

Problems of identification in information systems

УДК 004.932.2:004.93¹doi: <https://doi.org/10.20998/2522-9052.2021.3.01>В. О. Гороховатський¹, Н. І. Стяглик², В. В. Царевська²¹ Харківський національний університет радіоелектроніки, Харків, Україна² ННІ «Каразінський банківський інститут» ХНУ імені В.Н. Каразіна, Харків, Україна

КОМБІНАЦІЙНИЙ МЕТОД ПРИСКОРЕНОГО МЕТРИЧНОГО ПОШУКУ ДАНИХ У ЗАДАЧАХ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ

Анотація. Предметом досліджень статті є методи класифікації зображень за множиною дескрипторів ключових точок у системах комп'ютерного зору. Метою є підвищення продуктивності структурних методів класифікації шляхом впровадження індексованих хеш-структур на множині дескрипторів бази еталонних образів та узгодженого ланцюжкового поєднання кількох етапів аналізу даних у процесі класифікації. Застосовувані методи: детектор та дескриптори BRISK, засоби хешування даних, методи пошуку в об'ємних масивах даних, метричний апарат визначення релевантності векторів, програмне моделювання. Отримані результати: розроблено ефективний метод класифікації зображень на основі впровадження швидкісного пошуку із використанням індексованих хеш-структур, що прискорює обчислення в десятки разів; вигреш у часі обчислень збільшується при зростанні числа еталонів та дескрипторів у описах; особливості класифікатора полягають у тому, що здійснюється не точний пошук, а із врахуванням допустимого відхилення даних від еталону; експериментально перевірена результативність класифікації, що вказує на працездатність та ефективність запропонованого методу. Практична значущість роботи – побудова моделей класифікації у трансформованому 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 BRISK; хешування із врахуванням близькості, хеш-кошик, швидкодія класифікації.

Вступ

Структурні методи класифікації зображень, що впроваджуються у сучасних системах комп'ютерного зору, ґрунтуються на даних про образи візуальних об'єктів у вигляді множини дескрипторів ключових точок (КТ) [1-5]. Дескриптор КТ – це вектор розміром 64...512 бінарних компонентів, що є апроксимацією для фрагмента функції яскравості зображення. Побудова традиційного класифікатора зображень заснована на метричному визначенні ступеня релевантності виду множина-множина між образами розпізнаваного об'єкту і еталону та оптимізації значення релевантності у базі еталонів.

Загалом клас візуального об'єкта визначається як нескінченна множина зображень об'єкту з урахуванням його допустимих геометричних перетворень: зсувів, поворотів, масштабування в межах поля зору системи [1-4]. Представником класу виступає еталонне зображення, за яким комп'ютерна система визначає образ (множину ознак) еталона. Множина образів користувачем еталонних образів створює скінченну базу зразків, у рамках якої здійснюється класифікація об'єктів, що полягає у віднесенні вхідного зображення до одного із еталонних класів або у відмові від класифікації [1, 2].

Більш універсальним у аспекті забезпечення завадостійкості вважається по-елементний аналіз компонентів опису, що на основі формування статистичних класифікаційних характеристик для цілісного образу дає можливість у процесі класифікації знизити вплив завад [2].

Для способу по-компонентного аналізу традиційний підхід структурної класифікації зводиться до визначення для кожного дескриптора із опису розпізнаваного об'єкта деякої підмножини релевантних йому у комплектній множині дескрипторів бази еталонів та підрахунку числа дескрипторів – компонентів об'єкту, віднесених відповідно до кожного із класів (голосування). Ступінь релевантності визначається на підставі обчислення значення метрики. Зважаючи на те, що число дескрипторів у описі еталону досягає 300...500, а класів може бути десятки і навіть сотні, такий підхід призводить до процедури повного перебору та потребує досить громіздких обчислень [3-5].

Впровадження засобів хешування даних суттєво скорочує часові затрати у процесі практичної реалізації процедур метричного 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 [5-7]. Хешування реалізує перспективну ідею зниження розмірності аналізованих даних шляхом попередньої побудови для еталонної множини деякої індексованої структури, що спрощує оброблення та знижує обчислювальні затрати на класифікацію.

У роботах [8, 9] досліджено наближені методи метричного пошуку за зразком на підставі створення індексованих структур даних на множині ієрархічних ознак сегментованих зображень. Індексована структура на підґрунті кластеризації для об'ємних масивів

багатовимірних даних впроваджена у професійних системах пошуку [10], де задля покращення продуктивності застосована багатоетапна схема кластеризації з моделлю наближеного пошуку релевантних елементів всередині кластеру. Створення кластерних структур як сучасний підхід інтелектуального аналізу даних дає можливість краще адаптуватися до змісту еталонної інформації, що застосовано нами при визначенні релевантності візуальних об'єктів у технологіях «мішок слів» [2, 11, 16, 17].

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 [6-8, 11-14].

У роботі [6] детально вивчається спектр засобів швидкісного пошуку, серед яких можна виділити застосування принципу комбінаційного поєднання ряду методів у вигляді ланцюжка: хешування, розбиття даних на блоки (проекування), LSH-оброблення (locality-sensitive hashing, тобто хешування із врахуванням близькості) у вигляді сортування за значенням хеш-функції для блоків даних, логічний аналіз (узгоджена фільтрація) значень хеш-функції для блоків. Ідея LSH-оброблення зосереджує увагу тільки на парах даних, які можуть бути схожими, щоб не перевіряти кожен із пар [6, 15].

Впровадження будь-якого нового засобу оброблення даних вимагає від дослідників уваги в аспекті забезпечення стабільності для функціоналу класифікатора. 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 – підвищення продуктивності структурних методів класифікації зображень шляхом впровадження індексованих хеш-структур на множині дескрипторів бази еталонних образів та узгодженого ланцюжкового поєднання етапів аналізу.

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

Класифікація за множиною компонентів опису

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 дескрипторів КТ, $s = \text{card } E_i$ – число дескрипторів у множині [1, 2]. Компоненти опису – це вектори $e_k \in B^n$, скінченна множина яких створює опис об'єкту.

Процес класифікації передбачає наявність деякої бази E описів еталонів розмірністю

$N : E = \bigcup_{i=1}^N E_i$. Кожний еталонний опис E_i репрезентує окремий клас, представлений скінченною множиною дескрипторів КТ.

Опис розпізнаваного об'єкту є множина $Z \subset B^n$, $Z = \{z_w\}_{w=1}^s$. Поставимо задачу побудови класифікатора K як відображення $K : Z \rightarrow [1, 2, \dots, N]$ у множині номерів класів на основі попереднього конструювання індексованої структури на множині E .

Класифікацію K за множиною компонентів представимо як двоетапний процес $K = K_2 K_1$, 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 висновок про клас об'єкту Z . Такий принцип оброблення узагальнює множину рішень однорідних локальних класифікаторів за моделлю бустінгу [2, 14]. Етап K_1 тут можна розглядати як багатозначну характеристичну функцію для визначення еталонного класу E_i щодо окремого дескриптора із опису 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 [1, 2].

Конструктивно та інформативно 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 традиційним методом лінійного метричного пошуку (повний перебір) шляхом послідовного аналізу кожного елемента набору E , то застосовують конкурентне правило [2, 16, 17]

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

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 у випадку, якщо отриманий мінімум відстані в (1) не перевищує заданий поріг δ_ρ : $\rho_{\min} \leq \delta_\rho$. У протилежному випадку клас d_w не визначається, тобто аналізований дескриптор вважається хибним. Необхідну кількість Q обчислених 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 χ , що підраховує число не співпадаючих бітів для векторів [2]

$$\chi(z_w, e_v(i)) = \sum_{j=1}^n 1[z_{wj}, e_{vj}(i)], \quad (2)$$

де логічна функція $1[\dots]$ приймає значення 1, якщо j -ті біти векторів z_{wj} , $e_{vj}(i)$ не співпадають, і значення 0 – у протилежному випадку.

Етап K_2 полягає у тому, що на кожному кроці за правилом (1) за значенням d_w інкрементується число r_i голосів елементів, віднесених до i -го класу:

$$\begin{cases} r_i + 1, & d_w = i; \\ r_i, & d_w \neq i, \end{cases} \quad (3)$$

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

$$i_0 = \arg \max_{i=1, \dots, N} r_i. \quad (4)$$

Значення $\{r_i\}_{i=1}^N$ відображають гістограму класів за числом голосів елементів із Z . Вирази (3), (4) 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, 18].

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 підходом ближнього сусіда. Зосередимося на її вирішенні шляхом застосування у двоетапній процедурі класифікації спеціалізованої структури даних, що націлена на зменшення обсягу обчислень Q .

Індексована структура і трансформація даних для 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 $T_k(E)$:

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

$$T_k(E) \cap T_j(E) = \emptyset. \quad (5)$$

На сьогодні існує два найбільш поширених способи розбиття: хешування та кластеризація [3, 5, 6, 18]. Впровадження хешування пов'язують із точними засобами оброблення, а кластеризацію – із наближеними способами самонавчання. Трансформація (5) здійснює попередню класифікацію і повністю зберігає усю сукупність аналізованих даних, яка тепер розподіляється між групами $T_k(E)$. У результаті перетворення (5) 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

$$t_{i,k} = \text{card}\{e_v \mid e_v \in E_i \ \& \ e_v \in T_k\} \quad (6)$$

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 кожної корзини даних

$$b_i(T_k) = \frac{t_{i,k}}{\text{card } T_k}. \quad (7)$$

Вираз (7) 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 використання вагового впливу коефіцієнтів b_i для найбільш значущих класів у кошику [2].

Для хешування ознакою окремого кошику даних виступає значення хеш-функції. 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 [6, 9, 10, 14].

Якщо в одному кошику в середньому число елементів є $N \cdot s / M$, 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 $s = 500$, $N = 10$, $M = 10$ виграш $\beta = Q_1/Q_2$ складає приблизно 450 разів і зростає зі збільшенням обсягів N, s даних.

Метод оброблення даних з використанням хешування [5, 6] фактично здійснює попередню класифікацію на окремі кошики та має свої особливості.

Число кошиків для цього методу, як правило, задане і пов'язане із діапазоном значень хеш-ключа $h(e_v)$ для хешування, яке задає користувач.

Ключ – це, як правило, ціле число, що приймає множину із L фіксованих значень $h(e_v)$, $h(e_v) \in H_0, H_0 = \{h_i\}_{i=1}^L$ – множина значень хеш-функції [5, 7]. Для векторів із B^n таким ключем може бути, наприклад, число одиничних бітів у складі дескриптора, число вибраних комбінацій бітів довільної довжини та ін. Взагалі ознаку $h(e_v)$ для хешування можна корисно адаптувати до значень описів наявної бази еталонів.

Із впровадженням хешування розбиваємо множину даних E на непересічні групи $H_k(E)$ (кошики) числом L з рівнозначним значенням хеш-ключа

$$E = H(E) = \bigcup_{k=1}^L H_k(E),$$

$$H_i(E) \cap H_j(E) = \emptyset,$$

$$H_k(E) = \{e_v \in E \mid h(e_v) \in h_k\}. \quad (8)$$

Впровадимо тепер конкретний ланцюжок засобів попереднього оброблення даних задля забезпечення прискореного пошуку. Застосуємо блокове подання даних множини дескрипторів, обчисливши два значення хеш-функції $h(1), h(2)$ для кожного із дескрипторів бази еталонів. Конкретно для дескриптора BRISK розміром 512 бітів обчислимо значення $h(1), h(2)$ для першої та другої його половини за схемою «256+256».

Далі впровадимо LSH-оброблення [6]: відсортуємо еталонний масив $E = \{e_v\}_{v=1}^{s \cdot N}$ дескрипторів усієї бази за значенням хеш-функції $h(1)$ для 1-го блоку.

Класифікація на основі синтезованої структури

Застосуємо детектор КТ до зображення розпізнаного 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.

Пропонований спосіб із впровадженням ланцюжка засобів включає такі дії.

1) Для будь-якого дескриптора $z^* \in Z$ вхідного образу обчислюємо значення хеш-ключів $h^*(1), h^*(2)$ як значення хеш-функції для аргументу z^* .

2) При визначенні у базі дескриптора, ідентичного аналізованому $z^* \in Z$ зі значеннями $h^*(1), h^*(2)$, застосуємо лінійний пошук у «полосі даних» за значенням хеш-функції $h(1)$ для 1-го блоку, що обмежується заданою границею $h^*(1) \pm \Delta$. Результат визначається введеним параметром Δ , який можна назвати точністю пошуку. Крім того, параметр Δ забезпечує ідентифікацію варіативних даних, що незначно відрізняються від еталонних.

3) Логічне узгоджене оброблення. Здійснюємо зіставлення $\rho(z^*, e_v)$ за метрикою Хемінга (2) і пошук мінімуму

$$b = \arg \min_v \rho(z^*, e_v)$$

тільки у випадку, якщо значення хеш-функції для 2-го блоку знаходиться у межах $h^*(2) \pm \Delta$.

4) За результатом пошуку на етапі 3 для елемента z^* визначаємо клас d^* .

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

6) За виразом (4) класифікуємо розпізнаваний об'єкт до класу, що набрав найбільшу кількість голосів компонентів.

Запропонований продуктивний за швидкістю спосіб класифікації на підґрунті впровадження хешування може реалізувати як точний, так і наближений види пошуку.

Параметрами є вид хеш-функції, точність Δ , кількість блоків, число кошиків хешу, побудованих для компонентів даних в індексованій структурі. Цими параметрами можна керувати, виходячи із наповненості кошиків для конкретних даних.

Розподіли (6), (7) 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 значень для окремих класів).

Результати комп'ютерного моделювання

Моделювання виконано у середовищі PyCharm Community Edition 2020.2.3 із використанням засобів бібліотеки Open CV та мови програмування Python. Апаратне забезпечення – ноутбук з процесором Intel Pentium CPU N3540 2.16 ГГц та ОЗП 4 ГБ. Для визначення дескрипторів КТ застосовано детектор ключових точок BRISK [19] з дескриптором розмірністю $n = 512$. Розроблені моделі класифікатора вжито на

прикладі зображень шаржів людей (політик, артист, вчений), 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 з координатами КТ
(Fig. 1. Reference images with coordinates of key points)

У якості хеш-функції $h(\dots)$ взята кількість одиниць у двійковому коді вектору, параметр точності $\Delta = 13$ для обох блоків використано як $\pm 5\%$ від розміру блока 256, що відображає допустиме відхилення для значення хеш-функції.

Зважаючи на бінарний вид аналізованих даних, для зіставлення векторів застосовано метрику Хемінга. Всередині полоси фіксованих значень хешу у кошиках у експерименті застосовано традиційний лінійний пошук.

З метою зіставлення характеристик швидкодії програмно змодельований метод лінійного пошуку на усій множині еталонних даних (1500 дескрипторів) без застосування індексованих структур. Для множини еталонних зображень із застосуванням індексованих хеш-структур на 256 кошиків та для лінійного пошуку отримано максимально точний результат: усі 500 дескрипторів кожного із еталонів правильно віднесено до свого класу. При цьому час оброблення без індексного хешування склав приблизно 500 сек, а із комбінаційним хешуванням – лише 49 сек.

Для пошуку полоси даних у відсортованому масиві за значенням $h(1)$ можна замість лінійного пошуку запровадити, наприклад, метод дихотомії [7, 13], що зробить виграш ще суттєвішим.

Як бачимо, виграш у часі обчислень для запропонованого підходу (при заданому числі 3 еталонів та $s=500$) складає більше 10 разів у порівнянні із традиційним.

При цьому ясно, що виграш у часі обчислень збільшується при зростанні числа еталонів та дескрипторів у описах. Крім того, час обчислень суттєво залежить від створеної програмної моделі, типу комп'ютера, способу доступу до використаного програмного забезпечення, тому об'єктивними є лише відносні показники.

Хешування належить до точних методів трансформації даних.

Природно зрозуміло, що похибка класифікації із впровадженням хешування може виникнути через те, що всередині корзини згрупованих даних можуть зустрітися рівноцінні або близькі за хеш-функцією 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 $\delta_p \in [0, \dots, 256]$.

Експерименти з впливом перетворення повороту (найбільшим чином викривляє зображення, рис. 2) для вхідних еталонних зображень показали, що пропонувані метод успішно класифікує усі еталонні зображення.

Гістограми голосів для зображення рис. 2 містить значення 415, 52, 33, (для інших двох трансформованих зображень – (104, 330, 66) та (60, 59, 381)), що підтверджує впевнену класифікацію для кожного із класів. Максимум числа голосів за «правильний» еталон більше ніж утричі перевищує найближче значення.



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

Задля більш детального аналізу кількісного складу кошиків за показниками (6), (7) проведено обчислення цих значень шляхом розбиття відсортованої множини даних на 8 рівноцінних блоків за значенням хеш-функції $h(1)$ у діапазоні 0...255.

Табл. 1 містить значення показника (6) для 8-ми рівнозначних діапазонів значень $h(1)$. Із табл. 1 видно, що задля подальшого скорочення обчислень на класифікацію для даного набору еталонів деякі діапазони (0...63, 192...255) можна взагалі виключити із аналізу, для діапазонів 64...95, 160...191 встановити значущу перевагу одного із класів, а для решти діапазонів застосовувати розроблений метод. Це дає можливість додатково скоротити час класифікації.

Таблиця 1 – Числові показники класів для значень хеш-функції $h(t)$

Класи	Діапазони значень $h(t)$							
	31	63	95	127	159	191	223	255
1	0	0	43	140	224	92	1	0
2	0	0	25	143	257	72	3	0
3	0	0	6	111	254	122	7	0

У всіх проведених експериментах класифікація здійснювалася вірно, тобто усі вхідні зображення з перетвореннями правильно отримали свій номер класу. Отримані результати підтверджують ефективність розробленого методу.

Таким чином, із введенням індексної структури та хешування у експерименті отримано суттєвий вииграш у часі класифікації більше ніж у 10 разів. Основним недоліком впровадження хешування традиційно вважається деяке зростання необхідних обсягів пам'яті комп'ютера [5-10].

Висновки

У запропонованому підході з метою скорочення обсягів метричного пошуку для класифікації застосовано поєднання ланцюжка методів: хешування, розбиття на блоки (проекування), LSH-оброблення у вигляді сортування за значенням хеш-функції, логічний аналіз (узгоджена фільтрація) значень хеш-функції для блоків. Переваги запропонованого методу – лінійний пошук тільки у обмеженій і параметрично керованій полосі, зниження розмірності даних, пристосування до особливостей прикладної задачі класифікації із врахуванням відхилення значень схожості у певних межах.

Впровадження додаткового аналізу еталонних даних на попередньому етапі оброблення сприяє подальшому скороченню часу класифікації.

Особливості класифікатора полягають у тому, що здійснюється не точний пошук, а з можливостями допустимого відхилення.

Розглянуті методи класифікації базуються на принципі «порівняння з еталоном» і можуть бути універсально застосовані до довільних векторів даних бітового виду.

При застосуванні хеш-структур ключовим моментом є вибір результативної та адаптованої до даних хеш-функції, яка здійснює часткову класифікацію, не змінюючи дані. Більш детальний аналіз значень хеш-функції для конкретних даних дає можливість щодо подальшого скорочення необхідних обчислювальних витрат.

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

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

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

СПИСОК ЛІТЕРАТУРИ

1. Daradkeh Y.I., Tvoroshenko I., Gorokhovatskiy V., Latiff, L.A., Ahmad N. Development of Effective Methods for Structural Image Recognition Using the Principles of Data Granulation and Apparatus of Fuzzy Logic, *IEEE Access*, 2021, 9, pp. 13417-13428, DOI: <http://dx.doi.org/10.1109/ACCESS.2021.3051625>.
2. 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>.
3. Svyrydov A., Kuchuk H., Tsiara 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>.
4. Гороховатський В. О., Пупченко Д. В., Солодченко К. Г. Аналіз властивостей, характеристик та результатів застосування новітніх детекторів для визначення особливих точок зображення. *Системи управління, навігації та зв'язку*. 2018. №1 (47). С. 93–98, DOI: <http://dx.doi.org/10.26906/SUNZ.2018.1.093>.
5. 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>.
6. Лесковец Юре, Раджараман Ананд, Ульман Джеффри Д. Анализ больших наборов данных. Москва : ДМК Пресс, 2016. 498 с.
7. Ахо А., Хопкрофт Д., Ульман Д. Структуры данных и алгоритмы. Москва : Изд. дом «Вильямс», 2003. 384 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 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>.
12. Manning C.D., Raghavan P., and Schütze H. Introduction to Information Retrieval. Cambridge : University Press, 2008. 528 p.
13. Кормен Т., Лейзерсон Ч., Ривест Р., Штайн К. Алгоритмы: построение и анализ. М.: ИД Вильямс, 2005. 1296 с.
14. Флах П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных. Москва : ДМК Пресс, 2015. 400 с.

15. Gionis A., Indyk P., Motwani R. Similarity search in high dimensions via hashing, *Proc. Intl. Conf. on Very Large Databases*. 1999. pp. 518–529.
16. Гороховатский В.А., Путятин 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.
17. Gorokhovatsky V.A., Putyatin Ye. P. Image Likelihood Measures of the Basis of the Set of Conformities. *Telecommunications and Radio Engineering*. 2009. 68 (9). P. 763–778.
18. Nong Ye. *Data Mining: Theories, Algorithms, and Examples* (1st. ed.). CRC Press, Inc., USA. 2013.
19. Leutenegger Stefan, Chli Margarita, Roland Y. Siegwart. BRISK: Binary Robust Invariant Scalable Keypoints. *Computer Vision (ICCV)*. 2011. pp. 2548 – 2555.
20. 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>.
21. 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.

REFERENCES

1. Daradkeh, Y.I., Tvoroshenko, I., Gorokhovatskiy, V., Latiff, L.A., and Ahmad, N. (2021), “Development of Effective Methods for Structural Image Recognition Using the Principles of Data Granulation and Apparatus of Fuzzy Logic”, *IEEE Access*, 9, pp. 13417-13428, DOI: <http://dx.doi.org/10.1109/ACCESS.2021.3051625>.
2. 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>.
3. 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>.
4. Gorokhovatskiy, V.O., Pupchenko, D.V. & Solodchenko, K.G. (2018), “Analysis of properties, characteristics and results of application of the newest detectors for definition of special points of the image”, *Control, navigation and communication systems*, No. 1 (47), pp. 93-98, DOI: <http://dx.doi.org/10.26906/SUNZ.2018.1.093>.
5. 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>.
6. Leskovets, Yure, Radzharaman, Anand and Ulman, Dzhaffry D. (2016), *Analyzing large datasets*, DMK Press, Moscow, 498 p.
7. Akho, A., Hopcroft, D. and Ulman, D. (2003), *Data structures and algorithms*, Williams, Moscow, 384 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, 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>.
12. Manning, C.D., Raghavan, P. and Schutze, H., (2008), *Introduction to Information Retrieval*, Cambridge, University Press, 528 p.
13. Cormen, T., Leizerson, Ch., Ryvest, R. and Shtain, K. (2005), *Algorithms: construction and analysis*, Publishing house "Williams", Moscow, 1296 p.
14. Flach, P. (2015), *Machine learning. The science and art of building algorithms that extract knowledge from data*, DMK Press, Moscow, 400 p.
15. Gionis, A., Indyk, P. and Motwani, R. (1999), “Similarity search in high dimensions via hashing”, *Proc. Intl. Conf. on Very Large Databases*, pp. 518–529.
16. 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.
17. Gorokhovatsky V.A. and Putyatin, Ye. P. (2009), “Image Likelihood Measures of the Basis of the Set of Conformities”, *Telecommunications and Radio Engineering*, 68 (9), pp. 763–778.
18. Nong, Ye. (2013), *Data Mining: Theories, Algorithms, and Examples*, CRC Press, Inc., USA.
19. Leutenegger, Stefan, Chli, Margarita and Roland Y., Siegwart (2011), “BRISK: Binary Robust Invariant Scalable Keypoints”, *Computer Vision (ICCV)*, pp. 2548 – 2555,
20. 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>.
21. 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 (CVPR)*, pp. 6818-6826.

Received (Надійшла) 28.06.2021

Accepted for publication (Прийнята до друку) 25.08.2021

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

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

Volodymyr Gorokhovatskyi – 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>.

Стяглик Наталя Іванівна – кандидат педагогічних наук, завідувач кафедри інформаційних технологій і математичного моделювання, Навчально-науковий інститут «Каразінський банківський інститут» Харківського національного університету ім. В.Н. Каразіна, Харків, Україна;

Natalia Stiahlyk – Candidate of Pedagogical Sciences, Head of Department of Information Technology and Mathematical Modeling, Educational and Research Institute “Karazin Banking Institute” of V. N. Karazin Kharkiv National University, Kharkiv, Ukraine;

e-mail: natastyaglick@gmail.com; ORCID ID: <http://orcid.org/0000-0001-5938-2013>.

Царевська Віталія Віталіївна – студентка кафедри інформаційних технологій і математичного моделювання, Навчально-науковий інститут «Каразінський банківський інститут» Харківського національного університету ім. В.Н. Каразіна, Харків, Україна;

Vytaliia Tsarevska – student of Department of Information Technology and Mathematical Modeling, Educational and Research Institute “Karazin Banking Institute” of V. N. Karazin Kharkiv National University, Kharkiv, Ukraine;

e-mail: vitaliya.tsarevska42@gmail.com, ORCID ID: <http://orcid.org/0000-0003-1865-016X>

Комбинационный метод ускоренного метрического поиска данных в задачах классификации изображений

В. А. Гороховатский, Н. И. Стяглик, В. В. Царевская

Аннотация. Предмет исследований статьи – методы классификации изображений на множестве дескрипторов ключевых точек в системах компьютерного зрения. **Целью** является повышение производительности структурных методов классификации путем внедрения индексированных хэш-структур на множестве дескрипторов базы эталонных образов и согласованного цепного сочетания нескольких этапов анализа данных в процессе классификации. **Применяемые методы:** детектор и дескрипторы BRISK, средства хэширования данных, методы поиска в больших массивах данных, метрический аппарат определения релевантности векторов, программное моделирование. **Полученные результаты:** разработан эффективный метод классификации изображений на основе внедрения скоростного поиска с использованием индексированных хэш-структур, что ускоряет вычисления в десятки раз; выигрыш во времени вычислений увеличивается при возрастании числа эталонов и дескрипторов в описаниях; особенность классификатора состоит в том, что осуществляется не точный поиск, а с учетом допустимого отклонения данных от эталона; экспериментально проверена результативность классификации, что указывает на работоспособность и эффективность предложенного метода. **Практическая значимость работы** – построение моделей классификации в трансформированном 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 BRISK; хеширование с учетом близости; хэш-корзина; быстродействие классификации.

Combination method of accelerated metric data search in image classification problems

Volodymyr Gorokhovatskyi, Natalia Stiahlyk, Vytaliia Tsarevska

Annotation. The subject of research of the paper is the methods of image classification on a set of key point descriptors in computer vision systems. **The goal** is to improve the performance of structural classification methods by introducing indexed hash structures on the set of the dataset reference images descriptors and a consistent chain combination of several stages of data analysis in the classification process. **Applied methods:** BRISK detector and descriptors, data hashing tools, search methods in large data arrays, metric models for the vector relevance estimation, software modeling. **The obtained results:** developed an effective method of image classification based on the introduction of high-speed search using indexed hash structures, that speeds up the calculation dozens of times; the gain in computing time increases with an increase of the number of reference images and descriptors in descriptions; the peculiarity of the classifier is that not an exact search is performed, but taking into account the permissible deviation of data from the reference; experimentally verified the effectiveness of the classification, which indicates the efficiency and effectiveness of the proposed method. **The practical significance of the work** is the construction of classification models in the transformed space of the hash data representation, the efficiency confirmation 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; BRISK descriptor; hashing based on proximity; hash; classification performance.