

В. О. Гороховатський¹, Н. В. Власенко², М. О. Рибалка¹

¹ Харківський національний університет радіоелектроніки, Харків, Україна

² Харківський національний економічний університет імені С. Кузнеця, Харків, Україна

ЗАСТОСУВАННЯ ЗАСОБІВ ХЕШУВАННЯ ДАНИХ ДЛЯ ПРИСКОРЕННЯ КЛАСИФІКАЦІЙНИХ РІШЕНЬ У СТРУКТУРНИХ МЕТОДАХ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

Анотація. Предметом досліджень є методи класифікації зображень за множиною дескрипторів ключових точок. Метою є підвищення продуктивності методів класифікації, зокрема, прискорення часових показників класифікації шляхом впровадження засобів хешування для подання еталонних даних. **Методи, що застосовуються:** детектор та дескриптори ORB, засоби хешування даних, методи пошуку в масивах даних, апарат визначення релевантності векторів на основі метрик, програмне моделювання. **Отримані результати:** розроблено ефективний метод класифікації зображень на основі впровадження швидкісного пошуку із використанням хеш-структур, що прискорює обчислення в десятки разів; час класифікації для розглянутих експериментальних описів лінійно зростає зі зменшенням числа хеш-кошиків; вибір порогу для значення мінімуму метрики при встановленні класу для дескрипторів об'єкту суттєво впливає на точність класифікації; вибір такого порогу може бути оптимізовано для фіксованих баз зразків; експериментально досягнута точність класифікації вказує на працездатність та результативність запропонованого методу на підґрунті хешування даних. **Практична значущість роботи** – побудова моделей класифікації у просторі хеш-подання даних, підтвердження працездатності запропонованих модифікацій класифікаторів на прикладах зображень, розроблення прикладних програмних моделей для впровадження запропонованих методів класифікації у системах комп'ютерного зору.

Ключові слова: комп'ютерний зір; структурні методи класифікації зображень; дескриптор ORB; хешування, лінійний пошук, хеш-кошик, швидкодія оброблення, точність класифікації.

Вступ

Сучасні структурні методи класифікації зображень, що застосовуються у системах комп'ютерного зору, ґрунтуються на даних про образи візуальних об'єктів у вигляді множини дескрипторів ключових точок (КТ) [1-3]. Дескриптор КТ – це, як правило, вектор розміром 64...512 бінарних компонентів. Побудова класифікатора заснована на визначенні ступеня релевантності типу множини-множина між описами розпізнаваного об'єкту та еталону та оптимізації релевантності на множині еталонів.

В загальному випадку клас візуального об'єкта при цьому визначається як нескінченна множина зображень об'єкту з урахуванням його допустимих геометричних перетворень: зсувів, поворотів, масштабування в межах поля зору комп'ютерної системи [1]. Представником класу для класифікатора є еталонне зображення, за яким і визначаємо образ (множину ознак) еталона. Набір обраних користувачем еталонних образів створює скінченну базу зразків. У рамках даних бази еталонів здійснюється класифікація об'єктів, що полягає у віднесенні вхідного зображення до одного із еталонних класів або у відмові від класифікації у разі недостатнього для прийняття рішення рівня релевантності з еталонами [1-4].

При несуттєвому рівні впливу фону та завад на аналізованому зображенні отримав застосування класифікаційний підхід, де на підставі значень ком-плектної множини дескрипторів опису формується система інтегрованих характеристик [1]. Зрозуміло, що впровадження інтегрованої системи ознак суттєво скорочує обчислювальні витрати на класифікацію за рахунок принципу узагальнення інформації. Але більш універсальним у аспекті завадостійкості все-

таки вважається по-елементний аналіз компонентів опису, що на основі формування статистичних характеристик цілісного образу дає можливість у процесі класифікації частково знизити вплив перешкод [2].

Для способу по-елементного аналізу традиційний підхід структурної класифікації зводиться до пошуку для кожного дескриптора із опису розпізнаваного об'єкта деякої підмножини релевантних йому у комплектній множині дескрипторів бази еталонів та підрахунку числа дескрипторів об'єкту, віднесених відповідно до кожного із класів. Зважаючи на те, що число дескрипторів у описі кожного еталону досягає 300...500, а класів може бути десятки і навіть сотні, такий підхід потребує досить громіздких обчислень [1, 3, 5].

Впровадження засобів хешування даних здатне суттєво скоротити часові затрати у процесі практичної реалізації процедур пошуку та класифікації. При хешуванні дані бази еталонів попередньо розкладають на групи подібних за деякою ознакою елементів, що значно (у тисячі разів) прискорює процес оброблення за рахунок відповідного зростання вимог до об'ємів пам'яті [3, 6, 7]. Хешування реалізує перспективну ідею зниження розмірності шляхом попередньої побудови для еталонної множини деякої індексованої структури даних, що загалом сприяє спрощенню оброблення та знижує обчислювальні затрати на класифікацію.

У роботах [8, 9] розвинуто наближені методи швидкого метричного пошуку за зразком на підставі створення індексованих структур даних на множині ієрархічних ознак зображень, для яких здійснюється сегментація значень функції яскравості. Спеціалізована індексована структура на підґрунті принципу

кластеризації об'ємних масивів багатовимірних даних впроваджена також у розробках професійних систем пошуку [10], де задля покращення продуктивності пошуку компонентів візуальних даних застосована багатоетапна схема кластеризації з моделлю наближеного пошуку релевантних елементів всередині кластеру. Створення кластерних структур даних як сучасний підхід інтелектуального аналізу дає можливість краще адаптуватися до змісту еталонної інформації, що застосовано нами при визначенні релевантності візуальних об'єктів у технологіях «мішок слів» [1, 11, 16]. Алгоритмічні та програмно-апаратні засоби швидкого пошуку у об'ємних сховищах даних (наприклад, у веб-застосунках) успішно розвиваються і знаходять впровадження у сучасних прикладних пошукових системах [5, 10, 12, 14].

Зрозуміло, що для систем комп'ютерного зору ключовим критерієм при впровадженні засобів швидкого пошуку є все-таки гарантоване забезпечення достатніх показників результативності класифікації.

Мета роботи – розвинення та підвищення продуктивності методів класифікації зображень шляхом впровадження індексованих хеш-структур на множині дескрипторів бази еталонних образів.

Завдання дослідження полягають у побудові хеш-структур даних на множині дескрипторів бази еталонів, впровадження цих структур у процес класифікації, вивчення результативності та часових показників розроблених модифікацій класифікаторів.

Двоетапна класифікація

Розглянемо простір B^n багатовимірних бінарних векторів розмірністю n , у якому будемо конструювати образи розпізнаваного об'єкту і еталонів. Зафіксуємо деяку мультимножину векторів $E_i \subseteq B^n$ як опис еталону $E_i = \{e_v(i)\}_{v=1}^s$ у просторі множин дескрипторів КТ, $s = \text{card } E_i$ – число дескрипторів у множині [1]. Окремі ознаки – це вектори $e_k \in B^n$, скінченна множина яких створює опис об'єкту.

Класифікація передбачає наявність деякої бази E описів еталонних зображень розмірністю $N : E = \{E_1, E_2, \dots, E_N\}$. Кожний еталонний опис E_i репрезентує окремий клас, представлений скінченною множиною дескрипторів КТ. Тепер розглянемо довільний опис $Z \subset B^n$, $Z = \{z_w\}_{w=1}^s$ розпізнаваного об'єкту. Поставимо задачу побудови класифікатора K як відображення $K : Z \rightarrow [1, 2, \dots, N]$ на основі попереднього конструювання деякої індексованої структури на множині E .

Класифікацію K представимо у вигляді двоетапного процесу $K = K_2 K_1$, де на першому етапі $K_1 : B^n \rightarrow [1, 2, \dots, N]$ здійснюємо визначення класу d_w для кожного дескриптора $z_w \in Z$, а на другому етапі $K_2 : D \rightarrow [1, 2, \dots, N]$ із множини $D = \{d_w\}_{w=1}^s$ отриманих локальних рішень першого

етапу формуємо результуючий висновок про клас об'єкту Z . Такий принцип оброблення узагальнює множину рішень однорідних локальних класифікаторів [2, 14]. Фактично K_1 тут можна розглядати як багатозначну характеристичну функцію для визначення еталонного класу E_i щодо окремого дескриптора із опису розпізнаваного об'єкта.

Етапи K_1 , K_2 можна побудувати і в інших варіантах, наприклад, шляхом створення на етапі K_1 деякого ймовірнісного розподілу за класами еталонів, включаючи процедури логічного оброблення таких розподілів [2].

Конструктивно реалізація K спирається на апріорні дані наявної бази E , так як належність усіх $e_v(i)$ до відповідного образу E_i всередині бази уже відома на початку класифікації.

Якщо класифікацію K_1 здійснювати традиційно методом лінійного пошуку (повний перебір) шляхом послідовного аналізу кожного елемента набору E , то застосовують конкурентне правило [2, 13-15]

$$d_w = \arg \min_{i,v} \rho(z_w, e_v(i)), \quad (1)$$

де d_w – номер еталону E_i , до якого буде віднесено дескриптор z_w об'єкту, $d_w \in \{1, \dots, N\}$, $\rho(z_w, e_v(i))$ – метрика у векторному просторі.

Важливим задля результативної класифікації є впровадження логічного аналізу значення мінімуму, що досягнуто при здійсненні (1) [16]. Значення d_w визначається тільки тоді, якщо отриманий мінімум відстані не перевищує заданий поріг δ_ρ : $\rho_{\min} \leq \delta_\rho$. У протилежному випадку клас d_w не визначається, тобто дескриптор вважається хибним.

Загалом необхідну кількість Q обчислень значень метрики в (1) у повній базі еталонів лінійним пошуком можна оцінити значенням параметру $Q = N \cdot s^2$, вважаючи обсяги описів еталонів та об'єкта рівноцінними. Для векторів простору B^n в (1) може бути застосована проста в обчислювальному сенсі метрика Хемінга, що підраховує число не співпадаючих бітів для векторів [1].

Етап класифікації K_2 зводиться до того, що на кожному кроці аналізу опису Z по правилу (1) за значенням d_w інкрементується число r_i голосів елементів, віднесених до i -го класу

$$\begin{cases} r_i + 1, & d_w = i; \\ r_i, & d_w \neq i, \end{cases} \quad (2)$$

а клас i_0 образу Z об'єкта визначаємо за максимумом числа голосів:

$$i_0 = \arg \max_{i=1, \dots, N} r_i. \quad (3)$$

Значення $\{r_i\}_{i=1}^N$ відображають гістограму класів за числом голосів елементів із Z . Вирази (2),

(3) конкретизують етап K_2 , що полягає у обробленні голосів для компонентів опису Z . Розглянута двоетапна процедура класифікації базується на основоположному принципі інтелектуального аналізу даних, що полягає у підрахунку числа позитивних рішень (підтримка, рейтинг зустрічальності) на аналізованій множині даних [14, 17].

Розглянуті варіанти побудови класифікатора природно можна трактувати в рамках теорії ансамблевих моделей [2,14], де за рахунок створення та агрегування відгуків компонентних класифікаторів (локальних рішень) синтезується «сильний» класифікатор із гарантовано вищою результативністю прийняття рішень. Це відповідає моделі бустінгу [14, 17].

Ключовою обчислювальною проблемою при побудові класифікатора для об'ємних баз даних є реалізація пошуку (1), що фактично є традиційним підходом ближнього сусіда. Зосередимося на її вирішенні шляхом застосування у двоетапній процедурі класифікації спеціалізованої структури даних, що спрямована на зменшення обсягу обчислень Q .

Індексована структура даних для організації швидкого пошуку

На етапі попереднього оброблення створимо у еталонній множині E образів спеціалізовану структуру задля забезпечення швидкісної результативної класифікації.

Здійснимо розбиття T на множині E дескрипторів бази зображень, отримуємо множину із M непересічних груп $T_k(E)$:

$$E = T(E) = \bigcup_{k=1}^M T_k(E),$$

$$T_k(E) \cap T_j(E) = \emptyset. \quad (4)$$

Будемо мати на увазі два основних та найбільш поширених способи розбиття: хешування та кластеризацію [3, 16]. Загалом впровадження хешування пов'язують із точними засобами оброблення, а кластеризацію – із наближеними способами самонавчання для вирішення прикладних задач [4, 9, 12]. Трансформація (4) здійснює попередню класифікацію і повністю зберігає усю сукупність аналізованих даних, яка тепер розподіляється між групами $T_k(E)$. У результаті перетворення (4) кожний дескриптор $e_v \in E$ бази еталонів отримує параметр k номеру групи (кошика).

Зважаючи на уже існуюче розбиття $E = \bigcup_{i=1}^N E_i$

даних на окремі еталонні образи, визначимо величину

$$t_{i,k} = \text{card}\{e_v \mid e_v \in E_i \ \& \ e_v \in T_k\} \quad (5)$$

як число дескрипторів i -го класу, що потрапили до кошика T_k . На підставі $t_{i,k}$ маємо вагову кількісну характеристику $b_i(T_k)$ для кожного кластеру

$$b_i(T_k) = \frac{t_{i,k}}{\text{card} T_k}. \quad (6)$$

Вираз (6) визначає статистичний розподіл елементів кожного сегменту даних за класами еталонів у вигляді вагових коефіцієнтів класів, причому $\sum_{i=1}^N b_i = 1$. Розподіл b є спільною характеристикою усіх елементів сегменту даних, отриману за результатами аналізу (навчання) для бази E . На етапі навчання ефективним для підвищення результативності класифікації може бути запровадження логічного оброблення вектору b задля спрощення аналізу та підсилення впливу вагових коефіцієнтів b_i для найбільш значущих класів у кошику [1, 2].

Конкретно для хешування ознакою окремого кошику (кластеру) даних виступає значення хеш-функції.

Ще одним важливим параметром, що впливає на швидкодію пошуку, є число кошиків M . Чим менше число кошиків, тим вища швидкодія переходу до кошику. Чим більше число кошиків, тим менше об'єм даних для аналізу всередині кошиків. Граничними ситуаціями є один кошик ($M = 1$) та повна відсутність хешування ($M = N \cdot s$), що відповідають лінійному пошуку.

За думкою дослідників можна формулювати задачу оптимізації числа M кошиків, де критерієм виступає число Q обчислень метрики для компонентів опису [9, 10].

Якщо в одному кошику в середньому число елементів є $\frac{N \cdot s}{M}$, то кількість обчислень (1) всередині індексної структури з використанням значень хеш-коду пропорційна $Q_1 = s \cdot M + \frac{N \cdot s}{M}$, що значно менше, ніж $Q_2 = N \cdot s^2$ для традиційного лінійного пошуку. Для конкретних значень $s = 500$, $N = 10$, $M = 10$ вираш $\beta = Q_1/Q_2$ складає приблизно 450 разів і зростає зі збільшенням обсягів N, s даних.

Задля подальшого прискорення процедури оброблення можна застосувати додаткове упорядкування отриманої структури даних, впровадивши, наприклад, сортування даних всередині кошиків T_k за деяким ключем $g(e_v)$ [1].

Метод оброблення даних з використанням хешування [3, 6], який фактично здійснює попередню класифікацію на окремі кошики та має свої особливості. Число кошиків для цього методу, як правило, фіксоване і пов'язане із діапазоном значень хеш-ключа $u(e_v)$ для хешування, яке задає користувач та вважають заданим.

Ключ – це, як правило, ціле число, що приймає множину фіксованих значень $u(e_v) \in U_0$, $U_0 = \{u_k\}_{k=1}^L$ – множина значень хеш-функції [6]. Для векторів із B^n таким ключем може бути, на-

приклад, число одиничних бітів у складі дескриптору, число вибраних комбінацій бітів довільної довжини та ін. Взагалі ознаку $u(e_v)$ для хешування можна корисно адаптувати до значень описів наявної бази еталонів.

Із впровадженням хешування розбиваємо множину даних E на непересічні групи $U_k(E)$ (ланки, кошики, полиці) числом L з рівнозначним значенням хеш-ключа $u(e_v) = u_k$

$$E = U(E) = \bigcup_{k=1}^L U_k(E),$$

$$U_i(E) \cap U_j(E) = \emptyset,$$

$$U_k(E) = \{e_v \in E \mid u(e_v) \in u_k\}. \quad (7)$$

Класифікація на основі хеш-структури

Застосовуємо детектор КТ до зображення розпізнаваного об'єкту, формуємо опис $Z = \{z_w\}_{w=1}^s$ об'єкту як множину дескрипторів. Побудуємо класифікатор з використанням швидкісного пошуку у базі E на підставі створеної структури.

Спосіб із впровадженням хешування включає такі дії.

1. Для дескриптора z_w вхідного образу обчислюємо величину хеш-ключа $u_w = u(z_w)$, $u_w \in U_0$ як значення хеш-функції для аргументу z_w .

2. Переходимо до корзини хешу зі значенням u_w , а далі реалізуємо один із способів пошуку в межах кошика: лінійний, найближчих сусідів, за розподілами даних чи інші. Лінійний пошук та визначення найближчих сусідів в межах ланки реалізують точну процедуру. Використання оцінок розподілів за класами для окремих ланок хешу сприяє прискоренню обчислень і відповідає наближеному оціночному способу на підставі навчання.

3. За результатом пошуку на етапі 2 для елемента z_w визначаємо клас d_w .

4. На підставі аналізу усієї множини Z дескрипторів об'єкту накопичуємо лінійку $\{r_i\}_{i=1}^N$ значень голосів за кожний з еталонних класів.

5. За виразом (3) класифікуємо об'єкт до класу, що набрав найбільшу кількість голосів.

Як бачимо, запропонований продуктивний за швидкістю спосіб класифікації на підґрунті впровадження хешування може реалізувати як точний, так і наближений види пошуку. Параметрами є число груп (кошиків хешу), побудованих для компонентів даних в індексованій структурі. Цим параметром можна керувати, виходячи із наповненості кошиків для конкретних даних.

Розподіли за класами еталонів для створених груп даних теж несуть важливу інформацію для класифікації, на підставі чого можна видалити із аналізу окремі кошики (якщо класи всередині них розподілені суто рівномірно), або за значеннями розподілу приймати однозначне класифікаційне

рішення (якщо спостерігаються суттєві переваги для окремих класів).

За результатами попереднього оброблення еталонної інформації (навчальна, контрольна вибірка) можна оцінити результативність застосування класифікатора. Точність $prec$ класифікації можна оцінити відношенням гуртового числа ТР дескрипторів бази еталонів, для яких правильно визначено клас, до загального їх числа у відповідності до моделі [2, 14]

$$prec = \frac{TP}{TP + FP}, \quad (8)$$

де TP – число дескрипторів з правильно визначеним класом, FP – число дескрипторів з неправильно визначеним класом. Найвища точність відповідає значенню $prec = 1$.

Результати комп'ютерного моделювання

Моделювання виконано у середовищі PyCharm 2018 із використанням засобів бібліотеки Open CV та мови програмування Python [18-21]. Для визначення дескрипторів КТ застосовано детектор ключових точок ORB [18, 20] з дескриптором розмірністю $n=256$. Розроблені моделі класифікатора вжито на прикладі зображень порід собак (німецька вівчарка, хаскі, сербернар), розмір зображень масштабовано до 500x500 пікселів. Ілюстрація класів еталонних зображень та сформовані координати КТ показані на рис. 1. Кількість обчислених дескрипторів у описі кожного із еталонів складає $s=500$.

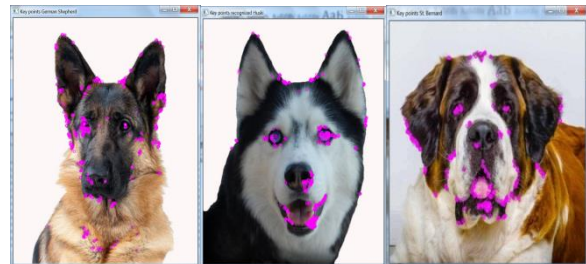


Рис. 1. Еталонні зображення порід собак з координатами ключових точок
(Fig. 1. Reference images of dog breeds with the coordinates of key points)

Як бачимо, візуально ці зображення за зовнішніми ознаками досить близькі. Якщо у якості хеш-функції $u(z_w)$ взяти число одиничних бітів у векторі розмірністю 256, то кількість вибраних кошиків змінюється у межах від 1 до 257, тобто найбільша множина U_0 містить значення 0, 1, 2, ..., 256. Всередині кошиків у експерименті для спрощення застосовано традиційний лінійний пошук. Зважаючи на бінарний вид аналізованих даних, для зіставлення векторів застосовано метрику Хемінга.

З метою зіставлення характеристик швидкодії програмно змодельований метод лінійного пошуку на усій множині еталонних даних (1500 дескрипторів) без застосування хешування. Для еталонних зображень із застосуванням хешування на 256 кошиків та без хешування отримано максимально

точний результат: усі 500 дескрипторів віднесені до свого еталону. При цьому час оброблення без хешування склав приблизно 40 сек, а із хешуванням – лише 0,6 сек. Як бачимо, вигреш у часі обчислень для запропонованого підходу (при заданому числі 3 еталонів та $s=500$) складає більше 65 разів у порівнянні із традиційним. При цьому ясно, що вигреш у часі обчислень збільшується при зростанні числа еталонів та дескрипторів у описах Аналогічні точні результати класифікації еталонів отримано і для іншої кількості кошиків.

Вивчення отриманої множини дескрипторів за визначеною хеш-функцією для зображень рис. 1 показало, що в межах її значень від 0 до 92 та від 217 до 255 взагалі немає жодного дескриптора. Розподіл дескрипторів для досліджуваної бази еталонів на 3 кошки, рівноцінні за діапазоном значень хеш-функції, показано у табл. 1.

Таблиця 1 – Кількісний розподіл дескрипторів бази

Значення хеш-функції	Число дескрипторів
93 - 133	375
134 - 174	1004
175 - 216	121

Експериментальний результат табл. 1 відкриває перспективу подальшого скорочення об'єму обчислень, так як допускає виключення із аналізу суттєвої не використовуваної частини від загального числа кошиків. У той же час ця можливість сприяє забезпеченню завадостійкості роботи методу з хешуванням.

Хешування відносять до точних методів трансформації даних. Природно зрозуміло, що похибка класифікації із впровадженням хешування може виникнути із-за того, що всередині корзини згрупованих даних можуть зустрітися рівноцінні дескриптори різних класів.

Важливим для забезпечення результативної класифікації представляється вплив параметра порогу δ_p для значення мінімальної відстані при прийнятті рішення про клас дескриптора об'єкта за правилом (1). Діапазон значень для вибраної метрики складає $\delta_p \in [0, \dots, 256]$.

Наше моделювання для навчальної вибірки показало, що зі зменшенням порогу відсікається більша частина «чужих» дескрипторів (інших еталонів), але при цьому одночасно губиться і значна частина «своїх». Зі зростанням порогу δ_p спостерігається збільшення числа «своїх» з одночасним збільшенням числа «чужих». Як бачимо, виникає конкурентна ситуація, розв'язання якої може бути досягнуто шляхом оптимізації значення порога конкретно для наявної бази еталонів, а критерієм виступає точність (8) класифікації, результативним фактом при цьому є переважаюче число правильно класифікованих дескрипторів для свого класу. Оптимізація універсально може бути реалізована експериментальним шляхом для прикладної бази зображень. У нашому дослідженні найкращі результати досягнуто при значенні $\delta_p=100$.

Найбільш критично вибір порога впливає на результат класифікації при геометричних перетвореннях вхідного зображення. Експерименти з дією перетворення повороту (найбільшим чином викривляє зображення, рис. 2) для вхідних еталонних зображень показали, що при низьких порогах втрачається статистично значиме число правильно класифікованих дескрипторів, що потребує збільшувати значення порогу.

У той же час навіть при невеликому значенні порогу $\delta_p = 25$ вхідні зображення еталонів (навчальна вибірка) класифікуються вірно, незважаючи на суттєве зниження числа виявлених «своїх» даних (від 330 до 129 для різних еталонів).

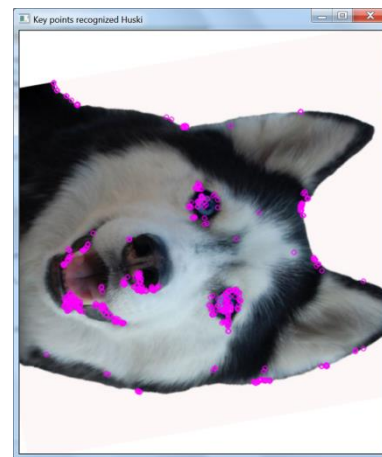


Рис. 2. Зображення з координатами ключових точок при повороті (Fig. 2. Images with coordinates of key points when turning)

Таблиця 2 – Залежність часу класифікації від числа кошиків

Число кошиків	257	128	64	32	16	8	1
Час класифікації	0,6	1,3	2,5	4,9	9,5	18,4	40

Дослідимо тепер детальніше залежність часу класифікації (в сек) від числа кошиків. Експериментальні дані поміщені в табл. 2, число кошиків 1 відповідає варіанту без хешування.

Як бачимо, експериментальний час класифікації для розглянутих описів лінійно зростає зі зменшенням числа кошиків при застосуванні хешування.

Таким чином, при необхідності прийняття найбільш швидкісних рішень треба вибирати максимально можливе число кошиків. Отримана залежність властива виключно для аналізованих еталонних зображень.

У загальному плані необхідно додатково вивчати цю залежність для досліджуваної бази даних. Найбільш несприятливі для обсягів часу є ситуації, коли, наприклад, значна кількість дескрипторів різних класів зосереджуються у спільній корзині. Для вибраної нами хеш-функції стосовно досліджу-

ваних описів найкращим вибором виглядає число кошиків 257.

Час класифікації спадає зі збільшенням числа кошиків через те, що зменшуються кількості дескрипторів всередині кошиків, а після знаходження кошика обчислюється відстань Хемінга уже для меншої кількості дескрипторів. Пошук кошика за значенням хеш-функції відбувається значно швидше у порівнянні з обчисленням метрики для множини дескрипторів всередині кошиків.

Задля оцінювання ступеня відмінностей між описами еталонних даних обчислено манхетенську відстань для числа кошиків $L=257$. Ці відстані складають 234, 196, 274, що підкреслює досить значиму близькість еталонів у просторі ознак, так як максимум такої відстані дорівнює 1000 за числом дескрипторів для пари описів. Найменша відстань досягнута у порівнянні 1 та 3-го еталонів рис. 1.

Для навчальної вибірки, коли на вхід подаються еталони, значення параметра точності $prec=1$ як для традиційного, так і для пропонованого методу з використанням хешування. Але в реальній ситуації, наприклад, при перетвореннях повороту, величина $prec$ коливається у межах до 0,47 з хешуванням та 0,53 без хешування з правильним визначенням класу.

Незважаючи на деяке зниження показника точності, у проведених експериментах класифікація здійснювалася вірно, тобто усі вхідні зображення з перетвореннями правильно отримали свій номер класу. Ці результати підтверджують ефективність розробленого методу.

Таким чином, із введенням хешування і зростанням числа кошиків точність незначно знижується, але практично не залежить від числа кошиків. Основним недоліком впровадження хешування традиційно вважається зростання необхідних обсягів пам'яті комп'ютера [3, 6, 7].

Висновки

Розглянуті у роботі методи класифікації базуються на принципі «порівняння з еталоном» і можуть бути універсально застосовані до довільних векторів даних бітового виду. Ці методи базуються на двох різновидах класів – для цілісного опису (об'єкт та еталони) та для окремих компонентів опису. При застосуванні хеш-структур ключовим моментом є вибір результативної та адаптованої до даних хеш-функції, яка здійснює часткову класифікацію, не змінюючи дані.

Під час реалізації запропонованих підходів можливі варіанти оброблення з використанням наявної структурованості еталонних даних, тобто за окремими описами еталонів. Це дає можливість подальшого скорочення обсягів обчислень шляхом введення логічного аналізу за результатом послідовного обчислення релевантності для компонентів бази еталонів.

Наукову новизну проведеного дослідження складає розроблення продуктивного методу класифікації зображень на основі впровадження швидкісного пошуку із використанням хеш-структур, що зменшує обсяг обчислень в десятки разів.

Практична значущість роботи полягає у побудові моделей класифікації у структурованому просторі даних, підтвердженні працездатності запропонованих модифікацій на прикладах зображень, створенні програмних моделей для впровадження розроблених методів класифікації у системах комп'ютерного зору.

Перспективи дослідження можуть бути пов'язані із введенням логічного оброблення розподілів даних всередині хеш-кошиків, вивченням завадостійкості розроблених методів та оцінюванні їх прикладної результативності стосовно об'ємних колекцій зображень.

СПИСОК ЛІТЕРАТУРИ

1. Гороховатський В.О., Гадецька С.В. Статистичне оброблення та аналіз даних у структурних методах класифікації зображень (монографія). Харків : ФОП Панов А.Н., 2020. 128 с. DOI: <http://dx.doi.org/10.30837/978-617-7859-69-6>.
2. Гороховатський В.О., Гадецька С.В., Стяглик Н.І., Власенко Н.В. Класифікація зображень на підставі ансамблю статистичних розподілів за класами еталонів для компонентів структурного опису. *Радіоелектроніка, інформатика, управління*. 2020. №4. С. 85–94. DOI: <http://dx.doi.org/10.15588/1607-3274-2020-4-9>.
3. Gorokhovatskiy V.A., Gorokhovatskiy A. V., Peredrii Ye.O. Hashing of Structural Descriptions at Building of the Class Image Descriptor, Computing of Relevance and Classification of the Visual Objects. *Telecommunications and Radio Engineering*. 2018. Vol. 77 (13). P. 1159–1168. DOI: <http://dx.doi.org/10.1615/TelecomRadEng.v77.i13.40>.
4. Manning C.D., Raghavan P., and Schütze H. Introduction to Information Retrieval. Cambridge : University Press, 2008. 528 p.
5. Svyrydov A., Kuchuk H., Tsiapa O. Improving efficiency of image recognition process: Approach and case study. *Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies, DESSERT 2018, 2018*, pp. 593-597, DOI: <http://dx.doi.org/10.1109/DESSERT.2018.8409201>.
6. Ахо А., Хопкрофт Д., Ульман Д. Структуры данных и алгоритмы. Москва : Издательский дом «Вильямс», 2003. 384 с.
7. Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ. М.: ИД Вильямс, 2005. 1296 с.
8. Berman A., Shapiro L. A flexible image database system for content-based retrieval. *Computer Vision and Image Understanding*. Vol. 75, No. 1/2. 1999. P. 175–195.
9. Kinoshenko D., Mashtalir V., Yegorova E., Vinarsky V. Hierarchical partitions for content image retrieval from large-scale database. Machine Learning and Data Mining in Pattern Recognition / Perner P., Imlya A. (Eds.). *Lecture Notes in Artificial Intelligence*. Springer-Verlag, 2005. Vol. 3587. P. 445–455.
10. Babenko A., Slesarev A., Chigorin A., Lempitsky V. Neural codes for image retrieval. Conference Paper. *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2014, 8689 LNCS(PART 1), p. 584–599.

11. Gorokhovatskiy V., Putyatin Ye., Gorokhovatskiy O., Peredrii O. Quantization of the Space of Structural Image Features as a Way to Increase Recognition Performance. *The Second IEEE Int. Conf. on DataStream Mining & Proc. 21-25 August 2018*, Lviv, Ukraine. P. 464–467.
12. Лесковец Юре, Раджараман Ананд, Ульман Джефффри Д. Анализ больших наборов данных. Москва : ДМК Пресс, 2016. 498 с.
13. Гороховатский В.А., Пуятин Е.П., Столяров В.С. Исследование результативности структурных методов классификации изображений с применением кластерной модели данных. *Радиоэлектроника, информатика, управление*. 2017. №3 (42). С. 78–85.
14. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. Москва : ДМК Пресс, 2015. 400 с.
15. Гороховатский В.А. Иерархия пространственных отношений структурных признаков в задачах сопоставления визуальных объектов. Системи управління, навігації та зв'язку: зб. наук. пр. К.: ЦНДІ навігації і управління, 2008. Вип. 3(7). С. 85–89.
16. Gorokhovatskiy O., Gorokhovatskiy V., and Peredrii O. Analysis of Application of Cluster Descriptions in Space of Characteristic Image Features, *Data*, 2018, 3(4), 52. DOI: <http://dx.doi.org/10.3390/data3040052>.
17. Nong Ye. *Data Mining: Theories, Algorithms, and Examples* (1st. ed.). CRC Press, Inc., USA. 2013.
18. ORB feature detector and binary descriptor. URL: https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_orb.html.
19. Xu Zhang, Felix X. Yu, Svebor Karaman, Shih-Fu Chang. Learning Discriminative and Transformation Covariant Local Feature Detectors. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2017. P. 6818-6826.
20. Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski. ORB: “An efficient alternative to SIFT or SURF”, *International conference on computer vision*. 2011. IEEE. P. 2564-2571.
21. Yakovleva, O., Nikolaieva, K. Research of descriptor based image normalization and comparative analysis of SURF, SIFT, BRISK, ORB, KAZE, AKAZE descriptors. *Advanced Information Systems*. 2020. Vol. 4, No. 4. P. 89-101. DOI: <http://dx.doi.org/10.20998/2522-9052.2020.4.13>.

REFERENCES

1. Gorokhovatskiy, V.O. and Gadetska, S.V. (2020), *Statistical processing and data mining in structural image classification methods* (monograph), FLP Panov A.N., Kharkiv, 128 p., DOI: <http://dx.doi.org/10.30837/978-617-7859-69-6>.
2. Gorokhovatskiy, V.O, Gadetska, S.V., Styahlyk, N.I. and Vlasenko, N.V. (2020), “Classification of images based on an ensemble of statistical distributions by standard classes for structural description components”, *Radio Electronics, Computer Science, Control*, No. 4, pp. 85–94, DOI: <http://dx.doi.org/10.15588/1607-3274-2020-4-9>.
3. Gorokhovatskiy, V.A., Gorokhovatskiy, A.V. and Peredrii, Ye.O. (2018), “Hashing of Structural Descriptions at Building of the Class Image Descriptor, Computing of Relevance and Classification of the Visual Objects”, *Telecommunications and Radio Engineering*, Vol. 77 (13), pp. 1159–1168, DOI: <http://dx.doi.org/10.1615/TelecomRadEng.v77.i13.40>.
4. Manning, C.D., Raghavan, P. and Schütze, H., (2008), *Introduction to Information Retrieval*, Cambridge, University Press, 528 p.
5. Svyrydov, A., Kuchuk, H. and Tsiapa, O. (2018), “Improving efficiency of image recognition process: Approach and case study”, *Proceedings of 2018 IEEE 9th International Conference on Dependable Systems, Services and Technologies, DESSERT 2018*, pp. 593-597, DOI: <http://dx.doi.org/10.1109/DESSERT.2018.8409201>.
6. Akho, A., Hopcroft, D. and Ulman, D. (2003), *Data structures and algorithms*, Williams, Moscow, 384 p.
7. Cormen, T., Leizerson, Ch., Ryvest, R. and Shtain, K. (2005), *Algorithms: construction and analysis*, 2-nd edition, Publishing house "Williams", Moscow, 1296 p.
8. Berman, A. and Shapiro, L. (1999), “A flexible image database system for content-based retrieval”, *Computer Vision and Image Understanding*, Vol. 75, No. ½, pp. 175–195.
9. Kinoshenko, D., Mashtalir, V., Yegorova, E. and Vinarsky, V. (2005), “Hierarchical partitions for content image retrieval from large-scale database. Machine Learning and Data Mining in Pattern Recognition” / Perner, P., Imlya, A. (Eds.), *Lecture Notes in Artificial Intelligence*. Springer-Verlag. Vol. 3587. P. 445–455.
10. Babenko, A., Slesarev, A., Chigorin, A. and Lempitsky, V. (2014), “Neural codes for image retrieval”, Conference Paper. *Lecture Notes in Computer Science* (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8689 LNCS (PART 1), pp. 584-599.
11. Gorokhovatskiy, V., Putyatin Ye., Gorokhovatskiy, O. and Peredrii, O. (2018), “Quantization of the Space of Structural Image Features as a Way to Increase Recognition Performance”, *The Second IEEE International Conference on DataStream Mining & Processing 21-25 August 2018*, Lviv, Ukraine, pp. 464–467.
12. Leskovets, Yure, Radzharaman, Anand and Ulman, Dzheffry D. (2016), *Analyzing large datasets*, DMK Press, Moscow, 498 p.
13. Gorokhovatskiy, V.A., Putyatin, E.P. and Stolyarov, V.S. (2017), “Study of the effectiveness of structural methods of image classification using a cluster data model”, *Radioelectronics, informatics, management*, No. 3 (42), pp. 78–85.
14. Flach, P. (2015), *Machine learning. The science and art of building algorithms that extract knowledge from data*, DMK Press, Moscow, 400 p.
15. Gorokhovatskiy, V.O. (2008), “Hierarchy of spatial relations of structural features in the tasks of matching visual objects”, *Control, navigation and communication systems*, collection of scientific papers, Is. 3(7), pp. 85–89.
16. Gorokhovatskiy, O., Gorokhovatskiy, V. and Peredrii, O. (2018), “Analysis of Application of Cluster Descriptions in Space of Characteristic Image Features”, *Data*, 3(4), 52, DOI: <http://dx.doi.org/10.3390/data3040052>.
17. Nong, Ye. (2013), *Data Mining: Theories, Algorithms, and Examples*, 1st. ed., CRC Press, Inc., USA.
18. (2020), *ORB feature detector and binary descriptor*, available at: https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_orb.html.
19. Xu. Zhang, Felix, X. Yu, Svebor, Karaman and Shih-Fu, Chang (2017), “Learning Discriminative and Transformation Covariant Local Feature Detectors”, *The IEEE Conference on Computer Vision and Pattern Recognition*, pp. 6818-6826.

20. Ethan, Rublee, Vincent, Rabaud, Kurt, Konolige and Gary, Bradski (2011), ORB: "An efficient alternative to SIFT or SURF", *2011 International conference on computer vision*, IEEE, pp. 2564-2571.
21. Yakovleva, O. and Nikolaieva, K. (2020), "Research of descriptor based image normalization and comparative analysis of SURF, SIFT, BRISK, ORB, KAZE, AKAZE descriptors", *Advanced Information Systems*, Vol. 4, No. 4, pp. 89-101, DOI: <http://dx.doi.org/10.20998/2522-9052.2020.4.13>.

Received (Надійшла) 24.02.2021

Accepted for publication (Прийнята до друку) 09.04.2021

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Гороховатський Володимир Олексійович – доктор технічних наук, професор, професор кафедри інформатики, Харківський національний університет радіоелектроніки, Харків, Україна;

Volodymyr Gorokhovatsky – Doctor of Technical Sciences, Professor, Professor of Computer Science Department, Kharkiv National University of RadioElectronics, Kharkiv, Ukraine;

e-mail: gorohovatsky.vl@gmail.com; ORCID ID: <http://orcid.org/0000-0002-7839-6223>.

Власенко Наталія Володимирівна – кандидат технічних наук, старший викладач кафедри інформатики та комп'ютерної техніки, Харківський національний економічний університет імені С. Кузнеця, Харків, Україна;

Nataliia Vlasenko – Candidate of Technical Sciences, Senior lecturer at the Department of Informatics and Computer Engineering, Simon Kuznets Kharkiv National University of Economics, Kharkiv, Ukraine;

e-mail: gorohovatskaja@gmail.com; ORCID ID: <https://orcid.org/0000-0002-6841-7086>.

Рибалка Михайло Олексійович – студент кафедри інформатики, Харківський національний університет радіоелектроніки, Харків, Україна;

Mykhailo Rybalka – student of Computer Science Department, Kharkiv National University of Radioelectronics, Kharkiv, Ukraine;

e-mail: mykhailo.rybalka@nure.ua; ORCID ID: <https://orcid.org/0000-0002-4081-0281>.

Применение методов хеширования данных для ускорения классификационных решений в структурных методах распознавания изображений

В. А. Гороховатский, Н. В. Власенко, М. А. Рыбалка

Аннотация. Предметом исследований есть методы классификации изображений с помощью множества дескрипторов ключевых точек. Цель – повышение производительности методов классификации, в частности, ускорение временных показателей классификации путем внедрения средств хеширования для представления эталонных данных. Применяемые методы: детектор и дескрипторы ORB, средства хеширования данных, методы поиска в массивах данных, аппарат определения релевантности векторов на основе метрик, программное моделирование. Полученные результаты: разработан эффективный метод классификации изображений на основе внедрения скоростного поиска с использованием хэш-структур, что ускоряет вычисления в десятки раз; время классификации для рассмотренных экспериментальных описаний линейно возрастает с уменьшением числа хеш-корзин; выбор порога для значения минимума метрики при установлении класса для дескрипторов объекта существенно влияет на точность классификации; выбор такого порога может быть оптимизирован для фиксированных баз образцов; экспериментально достигнутая точность классификации указывает на работоспособность и результативность предложенного метода на основе хешированием данных. Практическая значимость работы – построение моделей классификации в пространстве хэш-представления данных, подтверждение работоспособности предложенных модификаций классификаторов на примерах изображений, разработка прикладных программных моделей для внедрения предложенных методов классификации в системах компьютерного зрения.

Ключевые слова: компьютерное зрение; структурные методы классификации изображений; дескриптор ORB; хеширование; линейный поиск; хеш-корзина; скорость обработки; точность классификации.

Application of data hashing tools to accelerate classification decisions in structural image recognition methods

Volodymyr Gorokhovatskyi, Nataliia Vlasenko, Mykhailo Rybalka

Анотація. The subject of this research is the image classification methods based on a set of key points descriptors. The goal is to increase the performance of classification methods, in particular, to improve the time characteristics of classification by introducing hashing tools for reference data representation. Methods used: ORB detector and descriptors, data hashing tools, search methods in data arrays, metrics-based apparatus for determining the relevance of vectors, software modeling. The obtained results: developed an effective method of image classification based on the introduction of high-speed search using hash structures, which speeds up the calculation dozens of times; the classification time for the considered experimental descriptions increases linearly with decreasing number of hashes; the minimum metric value limit choice on setting the class for object descriptors significantly affects the accuracy of classification; the choice of such limit can be optimized for fixed samples databases; the experimentally achieved accuracy of classification indicates the efficiency of the proposed method based on data hashing. The practical significance of the work is - the classification model's synthesis in the hash data representations space, efficiency proof of the proposed classifiers modifications on image examples, development of applied software models implementing the proposed classification methods in computer vision systems.

Keywords: computer vision; structural methods of image classification; ORB descriptor; hashing; linear search; hash; processing speed; classification accuracy.