

В.О. Гороховатський, К.Г. Солодченко

Харківський національний університет радіоелектроніки, Харків

ЗАСТОСУВАННЯ АПАРАТУ АНАЛІЗУ ТА ОБРОБЛЕННЯ БІТОВИХ ДАНИХ У МЕТОДАХ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ ЗА МНОЖИНОЮ КЛЮЧОВИХ ТОЧОК

Вирішена задача структурного розпізнавання візуальних об'єктів на підґрунті описів у вигляді множини ключових точок зображення. Запропоновано метод бінарного аналізу множин дескрипторів опису для формування центрів класів з метою класифікації у межах заданої бази еталонів. Обговорюються критерії оцінювання ефективності класифікації. Проведено програмне моделювання та дослідження методу у порівнянні з медіанними центрами, отримано підтвердження результативності розробленого методу для прикладної бази зображень.

Ключові слова: комп'ютерний зір, структурне розпізнавання, особливі точки зображення, детектор BRISK, дескриптори особливих точок, метод бінарного аналізу, матриця відстаней, база зображень, програмне моделювання, критерій правильної класифікації.

Вступ

У системах комп'ютерного зору при розпізнаванні візуальних об'єктів за їх структурним описом у вигляді множини дескрипторів особливих точок (ОТ) зображення постає задача побудови функції (правила, процедури) класифікації, що задає відношення еквівалентності на множині дескрипторів бази зображень, де класи представлені окремими описами еталонів.

Зважаючи на наявну подібність значень дескрипторів для різних візуальних об'єктів, така еквівалентність на практиці досягається лише наближено [1-6].

Останнього часу набули прикладного застосування такі детектори ОТ, як ORB та BRISK [4, 5], що отримують дескриптор особливих точок у вигляді бінарного вектора з розміром, кратним ступені двійки. Бінарне подання значно прискорює процес порівняння дескрипторів за рахунок можливості застосування двійкових операцій та відповідно спрощує апаратну реалізацію системи розпізнавання.

Крім того, бінарна арифметика дає потенцію застосувати ефективний апарат оброблення бінарних даних та синтезувати нові підходи для визначення подібності дескрипторів ОТ при побудові правил класифікації.

Метою статті є побудова методу класифікації зображень у просторі множин дескрипторів ключових точок з використанням апарату аналізу та оброблення бітових даних.

Задачами дослідження є створення математичних моделей оброблення бітових даних при побудові правил класифікації, вивчення їх властивостей та можливостей параметричного керування ними, оцінювання результативності процедур класифікації для прикладних баз зображень.

Постановка задачі класифікації для множини бітових рядків

Нехай $Z = \{Z^j\}_{j=1}^J$ – множина дескрипторів бази зображень, яку складають описи еталонів, Z^j – еталон, J – число класів, $s(j) = \text{card } Z^j$ – кількість елементів еталону Z^j , $Z \subseteq B^n$, B^n – множина бінарних векторів (ланцюжків, рядків) розмірності n . Кількість розпізнаваних класів визначається числом J еталонів.

Розглянемо правило класифікації $L: B^n \rightarrow [1, \dots, J]$ у вигляді функції такої, що $\forall z \in Z$ виконується $L(z) \in [1, \dots, J]$, тобто кожний елемент буде зараховано до одного із еталонів Z^j . При цьому конструктивно та параметрично функція опирається на апріорні дані конкретної бази, так як приналежність кожного з елементів $z \in Z$ до відповідного еталону всередині бази задана до початку класифікації.

Для побудови процедури класифікації застосуємо дискретну модель простору B^n ознак [1, 6], коли для кожного з еталонів Z^j попередньо сформуємо деякий «центр класу» $m_j \in B^n$, $j = 1, \dots, J$. У процесі здійснення класифікації довільний вектор будемо конкурентно відносити до класу v у відповідності до найменшої відстані

$$v = \arg \min_{j=1, \dots, J} \rho(b, m_j), \quad (1)$$

де $\rho(b, m_j)$ – метрика у просторі B^n .

Зрозуміло, якщо розпізнавання відбувається на універсумі B^n , то правило (1) на практиці треба параметрично доповнити логічним аналізом із застосуванням порогу [2]: віднесення дескриптора ОТ

до класу v здійснюється тільки у випадку, якщо для вектора b буде також виконано умову $\rho(b, m_j) \leq \delta$, де δ – заданий поріг еквівалентності елементів у просторі. При відсутності такого порогового обмеження до класу може бути віднесено зовсім не схожий елемент, хоча мінімум у (1) буде завжди досягнуто.

У результаті запропонованої процедури класифікації множина дескрипторів $O = \{o_1, \dots, o_v\}$, $o_i \in B^n$, що описує довільний візуальний об'єкт, буде подана у вигляді цілочисельного вектора $h[O] = \{h_1, \dots, h_J\}$, де $h_j = \text{card} \{o_i \in O \mid o_i \rightarrow Z^j\}$ – кількість елементів еталону, віднесених до класу з номером j . Одним із методів визначення класу d об'єкта $O \in$ парадигма

$$d = \arg \max_{j=1, \dots, J} h_j. \quad (2)$$

Правило (2) безпосередньо можна застосувати лише в ситуації, коли для подання $h[Z^j]$ кожного з еталонів максимум відповідає номеру j еталона, який класифікується. Коли ця умова не виконується навіть для даних бази еталонів, то треба застосовувати інші більш універсальні підходи [1]. В той же час побудова правила класифікації із застосуванням (2) дає ефективну можливість суттєво спростити прийняття рішення, так як кожний із структурних елементів безпосередньо може бути класифікований без додаткового аналізу їх сукупних властивостей. Застосування бінарного подання та оброблення дає можливість синтезувати саме такий спосіб.

Загалом підхід з визначенням центрів класів забезпечує значне підвищення швидкодії класифікації за рахунок переходу від обчислення міри релевантності множин до порівняння векторів чи навіть окремих їх елементів [6]. Бінарне подання даних додатково спрощує процес обробки, даючи можливість, наприклад в (1), застосовувати метрику Хемінга замість більш суттєвих за обсягом обчислень метрик.

Методи формування центрів класів

Традиційним підходом до обчислення центрів є застосування методів кластеризації, спираючись на ознаки, за якими різняться значення дескрипторів ОТ зображення. Зважаючи на те, що раніш використовували дескриптори SIFT, SURF взагалі нормовані, а вагомість норми для дескрипторів ORB, BRISK не є достатньо значимою характеристикою окремого класу, для кластеризації застосовувалися безпосередньо значення елементів множини Z [1].

У межах апарату бінарних даних викликає інтерес застосування націленого оброблення множини дескрипторів із визначенням одного з елементів, що

буде представляти множину як клас. Двійковий вид дескрипторів дає підвалини застосувати апарат логічного оброблення. Зважаючи на бінарний вид BRISK- та ORB-дескрипторів, для кожного з еталонів Z^j визначимо бінарний вектор центра класу на підставі логічного правила

$$m_j(a) = \begin{cases} 1, & f(Z^j, a) \geq s(j)/2, \\ 0, & \text{інакше,} \end{cases} \quad (3)$$

де $m_j(a)$ – значення біту на позиції з номером a для центру m_j , $f(Z^j, a)$ – функція, що підраховує кількість одиничних бітів на позиції з номером a у множині дескрипторів еталону Z^j , $a = 1, \dots, n$. Функція $f(Z^j, a)$ може бути обчислена безпосередньо додаванням бітів як

$$f(Z^j, a) = \sum_{d=1}^{s(j)} x_d(a), \quad x_d \in Z^j, \quad (4)$$

де $x_d(a)$ – біт з номером a для дескриптора з номером d в описі еталона.

За виразом (3) значення кожного з бітів центру m_j для j -го класу визначається переважно більшістю значень відповідних розрядів усіх дескрипторів ОТ, які належать еталону Z^j . Бітове подання можна ефективно використати при роботі з множинами даних. Підрахунок числа одиничних бітів у (4) можна здійснити також застосуванням відповідних бінарних функцій [11].

Аналізуємо елементи $x \in X$ загального вмісту структурних описів бази еталонів (навчальна вибірка) шляхом віднесення їх до відповідного класу з використанням конкуренції (1). У якості ρ застосуємо метрику Хемінга

$$\rho(x, m_j) = \sum_{a=1}^n |x(a) - m_j(a)|, \quad (5)$$

що визначає кількість розбіжних бітів для двійкових послідовностей однакової довжини.

У результаті оброблення з використанням попередньо обчислених центрів одержуємо подання для кожного еталона: $h[Z^j] = (h_1^j, \dots, h_n^j)$, яке відповідає розподілу елементів множини Z^j за класами еталонів, де h_a^j – цілі числа, причому $\sum_{a=1}^{s(j)} h_a^j = s(j)$, так як кожний з елементів опису буде віднесено до одного з фіксованих класів.

Важливою характеристикою множини отриманих центрів є симетрична матриця відстаней між ними $\Upsilon = \{\rho(m_i, m_j)\}_{i,j=1}^J$, яка відображає їх сумісну наближеність у просторі B^n та фактично визначає фінальну якість класифікації. Чим більше віддалені центри між собою, тим вище досяжна ступінь розрізненості дескрипторів із різних класів.

Загальним же критерієм якості квантування простору у відповідності до апріорних даних – опису бази зображень у вигляді множин дескрипторів ОТ може бути нормоване значення розкиду даних навколо сформованих центрів

$$E = \frac{1}{sN} \sum_{j=1}^J \sum_{v=1}^{s(j)} \rho(x_v, m_j), \quad (6)$$

де N – розмір дескриптора ($N = 512$ для BRISK), $s = \sum_{j=1}^J s(j)$ – загальне число елементів навчальної вибірки, а для спрощення та приведення до інтервалу $0 \dots 1$ замість традиційного квадрата метрики вибрано її значення [7].

Зважаючи на бітове подання даних, виникає можливість також застосувати по-байтовий аналіз, використавши замість виразів (5), (6) міри, обчислені на підставі подібності окремих байтів.

З метою порівняння показників класифікації також розглянемо у якості центру множини її узагальнену медіану [8]. Узагальнена медіана m скінченної множини $X = \{x(i)\}$ визначається шляхом мінімізації функціоналу D :

$$m = \arg \min_{v \in X} D(v), \quad D(v) = \sum_{x(i) \in X} \rho(x(i), v), \quad (7)$$

де $\rho(x(i), v)$ – відстань між елементами. Обчислити медіану можна шляхом агрегації значень рядків матриці взаємних відстаней елементів множини [6]. У результаті обчислення (7) отримуємо медіану m_j як один із елементів еталону $m_j \in Z^J$.

Вектор $h[Z^i]$ також може виступити критерієм якості класифікації на навчальній вибірці. Ідеальним класифікатором є той, для якого $\arg \max_j h_j^i = i$, тобто при реалізації (2) максимальне значення досягається для номера еталону, що класифікується.

Другим критерієм при умові правильної класифікації є відношення загального числа дескрипторів, які за результатом оброблення віднесені до правильного класу, до гуртового числа дескрипторів бази

$$\beta = \sum_{i=1}^J h_i^i / s. \quad (8)$$

Критерій (8) – це відношення кількості правильно класифікованих дескрипторів бази до їх загального числа. Ідеальним варіантом є значення $\beta = 1$.

Результати комп'ютерних експериментів

У середовищі C++ з використанням бібліотеки програм Open CV [10] нами розроблено програмну модель для класифікації зображень на підставі описів, сформованих детектором BRISK. У якості бази еталонів застосовано набір із 5 зображень мечів, два з яких показано на рис. 1. Кількість виявлених ОТ в кожному із еталонів розміром 256×256 обмежена числом 51, загальна кількість ОТ бази (обсяг навчальної вибірки) дорівнює 255. Через обчислення центрів класів у відповідності до логіки (3) сформовано вектори $h[Z^i]$ для кожного із еталонів. Приклади центрів показано на рис. 2.

Характеристикою, що відображає властивості множини центрів, є матриця відстаней між ними (табл. 1). Можна побачити, що отримані центри досить вдало розрізняють еталонні описи, так як мінімальне значення серед діагональних елементів дорівнює 69 (відстань між центрами 1 і 3 еталонів), що чимало віддалено від мінімального значення 0.

Зауважимо, що застосований детектор та бінарне оброблення (3) відіграють провідну роль при розрізненні даних конкретної бази. У результаті оброблення (3) та застосування (1) з метрикою (5) отримано матрицю з рядками $h[Z^i]$ по класах у вигляді табл. 2.

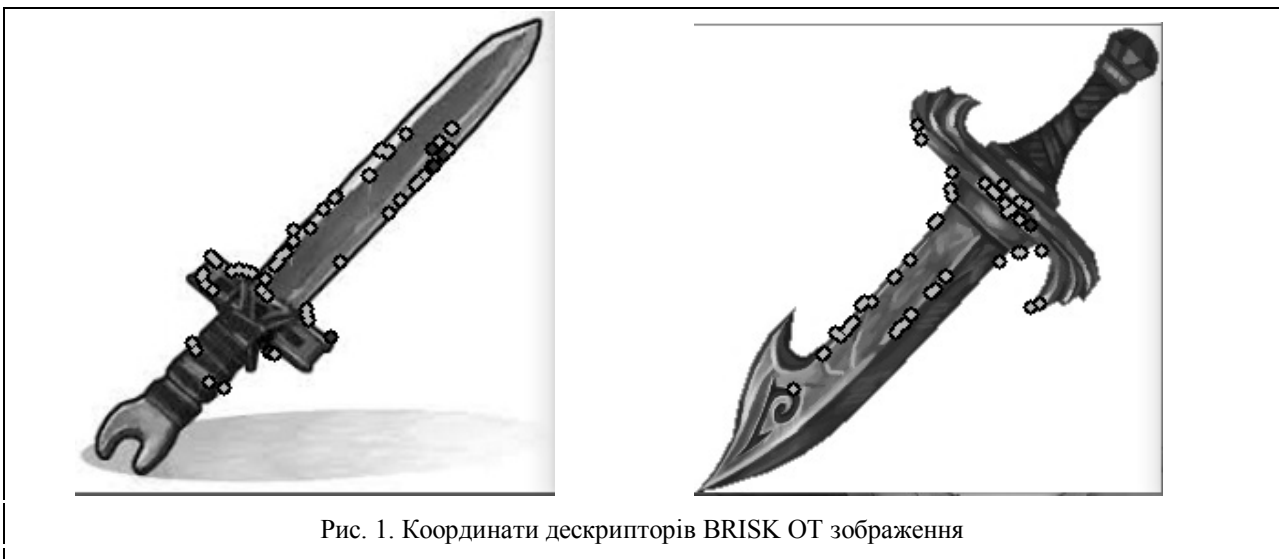


Рис. 1. Координати дескрипторів BRISK ОТ зображення



Рис. 2. Бітове подання (3) для центрів зображень рис. 1

Таблиця 1
Матриця відстаней між центрами

Центр	Відстань				
	Z^1	Z^2	Z^3	Z^4	Z^5
Z^1	0	135	69	98	121
Z^2	135	0	108	137	122
Z^3	69	108	0	99	116
Z^4	98	137	99	0	123
Z^5	121	122	116	123	0

Як бачимо із табл. 2, множина дескрипторів BRISK для даної бази зображень з використанням бінарного оброблення однозначно класифікує множину ОТ кожного з еталонів, так як в рядках таблиці кількість ОТ на діагоналі максимальна, причому максимальний елемент відповідає зразку та у 2,2 рази перевищує найближчий із локальних максимумів. Розкид (6) отримав значення $E = 0,26$, а відношення правильно класифікованих дескрипторів до загального числа склало $\beta = 0,55$.

Таблиця 2
Розподіл ОТ еталонів по класах

Еталони	Класи				
	1	2	3	4	5
Z^1	26	2	4	9	10
Z^2	9	32	7	1	2
Z^3	9	11	25	4	2
Z^4	5	4	7	30	5
Z^5	9	2	2	12	26

Час формування 5 центрів склав 3.8 мс, а класифікація множини дескрипторів склала 4.4 мс на комп'ютері з процесором I5-250M (2.50 GHz) та з оперативною пам'яттю розміром 4 Гб, операційна система Windows 10 x64.

Для бази інших 5 зображень отримано аналогічні позитивні результати, причому відношення склало $\beta = 0,71$, а значення $E = 0,31$. Експеримент, проведений для 10 еталонів (навчальна вибірка 510 дескрипторів), показав близькі результати: усі еталони класифікуються вірно, розкид (6) дорівнює приблизно $E = 0,28$, а відношення правильно класифікованих дескрипторів до загального числа склало $\beta = 0,51 \dots 0,54$, тобто дещо менше, ніж для 5 еталонів. Зі збільшенням кількості еталонів треба і більше число точок, щоб отримати стабільне рішення. На нашу думку, у випадку 10 еталонів мінімально необхідно хоча б по 100 ОТ для кожного класу.

Застосування у експериментах способу оброблення даних по байтам не дало якихось суттєвих переваг: для 5 еталонів розкид (6) отримав значення $E = 0,26$, а відношення (8) склало $\beta = 0,53$.

Аналіз із використанням медіани у якості центрів класів показав наступне. Для 5 еталонів розкид (6) отримав значення $E = 0,31$, а відношення правильно класифікованих дескрипторів до загального числа склало $\beta = 0,44$. Як бачимо, застосування медіани у порівнянні з моделлю (3) дещо збільшує помилку квантування та зменшує долю правильно класифікованих даних.

З метою поглибленого вивчення властивостей створеного методу проведено повторне застосування бітового оброблення (3) до кластерів зображень, отриманих застосуванням (3) до вхідних описів візуальних об'єктів. Дослідження показало деяке зниження помилки $E = 0,28$ з одночасним незначним зниженням критерію $\beta = 0,47$, що підкреслює стабільність та можливість ітераційного застосування моделі (3).

Загалом треба сказати, що детектор BRISK за рахунок суттєво інформативного значення розміру дескриптора, що дорівнює 512, гарантує досить детальне подання і можливість аналізу властивостей зображення з метою результативного розрізнення візуальних об'єктів. Підтвердженням цього є факт, що детектор ORB з розміром 256 для цих же даних не дає результату відповідної якості, тобто напругу не забезпечує розрізнення еталонів.

Висновки

У результаті дослідження запропоновано ефективний метод класифікації зображень з використанням бінарного оброблення.

При практичній побудові правил класифікації візуальних об'єктів із застосуванням детекторів ОТ треба звернути увагу на ефективний вибір та поєднання можливостей як детекторів ОТ, так і апарату для оброблення їх значень для досліджуваних баз зображень.

Розроблений підхід універсально може бути застосовано для класифікації множин бінарних даних довільної природи. Етап попередньої обробки з визначенням центрів на базі бінарного аналізу можна вважати різновидом хешування на множині дескрипторів ОТ еталону.

Серед досліджених варіантів обчислення центрів класів найкращі показники має логічний аналіз із визначенням окремих бітів на множині ознак еталону.

Наукова новизна дослідження полягає у синтезі нового продуктивного методу класифікації зображень з використанням бітового аналізу та оброблення даних у просторі дескрипторів ключових точок зображення.

Практична значущість роботи – отримання практичних програмних моделей для застосування та оцінювання ефективності методів класифікації з бінарним поданням даних і підтвердження їх результативності в прикладних базах зображень.

Список літератури

1. Гороховатский В.А. Исследование результативности структурных методов классификации изображений с применением кластерной модели данных / В.А. Гороховатский, Е.П. Путятин, В.С. Столяров // Радиоэлектроника, информатика, управление. – 2017. – №3 (42). – С. 78–85.

2. Гороховатский В.А. Структурный анализ и интеллектуальная обработка данных в компьютерном зрении / В.А. Гороховатский. – Х.: Компания СМІТ, 2014. – 316 с.

3. Duda R.O. Pattern classification / Duda R.O., Hart P.E., Stork D.G. – 2ed., Wiley, 2000.–738p.

4. Ethan Rublee, Vincent Rabaud, Kurt Konolige, Gary Bradski. ORB: an efficient alternative to SIFT or SURF. – Computer Vision (ICCV), IEEE International Conference on IEEE, pp. 2564 – 2571, 2011.

5. Stefan Leutenegger, Margarita Chli, Roland Y. Siegwart. BRISK: Binary Robust Invariant Scalable Keypoints. – Computer Vision (ICCV), pp. 2548 – 2555, 2011.

6. Gorokhovatsky V.A. Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions / V.A. Gorokhovatsky // Telecommunications and Radio Engineering. – 2016, Vol. 75, No 14. – P. 1271–1283.

7. Осовский С. Нейронные сети для обработки информации / С. Осовский. – М.: Финансы и статистика, 2002. – 344 с.

8. Кохонен Т. Самоорганизующиеся карты / Т. Кохонен. – М.: БИНОМ. Лаборатория знаний, 2013. – 655 с.

9. Shapiro, L. and Stockman, G., (2001), Computer vision, Prentice Hall. – 625 p.

10. OpenCV Open Source Computer Vision. [Электронный ресурс]. – Режим доступа: <https://docs.opencv.org/master/index.html>, свободный. – Загл. с экрана.

11. Солонина А.И. Алгоритмы и процессоры цифровой обработки сигналов / А.И. Солонина, Д.А. Улахович, Л.А. Яковлев. – СПб.: БХВ-Петербург, 2001. – 464 с.

Надійшла до редколегії 11.02.2018

Рецензент: д-р техн. наук, проф. Є.П. Путятін, Харківський національний університет радіоелектроніки, м. Харків.

ПРИМЕНЕНИЕ АППАРАТА АНАЛИЗА И ОБРАБОТКИ БИТОВЫХ ДАННЫХ В МЕТОДАХ КЛАССИФИКАЦИИ ИЗОБРАЖЕНИЙ ПО МНОЖЕСТВУ КЛЮЧЕВЫХ ТОЧЕК

В.А. Гороховатский, К.Г. Солодченко

Решена задача структурного распознавания визуальных объектов на основе описаний в виде множества ключевых точек изображения. Предложен метод бинарного анализа множеств дескрипторов описания для формирования центров классов с целью классификации в пределах заданной базы эталонов. Обсуждаются критерии оценки эффективности классификации. Проведено программное моделирование и исследование метода по сравнению с медианными центрами, получено подтверждение результативности разработанного метода для прикладной базы изображений.

Ключевые слова: компьютерное зрение, структурное распознавание, особые точки изображения, детектор BRISK, дескрипторы особых точек, метод бинарного анализа, матрица расстояний, база изображений, программное моделирование, критерий правильной классификации.

APPLICATION OF THE APPARATUS FOR ANALYSIS AND PROCESSING OF DATA BITS IN METHODS OF CLASSIFICATION OF IMAGES ON A SET OF KEY POINTS

V.O. Gorokhovatsky, K.G. Solodchenko

The problem of structural recognition of visual objects based on descriptions in the form of a set of key points of the image was solved. A method is proposed for binary analysis of sets of description descriptors for the formation of class centers with the aim of classifying them within a given base of standards. Criteria for evaluating the effectiveness of classification are discussed. The software modeling and research of the method was carried out in comparison with the median centers, the efficiency of the developed method for the applied image database was confirmed.

Keywords: computer vision, structural recognition, special image points, BRISK detector, special point descriptors, binary analysis method, distance matrix, image database, software modeling, criterion of correct classification.