у вигляді множини дескрипторів КТ зображення нами розроблено алгоритм та відповідний програмний засіб для класифікації, що задає відношення еквівалентності для фіксованої множини дескрипторів розпізнаваного об'єкту, спираючись на задану базу зображень, в рамках якої відбувається вирішення

задачі класифікації. Загальну схему роботи розробленого методу та алгоритму розпізнавання наведено на рисунку 1.2.



Рисунок 1.2 – Загальна схема роботи алгоритму класифікації

Як бачимо, на вхід поступає розпізнаване зображення, для якого формується множина дескрипторів КТ, далі для цієї множини формуються розподіли даних (1.4) та на їх основі ієрархічні ознаки (1.5), які потім порівнюються за формулами (1.6), (1.7) із відповідними еталонними ознаками. За результатом зіставлення ознак визначається клас вхідного зображення.

2 Результати дослідження

2.1 Опис програмної моделі

Модель класифікації зображень за допомогою пошуку відстані між ознаками вхідного зображення та ознаками бази еталонів з використанням ієрархічних структурних описів реалізована мовою Java у середовищі IntelliJ, під операційною системою Windows.

Обчислення, аналіз і обробка дескрипторів КТ реалізовані у вигляді програмного засобу на основі модуля features2d для крос-платформної бібліотеки OpenCV. У проєкті цей модуль забезпечує пошук та зберігання даних про КТ та дескрипторів зображення, а також виведення зображень з позначеними координатами КТ.

Для роботи програми сформовано базу еталонних зображень, з якими будуть порівнюватись вхідні зображення. За базу еталонів застосований набір з зображень гербів університетів Оксфорду, Гарварду, Пекіну та Массачусетса. Розмір зображень складає 600x600. Приклад можна побачити на рисунку 2.1. Для моделювання застосовано дескриптори ORB.

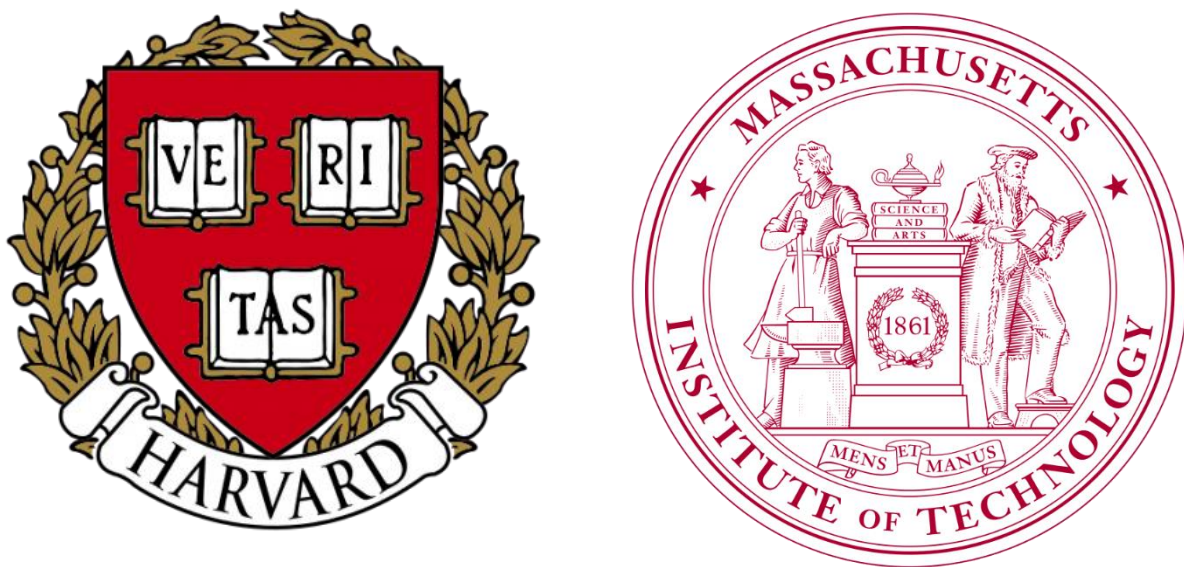


Рисунок 2.1 - Приклади зображень

З метою формування якісних структурних описів, стійких до фонових і випадковим перешкод, на основі використання стандартних функцій бібліотеки OpenCV проведена спеціальна обробка аналізованих зображень. В ході такої обробки зображення піддавалося фільтрації: зображення переводилося в градації відтінків сірого.

Дескриптори мають вигляд бінарного вектора розміром 256. Дані зберігаються в матриці, де кількість рядків дорівнює числу виявлених дескрипторів, а число стовпців дорівнює 32 (256 біт дескриптора трансформуються в 32 byte).

Після отримання зазначених дескрипторів переходимо до нового простору ознак, побудувавши однобітові розподіли. Далі будуємо ієрархічні ознаки для емпіричного опису зображення.

Також було програмно реалізовано додаткові функції: такі як побудова гістограм дескрипторів, виведення вхідного зображення для користувача, демонстрація зображення з нанесеними координатами ключових точок та найближчий у просторі даних еталон. Ці функції дають можливість оцінити результативність функціонування програмної моделі та методу, для якого вона розроблена.

2.2 Аналіз тестування комп'ютерної моделі

На рис. 2.2 наведено приклад знаходження 500 КТ для зображення Массачусетського гербу. На рис. 2.3 наведено приклад сформованого 1-бітового розподілу (256 ланок дескриптора ORB) для кількості нулів у 100 дескрипторах зображення Гарвардського герба.



Рисунок 2.2 – Синтезовані 500 КТ на зображенні герба Массачусетського університету

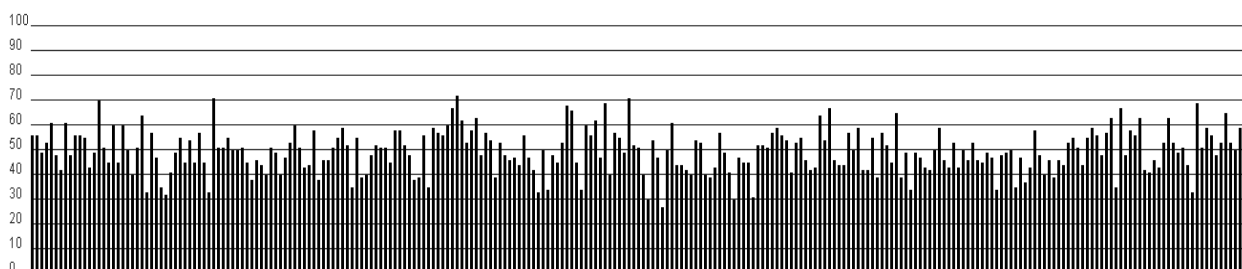


Рис. 2.3 – Приклад розподілу для 100 дескрипторів

У табл. 2.1 наведено значення манхеттенської відстані r для 1-бітових розподілів 100 дескрипторів чотирьох гербів.

Таблиця 2.1 – Відстані між 1-бітовими розподілами для 100 дескрипторів

	Оксфорд	Гарвард	Масачусет	Пекін
Оксфорд	0	1922	3782	5232
Гарвард	1922	0	2868	4182
Масачусет	3782	2868	0	2962
Пекін	5232	4182	2962	0

На рис. 2.4 показано схему побудови та значення ієрархічних ознак для емпіричного опису зображення рис.2.1 у складі 100 дескрипторів, де значення у нижніх колах відповідають 8-бітовому розподілу нулів, а верхнє значення дорівнює загальному числу нулів опису. Аналогічну структуру можна побудувати для довільного опису [19].

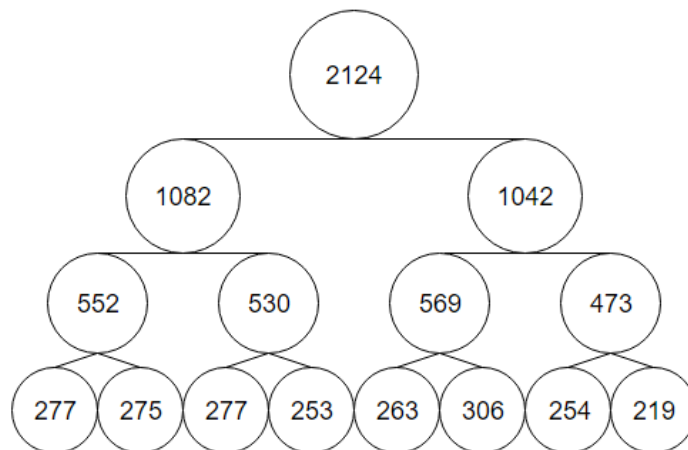


Рис. 2.4 – Значення ієрархічних даних на прикладі герба Гарварду

Таблиця 2.2 містить значення відстані (1.7) між розподілами зображень гербів Масачусетського та Оксфордського університетів для 100, 300, 500 дескрипторів та різним числом бітів у фрагментах розподілу. Також у таблиці 2.2 наведені відстані для скорочених розподілів цих гербів, де вибрано кожен другий та четвертий біти із однобітового розподілу.

При зменшенні кількості ключових точок майже пропорційно зменшуються й максимальна та фактична відстань між розподілами. Тобто збільшення числа дескрипторів від 100 до 300 не має суттєвого впливу на характеристики розпізнавання зображень. Значення таблиці 2.2 показують, що достатньо 100 ключових точок для розрізнення тестових зображень гербів. Для інших типів зображень або кількості розпізнаваних еталонів ці значення можуть відрізнятися.

Таблиця 2.2 – Значення нормованої відстані (1.7) для різної кількості дескрипторів та бітів розподілу

Кількість дескрипторів	Число бітів розподілу							
	1	2	4	8	128	256	Кожен 2-й	Кожен 4-й
500	0,142	0,098	0,073	0,060	0,046	0,046	0,138	0,144
300	0,139	0,095	0,071	0,061	0,045	0,045	0,147	0,147
100	0,147	0,102	0,074	0,061	0,046	0,046	0,105	0,155

У таблиці 2.3 наведено дані наявних обчислювальних витрат для реалізації варіантів розглянутих моделей при обчисленні релевантності описів зображень гербів Масачуссета та Оксфорду в залежності від кількості бітів у фрагментах розподілу. Таблиця 2.3 містить час опрацювання комп'ютерними моделями у мілісекундах.

Таблиця 2.3 – Оцінка часу при обчисленні релевантності

Число дескрипторів	Число бітів розподілу							
	1	2	4	8	128	256	Кожен 2-й	Кожен 4-й
500	66,92	33,94	26,22	22,09	20,38	19,91	32,49	26,51
300	46,46	65,17	34,17	22,32	19,61	19,91	33,13	26,08
100	47,01	32,58	24,54	21,47	18,57	16,29	32,13	25,38

Аналізуючи значення табл. 2.2 та оцінку часу оброблення з таблиці 2.3 для 100 КТ, можна сказати, що при обчисленні по правилу «кожен другий» час прогнозовано зменшився у порівнянні з 1-бітовим розподілом приблизно у 1,5 рази. А нормована відстань (1.7) дещо зменшилась від 0,147 до 0,105.

При обчисленні по правилу «кожен четвертий» час зменшився у порівнянні з 1-бітовим розподілом у 1,9 рази. А нормована відстань (1.7) практично не змінилась в межах 0,15. Ці показники стверджують про достатні властивості модифікацій побудованих ознак для розпізнавання зображень гербів зі значно меншим часом оброблення.

Відмітимо, що оцінка часу у таблиці 2.3 враховує тільки знаходження відстані, не включаючи затрати на обчислення розподілу, тобто програма орієнтована на порівняння з еталонними розподілами, побудованими на етапі попереднього оброблення. У цьому випадку, як видно із табл. 2.3, не має суттєвого значення кількість КТ, а важлива тільки кількість бітів у фрагментах розподілу. Чим більше використано бітів – тим повільніше працює програма. Як можна побачити із таблиці 2.3, з кожним збільшенням числа бітів у розподілі у 2 рази час виконання зменшується приблизно у 1,3 рази. Зрозуміло, що, якщо брати до уваги час, затрачений на обчислення розподілу, то він прямо пропорційний числу дескрипторів опису.

Зважаючи на те, що час обчислення релевантності за розподілами приблизно у 1000 разів менший, ніж для традиційної процедури голосування, бачимо, що за допомогою впровадження системи ієрархічних ознак можна добитися подальшого скорочення часу обчислень, забезпечуючи при цьому достатньо високу результативність розпізнавання.

Ступінь інтегрованості призводить до зменшення рівня відмінності ознак від 0,147 (для 1-бітового) до 0,06 для 8-бітового, що зменшує розрізненість описів у інтегрованому поданні. Але і для 256 бітів спостерігаються відмінності між ієрархічними описами гербів, що дозволяє їх розпізнавати між собою. При цьому час для 256 бітів зменшився майже в 3 рази.

2.3 Аналіз результатів дослідження

Проведені експерименти, результати яких показали доцільність формування ієрархічної системи ознак при розпізнаванні зображень. Для довільної бази описів зображень формуємо ієрархічну систему ознак. Якщо розпізнавання здійснюється впевнено за 1 спільною ознакою (загалом для 256 бітів), то застосовуємо цей спосіб, тому що маємо виграш у три рази за часом обчислення. Якщо ні – опускаємося на рівень нижче і так далі шукаємо задовільний результат.

При зменшенні кількості ключових точок майже пропорційно зменшуються й максимальна та фактична відстань між розподілами. Тобто збільшення числа дескрипторів від 100 до 300 не має суттєвого впливу на характеристики розпізнавання зображень. Значення таблиці 2.2 показують, що 100 ключових точок достатньо для розрізнення зображень гербів. Для інших типів зображень або кількості розпізнаваних еталонів ці значення можуть бути другими.

Зважаючи на те, що час обчислення релевантності за розподілами приблизно у 1000 разів менший, ніж для традиційної процедури голосування, бачимо, що за допомогою впровадження системи ієрархічних ознак можна добитися суттєвого скорочення часу обчислень, забезпечуючи при цьому достатньо високу результативність розпізнавання.

ВИСНОВКИ

У роботі з використанням просторової обробки структурних описів вирішена задача класифікації зображень на базі структурного опису у вигляді множини особливих точок, що визначаються детектором ORB. Для цього проаналізовано особливості детектору, запропоновано способи побудови ієрархічних ознак.

Трансформація ймовірнісних розподілів для фрагментів дескрипторів опису і зіставлення образів у просторі ієрархічних ознак забезпечують необхідну результативність розпізнавання. Таке оброблення реалізується значно швидше, ніж у просторі розподілів та у процедурах голосування.

Оброблення за значеннями згрупованих розподілів формує ієрархічну структуру даних зі змінюваним параметром деталізації та значимими властивостями опису.

Практичні рекомендації із проведеного дослідження полягають у результативному застосуванні інтегрованих ознак для груп бітів як оптимального з точки зору швидкодії обчислень.

Наукова новизна дослідження полягає в удосконаленні методу структурного розпізнавання зображень на основі впровадження блочної структури опису із використанням інтегрованих значень розподілу для фрагментів множини дескрипторів.

Практична значущість роботи – досягнення суттєвого рівня підвищення швидкодії при обчисленні релевантності, підтвердження результативності запропонованих ієрархічних ознак на прикладах зображень, отримання прикладних програмних моделей для дослідження та впровадження методів класифікації в системах комп'ютерного зору.

Перспективи дослідження можуть бути пов'язані із побудовою дерев рішень на основі ієрархічних ознак задля забезпечення потрібного рівня результативності системи розпізнавання.

ПЕРЕЛІК ПОСИЛАНЬ

1. Гороховатский В.А. Применение пространственных структур признаков для классификации изображений в компьютерном зрении (монография) / В.А. Гороховатский, Т.В. Полякова. – Х.: ФОРМ Панов А.Н., 2018. – 120 с.
2. Гонсалес Р. Цифровая обработка изображений [Текст] / Р. Гонсалес, Р. Вудс; перевод с англ. – М.: Техносфера, 2005. – 1072 с.
3. Stefan Leutenegger, Margarita Chli, Roland Y. Siegwart. BRISK: Binary Robust Invariant Scalable Keypoints. – Computer Vision (ICCV), pp. 2548 – 2555, 2011.
4. Gorokhovatsky V.A. Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions / V.A. Gorokhovatsky // Telecommunications and Radio Engineering. – 2016, Vol. 75, No 14. – pp. 1271–1283.
5. Gorokhovatsky, V.O. and Gadetska, S.V., (2019) Determination of Relevance of Visual Object Images by Application of Statistical Analysis of Regarding Fragment Representation of their Descriptions, Telecommunications and Radio Engineering, 78 (3), pp. 211–220.
6. Rublee, E., Rabaud, V., Konolige, K., and Bradski, G., (2011) ORB: an efficient alternative to SIFT or SURF, IEEE International Conference on Computer Vision (ICCV), Proceedings, pp. 2564-2571.
7. Гороховатский В. А. Построение и применение функции принадлежности для результативного структурного анализа изображений / В. А. Гороховатский, Ю. А. Куликов, А. Е. Берестовский // Бионика интеллекта. – 2015. – №1 (84). – С. 13–18.
8. Шапиро Л. Компьютерное зрение/ Л. Шапиро, Дж. Стокман.; [пер. с англ. А.А. Богуславского, под ред. С.М. Соколова]. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.

9. Гаганов В. Инвариантные алгоритмы сопоставления точечных особенностей на изображениях // В. Гаганов // Компьютерная графика и мультимедиа, 2009.–№1 (17).– [Электронный ресурс]. – Режим доступа: http://cgm.computergraphics.ru/issues/issue17/invariant_features

10. Прохоренок Н.А. OpenCV и Java. Обработка изображений и компьютерное зрение / Н. Прохоренок. – СПб.: БХВ–Петербург, 2018. – 320с.

11. Gadetska S.V. Statistical Measures for Computation of the Image Relevance of Visual Objects in the Structural Image Classification Methods / S.V. Gadetska, V.O. Gorokhovatsky // Telecommunications and Radio Engineering. – 2018, Vol. 77 (12), pp. 1041– 1053.

12. Гороховатский В. А. Структурное распознавание изображений с применением моделей интеллектуальной обработки и самоорганизации признаков / А. В. Гороховатский, А. Е. Берестовский // Радіоелектроніка, інформатика, управління. – 2016. – № 3. – С. 39–46

13. Han, J. Data Mining: Concepts and Techniques / J. Han, M. Kamber. – Amsterdam e.a.: Morgan Kaufmann Publishers, 2006. – 754 p

14. Путятин Е.П. Обработка изображений в робототехнике / Е.П. Путятин, С.И. Аверин. – М.: Машиностроение, 1990. – 320 с.

15. Чабан Л. Н. Методы и алгоритмы распознавания образов в автоматизированном дешифрировании данных дистанционного зондирования: учебное пособие. – М.: МИИГАиК, 2016, – 94 с

16. A. Zamula and S. Kavun, Complex systems modeling with intelligent control elements, Int. J. Model. Simul. Sci. Comput. 08(01), 1750009 (2017). [Электронный ресурс]. – Режим доступа: <https://www.worldscientific.com/doi/abs/10.1142/S179396231750009X>

17. Гороховатський В.О. Статистичні розподіли та ланцюжкове подання даних при визначенні релевантності структурних описів візуальних об'єктів / В.О. Гороховатський, С. В. Гадецька , Р.П. Пономаренко // Системи управління, навігації та зв'язку. – 2018. – №6 (52). – С. 87–92.

18. Сірик Т. О. Структурна класифікація зображень з використанням просторового оброблення описів/ Т. О. Сірик // Радіоелектроніка та молодь у XXI столітті: 23 міжнародний молодіжний форум, 16-18 квітня 2019 р.–154 с.

19. Гороховатський В.О. Дослідження системи ієрархічних ознак при блочному поданні опису у складі множини ключових точок зображення / В.О. Гороховатський, Д.О. Руденко, Т.О. Сірик // Системи управління, навігації та зв'язку. – 2019. – №2 (54). – С. 69–73.

20. Гороховатський, В.О., Запорожченко, А.П., Сірик, Т.О., Тарасенко О.П. (2020) Дослідження результативності застосування ознак розподілів даних для обчислення релевантності описів зображень. Системи управління, навігації та зв'язку. – №1 (подана до друку)