

В. О. Гороховатський¹, Н. І. Стяглик², Є. В. Мазур¹, А. Д. Вечірська¹

¹ Харківський національний університет радіоелектроніки, Харків, Україна

² ННІ Каразінський банківський інститут Харківського національного університету ім. В.Н. Каразіна, Харків, Україна

Способи метричної грануляції для опису зображень у задачі класифікації

Анотація. Предметом досліджень статті є методи класифікації зображень у системах комп'ютерного зору. Мета – удосконалення структурних методів класифікації задля підвищення їх швидкодії шляхом впровадження редукованої системи класифікаційних ознак на підрунті здійснення грануляції еталонних описів. Застосовані методи: детектори ключових точок *brisk* та *orb*, апарат теорії множин і векторних просторів, метричні моделі кластеризації і грануляції векторів, моделі «мішок слів» і голосування для класифікації, програмне моделювання. Отримані результати: розроблено моделі для трансформації опису зображень у вигляді набору кластерних центрів і побудови редукованого опису з використанням грануляції за подібністю векторів, досліджено швидкісні методи класифікації зображень на основі впровадження трансформованих описів. Результативність розроблених модифікацій класифікатора залежить від способу грануляції даних і від параметрів порогів на еквівалентність векторів, мінімум значення метрики і число голосів дескрипторів для класу – переможця у схемі голосування. Запропоновано побудову ієрархічної системи ознак з кількома рівнями грануляції. На підставі впровадження модифікацій вдалося в сотні разів скоротити обчислювальні витрати при забезпеченні результативності класифікації на навчальній вибірці даних. Практична значущість роботи полягає у побудові програмних моделей для здійснення швидкісної класифікації, підтвердженні працездатності і точності запропонованих модифікацій для прикладної бази зображень, створенні програмного застосування для впровадження модифікацій класифікаторів у комп'ютерному баченні.

Ключові слова: комп'ютерний зір; структурні методи класифікації; множина дескрипторів; кластеризація, грануляція векторів, точність класифікації.

Вступ. Аналіз літературних джерел

На сьогодні однією із найбільш вагомих проблем у системах комп'ютерного зору є забезпечення необхідних показників їх прикладного функціонування, зокрема, точності та швидкодії оброблення даних [1-4]. Особливо актуальною є вирішення цієї проблеми для задач класифікації візуальних об'єктів, де використовується значна кількість прототипів класів, образи яких подаються у формі множини векторів [2, 5]. Традиційно такі задачі відносяться до сфери *big data* і потребують здійснення об'ємного лінійного пошуку у багатовимірному просторі даних.

Зараз інтенсивно розвиваються структурні методи класифікації, де класифікаційне рішення базується на поданні зображення у формі множини дескрипторів ключових точок [2, 6]. Дескриптори ключових точок (КТ) формуються спеціальними фільтрами – детекторами [7, 8]. У структурних методах процес класифікації здійснюється за моделлю «мішок слів», а клас об'єкту визначається за результатом голосування складових компонентів – дескрипторів розпізнаваного об'єкту [2]. Такі методи мають переваги перед сучасними нейромережами, так як базуються на узгодженні з описом чи параметрами еталонів і не потребують довготривалого навчання. Крім того, при їх застосуванні є можливість оперативної зміни складу бази еталонів і прийняття рішення за окремими деталями візуального об'єкту. Функціонал апарату дескрипторів має необхідну в реальних умовах властивість інваріантності до геометричних перетворень об'єктів у полі зору.

Одним із способів зниження ресурсоемності при впровадженні структурних методів є застосування засобів грануляції даних із формуванням кластерів, хеш-кошиків чи оцінювання центрів [9]. Такі

способи можна вважати метричною грануляцією, так як вони базуються на оцінюванні близькості даних з використанням метрики або параметрів даних [10, 11]. За рахунок впровадження таких підходів суттєво покращується швидкодія оброблення з допустимим зниженням показника точності класифікації. Однак досягнення необхідних значень критерію результативності потребує більш глибокого вивчення усіх можливостей і нюансів прикладного застосування апарату грануляції з необхідністю підтвердження об'єктивності і універсальності цього апарату у сфері інтелектуального аналізу багатовимірних даних. Зокрема, є потреба у впровадженні більш простих у порівнянні з кластеризацією засобів формування результативних чи найбільш інформативних підмножин даних структурного опису [2, 11], а також використання роздільних схем для кластерного подання еталонів у створених модифікаціях швидкісних методів класифікації.

У статтях [3, 5] розглянуто три основні способи отримання редукованої системи ознак на прикладі метричного критерію інформативності для кожного елемента множини дескрипторів опису. Це такі, як:

- 1) відбір фіксованої кількості інформативних ознак;
- 2) визначення числа ознак із інформативністю вище заданого порогу;
- 3) формування інформативної підмножини ознак, що забезпечує необхідну результативність.

Розробляються також більш ресурсозатратні способи, наприклад *Relief* [12], де редукована множина формується випадковою процедурою аналізу на усій наявній множині даних з визначенням оптимальних відстаней до свого і чужого класів. Перспективним вважається також застосування оцінок інформативності на базі центрів даних [5, 13].

Кластерування є одним із найбільш ефективних прикладних підходів, що здійснює апроксимацію даних шляхом побудови деякої системи центрів кластерів. Швидкодія оброблення підвищується пропорційно зменшенню числа центрів. При цьому число центрів є параметром, яким можна керувати задля забезпечення необхідної результативності класифікації [10, 14].

Ефективна в обчислювальному плані ідея щодо побудови інтегральних ознак зображення реалізується пірамідальним поданням [16, 17]. Пірамідальна структура даних дає можливість суттєво скоротити багатомасштабний обсяг зорової інформації шляхом інтегрування яскравості для просторово близьких координат пікселів зображення. Такий підхід гранулювання доцільно поширити і на опис у формі множини дескрипторів. Загалом ідея грануляції даних набула популярності у системах розпізнавання образів [20]. Різні аспекти використання новітніх способів формування ефективного простору ознак у класифікаторах зображень обговорюються у роботах [22-25].

Зосередимося на вивченні і експериментальному впровадженні способів аналізу даних, що базуються на відношенні грануляції на множині дескрипторів. Для конкретизації розглянемо бінарне відношення для пар метрично найближчих елементів опису, в результаті грануляції у редукованому списку залишається тільки один із них.

Задачами дослідження є опрацювання новітніх модифікацій способів грануляції описів для впровадження у класифікаторах зображень, аналіз їх ефективності у порівняльному аспекті з традиційним методом. Увага зосереджена на апроксимації опису за рахунок скорочення кількості його інгредієнтів за принципом подібності, а також на роздільному способі кластерування для описів еталонів. Практичними моментами є вивчення результативності (швидкодії, точності) впровадження запропонованих модифікацій простору даних у методах класифікації шляхом проведення експериментального оцінювання за наслідками програмного моделювання для прикладної бази зображень.

Формалізація структурного методу

Будемо здійснювати класифікацію у рамках фіксованої бази із N еталонів (представники чи прототики класів) у формі деякої множини E описів еталонних зображень: $E = \{E_1, E_2, \dots, E_N\}$. Множина E – це фактично навчальна вибірка, яка одночасно є підґрунтям для побудови класифікатора на базі зіставлення з еталоном [2, 5]. Кожний еталонний опис E_k у формалізмі класифікатора репрезентує окремих клас. Клас k з описом E_k формально визначаємо як нескінченну множину зображень, отриманих із еталону (прототип класу з номером k) шляхом застосування до нього багатопараметричної групи геометричних перетворень, яка найчастіше у прикладних застосуваннях включає зміщення, поворот, масштабування, дія яких не виводять об'єкт інтересу із поля зору [18]. Сучасні детектори ключових точок, наприклад ORB, BRISK, AKAZE забезпечують інваріантність отриманого опису до такої групи перетворень [6–8].

Опис еталону $E_k = \{e_v(k)\}_{v=1}^s$, $E_k \subseteq V^n$, подаємо як скінченну множину потужністю в s дескрипторів КТ у просторі V^n бінарних векторів, $e_v(k) \in V^n$, $s = \text{card}E_k$ – число дескрипторів у множині. Усякий дескриптор $e_v(k)$ у складі бази E характеризується параметром k номеру класу, а загалом число ознак – дескрипторів у базовій множині E складає $\text{card}E = sN$.

Процес R структурної класифікації здійснимо у два етапи як $R = R_2 R_1$, де етап R_1 реалізує локальне рішення стосовно класу окремого дескриптора об'єкту, а етап R_2 на підставі рішень (голосів) компонентів усього складу опису отримує значення параметра класу для аналізованого об'єкту.

Класифікатор R_1 першого етапу реалізує по-елементний аналіз вхідного опису $Z = \{z_v\}_{v=1}^s$ об'єкта і відносить кожний дескриптор $z_v \in Z$ до одного із класів за правилом

$$R_1: z_v \rightarrow \{1, \dots, N\}. \quad (1)$$

Клас k аналізованого дескриптора z_v об'єкту у моделі (1) визначимо як аргумент мінімуму відстані на множині E дескрипторів усіх класів

$$R_1: k = \arg \min_{i=1, \dots, N; d=1, \dots, s} \rho(z_v, e_d(i)). \quad (2)$$

Тут $\rho: V^n \times V^n \rightarrow [0, \infty]$ – відстань у векторному просторі V^n .

У конкурентній моделі (2) доцільно використовувати відстань Хеммінга, що визначає число відмінних бітів у двох бінарних векторах [3, 4]. У виразі (2) мінімізується значення відстані за змінною i (номер класу) на множині дескрипторів бази E . У формулі (2) може бути організовано визначення мінімуму як за повним списком $s \cdot N$ дескрипторів бази, так і роздільно для кожного із N еталонів (з встановленим лімітом на величину мінімуму).

У роздільному способі аналізу для кожного із класів підраховується число дескрипторів об'єкту, що знайшли відповідність до фіксованого еталону. У такій моделі маємо множинне врахування голосів, коли один і той же дескриптор об'єкту може знайти відповідність для кількох еталонів різних класів.

Таким чином, модель (2) втілює багатокласову класифікацію для окремого дескриптора у векторному просторі V^n на підставі його найкращого узгодження з множинами еталонних дескрипторів.

Для побудови R_2 введемо вектор $\{h_i\}_{i=1}^N$ з цілими значеннями для накопиченої кількості голосів класів, отриманих застосуванням моделі (2) до усієї множини дескрипторів об'єкту Z . На підставі впровадження R_1 для кожного $z_v \in Z$ відповідно до (2) визначимо номер класу k , а потім інкрементуємо акумулятор голосів $h_k = h_k + 1$ для відповідного номеру класу. Фактично отриманий вектор $\{h_i\}_{i=1}^N$ є розподілом (гістограмою) числа голосів для множини дескрипторів об'єкту у системі класів.

За результатом оброблення опису Z об'єкту обчислюємо вектор $\{h_i\}_{i=1}^N$. Клас об'єкту визначимо правилом R_2 як аргумент максимуму

$$R_2: Z \rightarrow E_k | (k = \arg \max_{i=1, \dots, N} h_i) \ \& \ (h_k \geq \delta_n), \quad (3)$$

де δ_h – деякий поріг для мінімально допустимого числа голосів. Якщо умова $h_k \geq \delta_h$ не виконується, клас об'єкту не встановлюється (відмова від класифікації, немає підстави віднести об'єкт до жодного із класів).

Значення δ_h визначають експериментально для заданої бази E еталонів. Як правило, його обирають як мінімальне число голосів (з допуском), необхідне для впевненої класифікації тестової вибірки на основі еталонів. Із загальних теоретичних позицій δ_h є компромісом у задачі розрізнення еталонів від решти зображень, що можуть поступити на вхід системи класифікації [5, 9].

Послідовність правил R_1, R_2 реалізує класифікатор на підґрунті рішень набору класифікаторів, отриманих для множини складових компонентів об'єкту. Він впроваджує принцип структурного аналізу і забезпечує стійкість до просторових викривлень окремих компонентів (дескрипторів) із-за можливого впливу завод.

Модифікації простору ознак впровадженням грануляції даних

Здійснимо кластеризацію множини даних у базі E еталонних описів. Процедура кластеризації відносять до апарату грануляції даних з визначенням системи центрів. Будемо розглядати два альтернативних підходи до кластеризації з однаковим числом M кластерів – сукупно для усієї множини E (система центрів W^1) і роздільно для кожного із еталонів E_i , $i = 1, \dots, N$ (система центрів W^2). У результаті маємо два варіанти центрів

$$W^1 = \{w_j^1\}_{j=1}^M, \quad W^2 = \{w_j^2\}_{j=1}^M. \quad (4)$$

Однакове число кластерів у (4) обрано нами виключно задля можливості порівняння результативності роздільного і спільного кластерного подання для еталонних описів. Фактично центри (4) є апроксимацією множини E , вони створюють нові простори систем класифікаційних ознак.

При цьому для отримання векторного подання кожного еталону за побудованою системою центрів у першому варіанті (спільна кластеризація для множини E) треба додатково на підготовчому етапі виконати конкурентний аналіз опису для кожного із еталонів E_k . Для цього для усіх елементів $z \in E_k$ визначимо номер v найближчого центру у множині W^1 , $w_j \in W^1$

$$v = \arg \min_{j=1, \dots, M} \rho(w_j, z). \quad (5)$$

Схема (5) є втіленням моделі (2) з пошуком найближчого кластерного центру. У результаті застосування (5) отримаємо подання для еталону E_k у формі цілочислового вектору $g(E_k) = [g_1, g_2, \dots, g_M]$ з кількістю компонентів, що дорівнює числу кластерів. Тут g_d – число дескрипторів опису, що потрапили у кластер з номером d . Вектор g є розподілом даних опису за набором кластерів на відміну від $\{h_i\}_{i=1}^N$ – розподілу за системою класів. У процесі класифікації для аналізованого зображення з описом Z аналогічно

також треба отримати вектор $g(Z)$ кластерного подання.

Для системи центрів W^2 вже маємо закріплення фіксованого рівноцінного числа m центрів за кожним із еталонів, причому $M = mN$, так як їх кластеризація виконувалась роздільно. Тому додаткових дій з аналізу чи трансформації еталонного представлення даних не потрібно.

Загалом число центрів m є параметром, що залежить від складу даних.

Наряду з отриманням центрів на підґрунті здійснення кластеризації розглянемо процедуру гранулювання для складу еталонних описів на підставі критерію подібності дескрипторів. Схожі трансформації застосовуються у пірамідальному поданні зображень, де впроваджено інтегроване подання у формі ієрархічної структури [16]. Близьке перетворення використано нами при трансформації дескрипторів опису шляхом складання значень бітів в рамках системи фрагментів [2].

Визначимо для опису еталону E_k потужністю в s дескрипторів трансформацію

$$T(E_k) \rightarrow E_k^*, \quad \text{card}(E_k^*) \ll s, \quad (6)$$

результатом якої є відфільтрований опис E_k^* , отриманий відбором підмножини $E_k^* \subset E_k$ компонентів із E_k з використанням деякого критерію грануляції. У прикладному аспекті пропонується орієнтовне зниження потужності опису у два рази на одному етапі оброблення.

Принцип аналізу метричної близькості дескрипторів може бути основою для побудови різноманіття варіантів фільтрації (редукції) даних. Розглянемо конкретні варіанти побудови T на основі здійснення грануляції Ω даних на підставі значення метрики ρ для дескрипторів.

1. Грануляція за порогом еквівалентності. Визначимо для кожного із еталонних дескрипторів $e_d(k) \in E_k$ найближчий до нього за метрикою ρ дескриптор $e_*(k)$ всередині множини E_k , що відповідає моделі лінійного метричного пошуку

$$e_*(k) = \arg \min_{v \neq d} \rho(e_v(k), e_d(k)). \quad (7)$$

Після визначення дескриптора з мінімальним значенням метрики перевіримо умову еквівалентності $\rho(e_d(k), e_*(k)) \leq \delta_\rho$, δ_ρ – заданий граничний параметр для значення метрики, який наближено встановлює еквівалентність двох дескрипторів. У просторі багатовимірних векторів величину δ_ρ часто визначають як 25% від максимуму метрики [4].

Фактично тут грануляція реалізована як функція Ω двох аргументів $\Omega[e_a(k), e_b(k)] \rightarrow e_*(k)$, $a \neq b$, де значення $e_*(k) \in E_k^*$ – це один із аргументів, які визначено метрично еквівалентними між собою. Функція Ω повертає значення одного із аргументів, якщо вони визнані рівнозначними на підставі істинності побудованого бінарного відношення еквівалентності. Якщо ж еквівалентність за порогом δ_ρ не встановлено, то перший аргумент заносимо до E_k^* , виключаємо цей елемент із оброблення, а другий продовжує брати участь у аналізі.

Результат такого способу грануляції залежить від параметра порогу δ_ρ , а також від порядку слідування дескрипторів у описі.

2. Грануляція з відсіюванням за найменшою відстанню. Ще один спосіб можна збудувати на підґрунті аналізу повної матриці відстаней між дескрипторами опису E_k [19]. Застосуємо відбір необхідної кількості дескрипторів шляхом відсіювання, починаючи із значення найменшої відстані. Список E_k^* поповнюємо черговим дескриптором, виключаємо його відповідні рядок і стовпець із матриці, і переходимо до аналізу наступного дескриптора з більшим значенням відстані. Таким чином послідовно набираємо потрібну фіксовану кількість дескрипторів у множину E_k^* . Ця процедура дещо складніша за попередню, може містити сортування за відстанню, зате вона не залежить від порядку слідування дескрипторів і не потребує адаптації порогу δ_ρ до даних.

Зрозуміло, що обидва розглянуті способи грануляції можна об'єднати у єдиному більш складному різновиді аналізу.

Наприклад, другий спосіб може бути доповнено використанням порогу еквівалентності.

Запропоновані способи аналізу і трансформації даних, крім результативного скорочення складу опису еталону, створюють основу для побудови ієрархії ознак, так як перетворення (6) можна застосувати повторно до уже трансформованого опису E_k^* .

За результатом здійснення етапу грануляції елементам сформованого редукованого опису E_k^* можна призначити деякі вагові коефіцієнти, які можуть бути враховані під час класифікації. Наприклад, новоствореному елементу, який отримано грануляцією значень двох близьких дескрипторів, можна присвоїти коефіцієнт 2.

А елементу, що не знайшов найближчого із заданим обмеженням – залишити вагу 1.

Класифікація у трансформованому просторі

Традиційний метод класифікації реалізує послідовність моделей (2), (3).

Класифікатор з використанням спільної кластеризації на основі системи центрів W^1 спочатку отримує за моделлю (5) векторне подання

$$g(Z) = [g_1, g_2, \dots, g_M]$$

для розпізнаваного об'єкту, а далі з використанням метрики μ для векторів цілих чисел визначає клас об'єкту як аргумент мінімуму метрики μ на множині кластерних подань для еталонів бази E

$$R_2: Z \rightarrow E_v \mid v = \arg \min_{k=1, \dots, N} \mu(g(Z), g(E_k)). \quad (8)$$

Варіантом метрики μ в (8) може бути манхетенська відстань для простору векторів цілих чисел. Використання правила (8) замість правил (2), (3) для традиційного методу дає вигреш у обсязі обчислень приблизно пропорційний величині $\gamma = s/M$.

Відмітимо, що для відсіювання зображень поза базою оптимальне значення метрики μ повинне задовольняти умов $\mu_v \geq \delta_\mu$.

Класифікатор на підставі роздільної кластеризації на етапі R_1 втілює модель (5) на множині центрів $w_j^2 \in W^2$. Клас дескриптора при цьому однозначно визначається отриманим центром кластеру, а клас усього об'єкту оцінюється за моделлю (3). Вигреш у обсязі обчислень теж пропорційний значенню γ .

На підставі отриманих редукованих множин $\{E_k^*\}_{k=1}^N$ впровадимо класифікатор з використанням грануляції, що реалізує традиційну модель (2), (3). У той же час і до множин $\{E_k^*\}_{k=1}^N$ теж можна застосувати обговорювані схеми кластеризації, що призводить до ієрархічної структури даних опису.

При дії локальних завад класифікація у розглянутих методах фактично здійснюється на підставі невикривленої частини опису аналізованого об'єкту. У цьому випадку число голосів природно знижується, тому параметри δ_ρ , δ_n треба адаптувати до допустимого рівня перешкод.

Зважаючи на те, що при способі роздільної кластеризації апроксимація здійснюється шляхом формування центрів для еталонних даних, критерієм, що впливає і на результативність класифікації, можна прийняти середню відстань між системами центрів для пар еталонів. Якщо позначити вектори для центрів пари еталонів $A = \{a_i\}_{i=1}^m$ та $B = \{b_j\}_{j=1}^m$, то такий критерій β має вигляд

$$\beta(A, B) = [\sum_{i=1}^m \sum_{j=1}^m \rho(a_i, b_j)]/m^2. \quad (9)$$

Значення (9) знаходиться у межах визначення метрики ρ . Для бінарних векторів розмірністю n і метрики Хемінга показник (9) належить діапазону $[0, n]$. Чим більше значення (9), тим маємо більший розкид між системами сформованих центрів і тим вище ступінь розрізнення класів A, B між собою.

Результативність класифікації будемо оцінювати критерієм точності pr , який відображає відношення кількості q даних з правильною оцінкою класу до загального числа Q експериментів

$$pr = q/Q. \quad (10)$$

Експериментальні результати і обговорення

Основною метою проведення експериментів було підтвердження працездатності і результативності запропонованих способів грануляції у задачі класифікації в порівняльному аспекті з традиційним методом.

При програмному моделюванні класифікатора з використанням двох способів кластеризації використано базу із п'яти зображень картин української художниці М. Приймаченко [21], а також три її інші картини поза базою класифікації.

На рис. 1 наведено дві картини із використаної бази і координати КТ дескрипторів, отриманих детектором brisk (вектор із 512 бітів) [8, 26].

У решті експериментів використано детектор orb (вектор із 256 бітів) [7].

Як бачимо, художниця використовує оригінальну техніку для своїх картин, що в той же час робить їх візуально схожими.

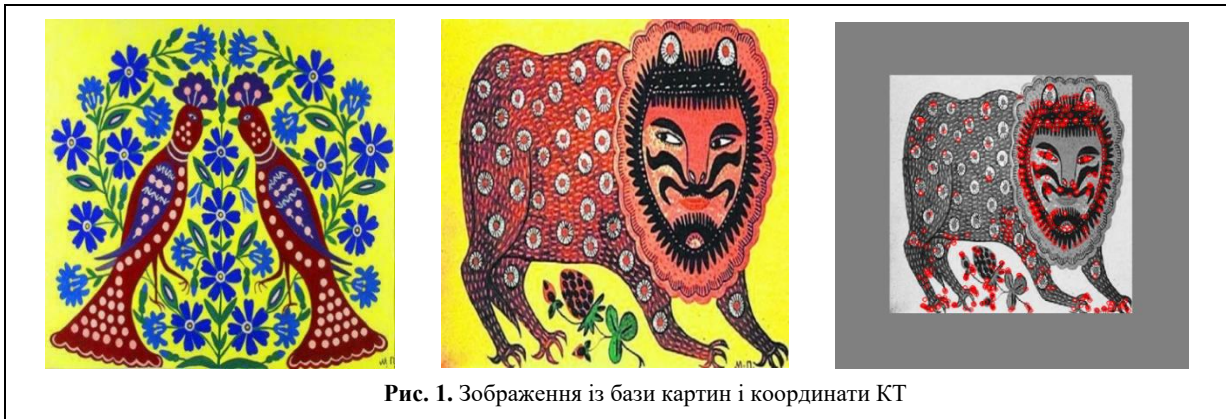


Рис. 1. Зображення із бази картин і координати КТ

Тестовий набір у експерименті склав 80 зображень і включав ці 8 картин та зображення, отримані із них геометричними перетвореннями зсуву (± 10 пікселів), повороту ($\pm 10\%$) та масштабування ($\pm 10\%$).

На рис. 2 показано тестове зображення і його координати КТ.

Для кластеризації використано $M = 15$ кластерів для бази, і $m = 3$ для кожного із еталонів.

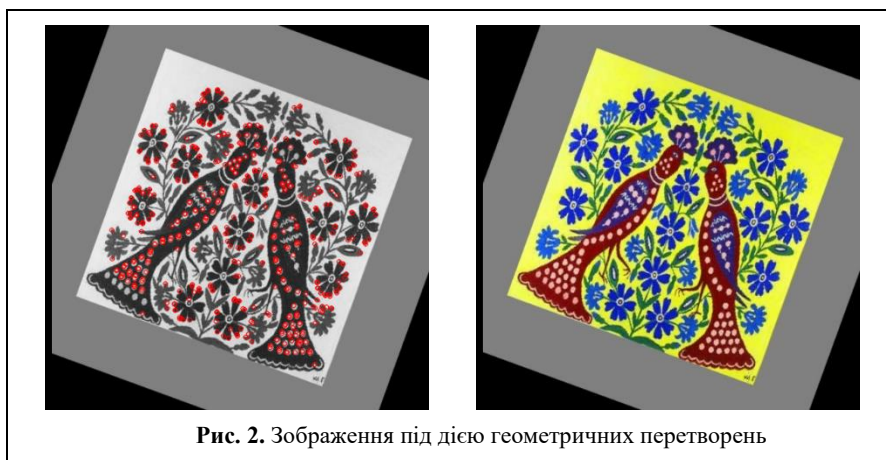


Рис. 2. Зображення під дією геометричних перетворень

Проведено попередні експерименти, щоб визначити для наявної бази зображень параметри порогів δ_p , δ_n для відстані Хемінга між дескрипторами та центрами і необхідного для прийняття рішення числа голосів дескрипторів.

Зважаючи на те, що результативність безпосередньо залежить від процедури кластерування, у обраній схемі кластеризації k -середніх початкові дані для центрів кластерів визначалися випадковим чином із складу опису. При цьому застосована модифікована процедура обчислення поточних центрів шляхом усереднення числа бітів з формуванням бінарного подання для новоствореного центру [10, 19].

У результаті проведених експериментів визначено параметри $\delta_p = 192$ для методів з кластеруванням, $\delta_n = 300$ для традиційного методу, $\delta_\mu = 180$ для методу з кластеруванням повної бази і $\delta_n = 140$ для методу з роздільним кластеруванням еталонів.

Результати експериментів засвідчили таке. Традиційний метод показав найвищу точність $pr = 1$, успішно класифікуючи усі зображення із бази і відділивши зображення не із бази. Модифікований метод з кластеризацією повної бази показав точність $pr = 0,82$, а з роздільною – $pr = 0,90$. Як бачимо, роздільна кластеризація дає дещо кращий рівень точності. При застосуванні числа кластерів $m = 5$

для кожного еталону точність методу з кластеризацією повної бази зросла до $pr = 0,88$, а з роздільною кластеризацією – до $pr = 0,98$.

Заміри часу класифікації одного зображення: традиційний метод – 6500 мс, для $m = 3$ із спільною кластеризацією – 41 мс, із роздільною – 42 мс.

Як бачимо, введення кластерування даних у процес класифікації суттєво знизило обчислювальні витрати (у нашому експерименті – у 154 рази!) за рахунок деякого зниження показника точності.

За моделлю (9) середні відстані між системами центрів першого і другого еталонів склали для трьох кластерів – 168,2, для п'яти – 182,8 (максимум відстані – 512). Як бачимо, відстань зростає зі збільшенням числа кластерів.

Зауважимо, що отримані показники точності безпосередньо залежать від обраної бази зображень, а також від набору альтернативних зображень не із бази, і можуть змінюватися у будь-яку сторону. Ключовим фактором є візуальна схожість аналізованих зображень між собою, що відображається і у близькості дескрипторів. У той же час значний вигравш у часі оброблення при застосуванні кластерування є досить об'єктивним фактором.

Моделювання за схемою грануляції (6) здійснювалося з використанням зображень птахів за анало-

гічною схемою із 80 зображень (5 із бази та 3 поза базою) і дією фіксованого набору геометричних перетворень (рис. 3). Експериментально вивчено обидва запропоновані способи грануляції – за порогом еквівалентності і з відсіюванням за найменшою відстанню. Значення порогу для визначення еквівалент-

ності для пари дескрипторів обрано як $\delta_p = 64$ як 25% від максимальної відстані 256.

Після здійснення грануляції на редукованому описі еталонів застосовано традиційний метод класифікації (2), (3) з підрахунком голосів для дескрипторів вхідного об'єкту.



Рис. 3. Зображення із бази птахів і координати ключових точок

Для підтвердження ефективності введення грануляції взято простий спосіб скорочення складу еталонних дескрипторів у два рази шляхом відсіювання через один без застосування апарату грануляції.

Встановлено, що точність класифікації для обраної бази зображень залишилася на найвищому рівні $rg = 1$ як для повного складу дескрипторів опису еталонів, так і для редукованого складу з грануляцією і без неї. При цьому число голосів для перетворених еталонів без застосування грануляції склало 340-440 із 500 і є приблизно на 5-10% нижче, ніж із використанням грануляції. Як бачимо, у нашому експерименті грануляція на базі зображень птахів дає дещо кращу результативність, ніж кластерне подання для бази картин. Проведено пробні експерименти із повторним впровадженням грануляції до уже гранульованого опису, що здійснює перехід на наступний рівень ієрархічного подання даних. Число дескрипторів із використанням адаптованого до бази порогу δ_p еквівалентності у 37% (значення 96) знизилось поетапно від 500 до 250...255, і далі до 140...160, що призвело до виграшу у швидкодії оброблення приблизно у три рази.

Усі еталонні зображення на двох рівнях гранульованого подання класифікуються правильно. У той же час максимальне число голосів для еталонних вхідних зображень з повним складом опису у 500 дескрипторів знижується на першому рівні до 410...450, а на другому рівні – до 300...400. Це вказує на перспективу впровадження наступних рівнів грануляції з ще більшим ступенем виграшу у швидкодії.

Висновки

Впровадження засобів грануляції даних у структурних методах класифікації зображень показало їх працездатність і достойний рівень результативності

із можливістю суттєвого прискорення часу оброблення. Обчислювальні затрати на класифікацію знижуються пропорційно ступеню скорочення опису еталонів і числа кластерів. Результативність класифікації за критерієм точності залежить від встановлених експериментально значень параметрів порогів та способу грануляції, обраного для конкретної бази розпізнаваних зображень.

Якщо порівнювати ефективність способів із сукупною і роздільною кластеризацією для описів бази еталонів, то із застосуванням кластерування к-середніх метод з роздільною кластеризацією показав вищий рівень точності класифікації. За обсягом часових затрат при однаковому числі кластерів ці способи приблизно рівноцінні. Обидва застосовують по-елементний метричний аналіз опису вхідного зображення, що забезпечує відсіювання завад. Рівень апроксимації для обох способів може бути покращено збільшенням числа кластерів, але за рахунок зростання обсягу обчислень.

Відпрацьована нами на практиці технологія визначення порогів, що використовуються у класифікаторі за множиною дескрипторів ключових точок, дає можливість універсально застосувати запропоновані методи для довільної бази зображень.

Перспективою дослідження може бути прикладне застосування розроблених способів для об'ємних колекцій зображень. Ще одним напрямком є вивчення ефективності щодо більш високих рівнів гранульованого подання описів.

Результати даного дослідження отримано за міжнародним науковим проектом INITIATE у рамках гранту № 101136775-HORIZON-WIDERA-2023-ACCESS-03.

СПИСОК ЛІТЕРАТУРИ

1. Tymchyshyn R., Volkov O., et al. (2018), Modern Approaches to Computer Vision, Control systems and computers, 6, 46-73.
2. Daradkeh Y. I., Gorokhovatskyi V., Tvoroshenko I., Gadetska S., and Al-Dhaifallah M. (2023), Statistical data analysis models for determining the relevance of structural image descriptions, IEEE Access, vol. 11, pp. 126938–126949, Nov. 2023, doi: 10.1109/ACCESS.2023.3332291.
3. Gorokhovatskyi V. Vlasenko N. (2021), The image description reduction in the set of descriptors on informativeness metric criteria base. Advanced Information Systems, 5 (4), 10–16. DOI: <https://doi.org/10.20998/2522-9052.2021.4.02>

4. P. Flach. (2012), Machine learning. The Art and Science of Algorithms that Make Sense of Data. New York, NY, USA: Cambridge University Press.
5. Gorokhovatskyi V. A. (2018), Image classification methods in the space of descriptions in the form of a set of the key point descriptors, Telecommunications and Radio Engineering, 77(9), 787-797.
6. Kuchuk H., Kovalenko A., Ibrahim B.F., Ruban I. (2019), Adaptive compression method for video information. International Journal of Advanced Trends in Computer Science and Engineering. Vol. 8 (1). P. 66–69. DOI: <http://dx.doi.org/10.30534/ijatcse/2019/1181.22019>
7. Rublee E., Rabaud V., Konolige K., and Bradski G., ORB: An efficient alternative to SIFT or SURF, in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 2564–2571, doi: 10.1109/ICCV.2011.6126544.
8. Leutenegger S., Chli M. and Siegwart R. Y. BRISK: Binary robust invariant scalable keypoints, in 2011 Int. Conf. on Computer Vision, Barcelona, Spain, pp. 2548–2555, 2011.
9. Gorokhovatskyi V., Tvoroshenko I. (2023), Identification of visual objects by the search request. International scientific symposium «Intelligent Solutions-S». Computational intelligence. Decision making theory: proc. of the international symposium, September 28, 2023, Kyiv-Uzhorod, Ukraine, pp. 25-27.
10. Gorokhovatskyi O., Gorokhovatskyi V. and Peredrii O. (2018), Analysis of application of cluster descriptions in space of characteristic image features, Data, vol. 3, no. 4, pp. 52.
11. Gorokhovatsky V., Putyatin Y. and Stolyarov V. (2017), Research of Effectiveness of Structural Image Classification Methods using Cluster Data Model, Radio Electronics, Computer Science, Control, vol. 3 (42), pp. 78–85.
12. Robnik-Sikonja, M., Kononenko, I. (2003), Theoretical and empirical analysis of ReliefF and RReliefF. Machine Learning 53 (1-2): 23-69.
13. Stańczyk U. (2015), Feature Evaluation by Filter, Wrapper, and Embedded Approaches. In: Stańczyk U., Jain L. (eds) Feature Selection for Data and Pattern Recognition. Studies in Computational Intelligence, Springer, Berlin, Heidelberg, vol. 584, 568 p.
14. Xiong, H. and Z. Li. (2014), Data Clustering: Algorithms and Application. 1st ed. Boca Raton: CRC Press.
15. Zhang X., Yu F. X., Karaman S. and Chang S.-F. (2017), Learning discriminative and transformation covariant local feature detectors,” in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 4923–4931.
16. Lowe D. G. (2004), Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision. 60 (2). [doi:10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94)
17. Crowley, J., Riff O. (2003), Fast computation of scale normalized Gaussian receptive fields, Proc. Scale-Space'03, Isle of Skye, Scotland, Springer Lecture Notes in Computer Science, 2695.
18. Putyanin, E.P., Averin, S.I. (1990), Image processing in robotics. Moscow, Mashinostroeniye, 320 p.
19. Gorokhovatskyi V., Tvoroshenko I., and Yakovleva O. (2024), Transforming image descriptions as a set of descriptors to construct classification features, Indonesian Journal of Electrical Engineering and Computer Science, vol. 33, no. 1, pp. 113-125, doi: 10.11591/ijeecs.v33.i1.
20. Martino A., De Santis E., Rizzi A. (2023), On Information Granulation via Data Filtering for Granular Computing-Based Pattern Recognition: A Graph Embedding Case Study. SN Computer Science, 4:314, <https://doi.org/10.1007/s42979-023-01716-1>
21. WikiArt. Енциклопедія візуальних мистецтв. <https://www.wikiart.org/uk/mariya-primachenko> (01.04.2024)
22. Gorokhovatsky V.A. Putyatin Y. P. (2009), Image Likelihood Measures of the Basis of the Set of Conformities. Telecommunications and Radio Engineering, 68 (9), p. 763–778.
23. Tvoroshenko I., and Zarivchatskyi R. (2020), Analysis of existing methods for searching object in the video stream, Abstracts of VI International Scientific and Practical Conference «About the problems of science and practice, tasks and ways to solve them» (October 26-30, 2020). Milan, Italy, pp. 500–505.
24. Yakovleva, O., Kovtunencko, A., Liubchenko, V., Honcharenko, V., & Kobylin, O. (2023), Face Detection for Video Surveillance-based Security System (COLINS-2023). In CEUR Workshop Proceedings (Vol. 3403). pp. 69-86.
25. Kuchuk, H., Podorozhniak, A., Liubchenko, N., and Onishchenko, D. (2021), System of license plate recognition considering large camera shooting angles, Radioelectronic and Computer Systems, 4(100), 82–91.
26. OpenCV. URL: <https://docs.opencv.org/>

Received (Надійшла) 08.04.2024

(Accepted for publication) Прийнята до друку 26.06.2024

Metric granulation methods for image description in the image classification problem

Volodymyr Gorokhovatskyi, Natalia Stiahlyk, Yehor Mazur, Anna Vechirska

Abstract. The subject of the research of the article is the methods of image classification in computer vision systems. The aim is to improve the structural methods of classification to increase their speed by implementing a reduced classification features system based on the granulation of reference descriptions. Applied methods: BRISK and ORB key point detectors, set theory and vector spaces apparatus, metric models of vector clustering and granulation, “bad of words” and voting for classification models, software modeling. Results are obtained: models for the transformation of image description in the form of a set of cluster centers and the construction of a reduced description using granulation based on the similarity of vectors were developed, high-speed image classification methods based on the implementation of transformed descriptions were researched. The effectiveness of the developed modifications of the classifier depends on the data granulation method and the threshold parameters for the equivalence of vectors, the minimum metric value and the number of votes of the descriptors for the winner class in the voting scheme. The construction of a hierarchical system of features with several granulation levels was proposed. Based on the implementation of modifications, it was possible to reduce the computational costs by hundreds of times while ensuring the effectiveness of classification on the training sample of data. The practical significance of the work includes the software models for high-speed classification, confirming the functionality and accuracy of the proposed modifications for the applied image dataset, creating software application for implementing modifications of classifiers in computer vision.

Keywords: computer vision; structural methods of image classification; set of descriptors; clustering; vector granulation; classification accuracy.