

В. О. Гороховатський, Д. О. Руденко, Т. О. Сірик

Харківський національний університет радіоелектроніки, Харків, Україна

ДОСЛІДЖЕННЯ СИСТЕМИ ІЄРАРХІЧНИХ ОЗНАК ПРИ БЛОЧНОМУ ПОДАННІ ОПИСУ У СКЛАДІ МНОЖИНИ КЛЮЧОВИХ ТОЧОК ЗОБРАЖЕННЯ

Предметом досліджень статті є ієрархічні моделі для встановлення ступеня релевантності описів зображень при розпізнаванні візуальних об'єктів у системах комп'ютерного зору. **Метою** є розроблення модифікації методу структурного розпізнавання на підставі впровадження блокових моделей даних із інтегруванням ймовірнісних розподілів. **Завдання:** розроблення математичних та програмних моделей для ефективного за швидкодією ієрархічного оброблення даних при визначенні релевантності структурних описів, вивчення властивостей цих моделей, оцінювання результативності при обробленні зображень. Застосовуваними **методами** є: детектор ORB для формування дескрипторів ключових точок, інтелектуальний аналіз даних, методи побудови розподілів бітових даних, апарат метричного визначення релевантності, програмне моделювання. Отримані такі **результати**. Перехід від опису множин дескрипторів до розподілів фрагментів, побудова ієрархічних ознак забезпечують необхідну результативність розпізнавання. Оброблення та аналіз даних виконується у кілька разів швидше, ніж на підставі розподілів. **Висновки.** Наукова новизна дослідження полягає в удосконаленні методу структурного розпізнавання зображень на основі впровадження блочної структури опису із використанням інтегрованих значень розподілу для фрагментів множини дескрипторів. Практична значущість – досягнення суттєвого рівня підвищення швидкодії при обчисленні релевантності, підтвердження результативності запропонованих ієрархічних ознак на прикладах зображень, отримання прикладних програмних моделей для дослідження та впровадження методів класифікації в системах комп'ютерного зору.

Ключові слова: структурні методи розпізнавання зображень, ключова точка, детектор ORB, дескриптор, розподіл даних фрагменту, релевантність описів, манхеттенська метрика, швидкодія визначення релевантності.

Вступ

Процес реалізації дієвих класифікаційних рішень у сучасних системах комп'ютерного зору потребує вирішення ряду проблем, пов'язаних із багатовимірною природою оброблюваних даних. При застосуванні структурних методів розпізнавання описи об'єктів подаються у вигляді наборів дескрипторів ключових точок (КТ) як скінченної множини числових векторів достатньо високої розмірності [1-8]. Так, бінарні описи, отримані детекторами ORB, BRISK [4, 9], містять 256 та 512 компонентів. У такому випадку перехід до подання даних у вигляді системи їх фрагментів меншої розмірності сприяє суттєвому спрощенню їх прикладного застосування [5]. Основним інструментом при цьому виступає апарат просторово-статистичного аналізу даних, який в аспекті розпізнавання базується на узагальненні даних опису об'єкта шляхом обчислення нових ознак для сукупності фрагментів [2; 5].

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

Метою статті є розвинення методу структурного розпізнавання зображень на підставі впровадження просторово-статистичного оброблення для кортежу значень блоків опису у складі множини дескрипторів.

Задачами дослідження є опрацювання моделей для обчислення релевантності модифікацій структурного опису, ефективних за швидкодією оброблення

даних, аналіз властивостей цих моделей та експериментальне оцінювання результативності запропонованих підходів за наслідками аналізу зображень.

Статистичні моделі у просторі ознак

Бітовий склад дескрипторів КТ у просторі B^n бінарних векторів (n – ступінь двійки) дає можливість запровадити аналіз дескриптора як кортежу послідовних елементів, діапазон значень яких відомий. Це дає змогу розглядати наявний структурний опис (множина бінарних векторів) у синтезованому просторі даних та здійснювати статистичний аналіз розпізнаваних зображень з урахуванням властивостей складу еталонних описів.

Опис Z розпізнаваного візуального об'єкту визначимо у вигляді скінченної множини $Z = \{z_v\}_{v=1}^s$, $z_v \in B^n$, $Z \subset B^n$ із s бінарних дескрипторів КТ, наприклад, сформованих детекторами ORB, BRISK, AKAZE [4, 5, 9].

Для фіксованого числа n подамо опис об'єкту бінарною матрицею $D = \{\{d_{i,j}\}_{i=1}^s\}_{j=1}^n$, сформовану довільною послідовністю отриманих детектором s дескрипторів КТ. Для спрощення аналізу вважаємо значення s однаковим для всіх еталонів, цього нескладно досягти випадковим або цілеспрямованим відбором необхідної кількості дескрипторів.

Кожний рядок матриці D , що є дескриптором КТ, подамо у виді m непересічних фрагментів (блоків), що послідовно слідуєть один за одним. Отримаємо новий простір даних, перевагою якого є можливість аналізувати внутрішню структуру та статистичні властивості опису, на основі чого його можна буде віднести до одного із еталонних класів. Застосуємо мережу блоків як основу для прийняття рішення про

клас об'єкту. Блочний спосіб формування ознак надає додаткову гнучкість при обробленні [1].

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

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

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

$$q_i = \text{card}\{z_v^k \in z_v, z_v \in Z \mid z_v^k = i\}. \quad (1)$$

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

Для кожного із m фрагментів побудуємо розподіл $q = \{q_1, \dots, q_w\}$, величина w визначена діапазоном значень даних для фрагмента, що визначається його розміром. Наприклад, для дескриптора ORB при розбитті на байти для $n = 256$ маємо $m = 32$, $w = 256$. Загалом множина Z під дією відображення Ω описується матрицею $Q = \{\{q_{i,v}\}_{i=1}^w\}_{v=1}^m$, що містить розподіли системи фрагментів.

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

Інтегрована ієрархічна система ознак

На основі одно-бітових розподілів побудуємо інтегровану в межах окремих фрагментів систему ознак. Для матриці D опису обчислимо значення сум вздовж стовпців (1-бітовий фрагмент), отримаємо вектор $t = (t_1, \dots, t_j, \dots, t_n)$, де $t_j = \sum_{i=1}^s d_{i,j}$, $j = \overline{1, n}$. Це і будуть ознаки найнижчого рівня, що у

повній мірі зберігають інваріантні властивості множини дескрипторів Z .

На базі ознак t_j обчислимо ознаки вищого рівня u_k для блоків як сукупностей стовпців

$$u_k = \sum_{j=k}^{k+b-1} t_j, \quad (2)$$

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

Ознаки (2) реалізують крос-кореляційне оброблення матриці D з прямокутною маскою розміром $b \times s$ [10]. У результаті обчислення (2) отримуємо цілочисельний вектор u_k розмірністю m . Параметр m є характеристикою системи фрагментів, він зменшується від n до 1 зі збільшенням розміру фрагмента від 1 до n .

Значення вектора $u = (u_1, \dots, u_k, \dots, u_m)$ можуть бути використані як самостійні структурні ознаки статистичного виду. Опрацювання (2) залежить від параметра b та реалізує просторове оброблення (інтегрування даних) на множині дескрипторів. Зважаючи на таку просту модель обчислення функцій (2), всі вони для довільного розміру фрагмента визначаються нескладно (логічно або додаванням цілих чисел), базуючись на значеннях статистичного розподілу.

На основі подання (2) може бути застосовано також ієрархічний спосіб розпізнавання, що використовує для зіставлення з еталонами систему ознак u_k з різним ступенем інтегрування даних, що визначається розміром блока.

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

Обсяг значень для ознак (2) змінюється від одного числа до вектора розміру n (розмір дескриптора) і зменшується зі збільшенням розміру фрагмента. При $m = 1$ маємо $b = n$, тобто кожний дескриптор містить один фрагмент, а величина u_1 визначається одним значенням і дорівнює сумі бітів усього опису. При $m = n$ маємо іншу граничну ситуацію: n 1-бітових фрагментів, а u_k обчислюється як сума стовпця матриці D .

Модель (2) реалізує процедуру зменшення інформаційної надмірності просторового сигналу за рахунок допустимого зниження (з точки зору якості розрізнення) ступеня розподільної здатності системи ознак подання опису.

Оброблення виду (2) є різновидом просторового аналізу значень для множини дескрипторів вздовж координат їх компонентів.

Іншими варіантами такого оброблення є фільтрація чи розклад дескриптора за системою ортогональних функцій [1, 10].

Релевантність r описів a та b визначимо на підставі зіставлення розподілів на підставі манхеттенської відстані між векторами $u(a), u(b)$:

$$r[u(a), u(b)] = \sum_{i=1}^m |u_i(a) - u_i(b)|. \quad (3)$$

Для більш точного оцінювання значення релевантності можна застосувати нормовані значення (3):

$$r^* = r/r_1, \quad (4)$$

де r_1 – максимально можлива відстань для обраного числа дескрипторів.

Аналіз результатів комп'ютерного моделювання

Розроблену модель побудови ієрархічних ознак застосовано на прикладі зображень гербів університетів Оксфорду, Гарварду, Пекіну та Массачусетса. Розмір зображень складає 600x600. Програмне середовище – Visual Studio 2015 із використанням

засобів бібліотеки Open CV [11]. Ілюстрація зображення герба Гарварду та Массачусетса показана на рис. 1. Для моделювання застосовано дескриптори ORB розмірністю $n = 256$.

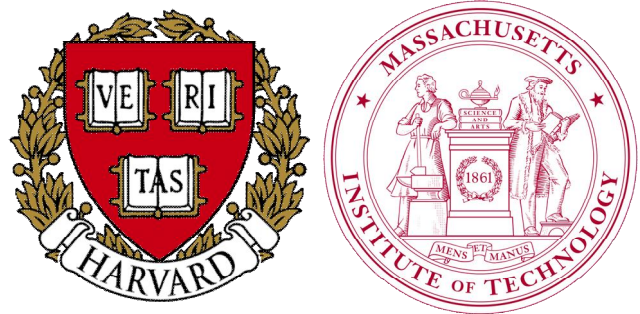


Рис. 1. Аналізовані зображення гербів

На рис. 2 наведено приклад сформованого 1-бітового розподілу (256 ланок дескриптора ORB) для кількості нулів у 100 дескрипторах герба рис 1.

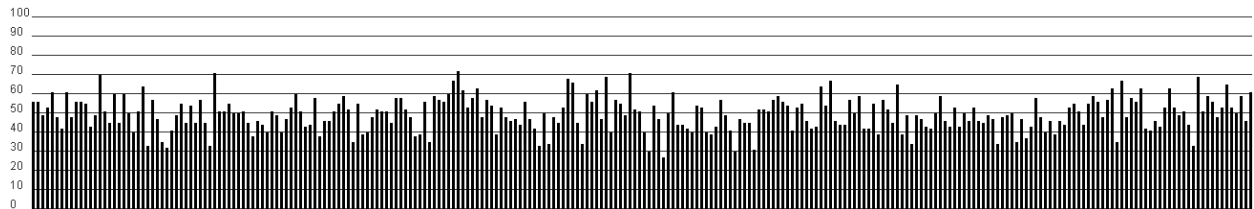


Рис. 2. Приклад розподілу для 100 дескрипторів

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

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

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

На рис. 3 показано схему побудови ієрархічних ознак для емпіричного опису зображення рис. 1 у складі 100 дескрипторів, де значення у нижніх колах відповідають 8-бітовому розподілу нулів, а верхнє значення дорівнює загальному числу нулів опису. Аналогічну структуру можна побудувати для довільного опису [12].

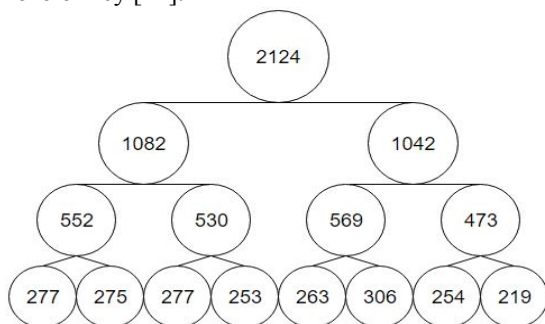


Рис. 3. Значення ієрархічних даних на прикладі герба Гарварду

Табл. 2 містить значення відстані (3) між розподілами зображень гербів Масачусетського та Оксфордського університетів для 100, 300, 500 дескрипторів та різним числом бітів у фрагментах розподілу. Також у табл. 2 наведені відстані для скорочених розподілів цих гербів, де вибрано кожен другий та четвертий біти із однобітового розподілу.

Таблиця 2 – Значення нормованої відстані (4) для різної кількості дескрипторів та бітів розподілу

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

При зменшенні кількості ключових точок майже пропорційно зменшуються й максимальна та фактична відстань між розподілами. Тобто збільшення числа дескрипторів від 100 до 300 не має суттєвого впливу на характеристики розпізнавання зображень. Значення табл. 2 показують, що 100 ключових точок достатньо для розрізнення зображень гербів. Для інших типів зображень або кількості розпізнаваних еталонів ці значення можуть відрізнятися.

У табл. 3 наведено дані наявних обчислювальних витрат для реалізації варіантів розглянутих моделей при обчисленні релевантності описів зображень гербів Масачусета та Оксфорду в залежності від кількості бітів у фрагментах розподілу. Табл. 3 містить час опрацювання комп'ютерними моделями у наносекундах.

Аналізуючи значення табл. 2 та оцінку часу оброблення з табл. 3 для 100 КТ, можна сказати, що при обчисленні по правилу «кожен другий» час прогнозовано зменшився у порівнянні з 1-бітовим розподілом приблизно у 1,5 рази. А нормована відстань (4) дещо зменшилась від 0,147 до 0,105.

Таблиця 3 – Оцінка часу при обчисленні релевантності

Кількість дескрипторів	Кількість бітів розподілу							
	1	2	4	8	128	256	Кожен 2-й	Кожен 4-й
500	66929	33944	26222	22095	20388	19913	32493	26511
300	46465	65172	34177	22327	19612	19914	33139	26087
100	47018	32586	24540	21472	18572	16293	32133	25384

При обчисленні по правилу «кожен четвертий» час зменшився у порівнянні з 1-бітовим розподілом у 1,9 рази. А нормована відстань (4) практично не змінилась в межах 0,15.

Ці показники стверджують про достатні властивості модифікацій побудованих ознак для розпізнавання зображень гербів зі значно меншим часом оброблення.

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

поділу, то він прямо пропорційний числу дескрипторів опису.

Зважаючи на те, що час обчислення релевантності за розподілами приблизно у 1000 разів менший, ніж для традиційної процедури голосування [5], бачимо, що за допомогою впровадження системи ієрархічних ознак можна добитися подальшого скорочення часу обчислень, забезпечуючи при цьому достатньо високу результативність розпізнавання.

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

Висновки

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

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

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

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

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

Перспективи дослідження можуть бути пов'язані із побудовою дерев рішень на основі ієрархічних ознак.

СПИСОК ЛІТЕРАТУРИ

1. Гороховатский В.А. Структурный анализ и интеллектуальная обработка данных в компьютерном зрении / В.А. Гороховатский. – Х.: Компания СМІТ, 2014. – 316 с.
2. Gorokhovatsky, V.O. and Gadetska, S.V., (2019) Determination of Relevance of Visual Object Images by Application of Statistical Analysis of Regarding Fragment Representation of their Descriptions, Telecommunications and Radio Engineering, 78 (3), pp. 211–220.
3. Gorokhovatsky V.A. Efficient Estimation of Visual Object Relevance during Recognition through their Vector Descriptions / V.A. Gorokhovatsky // Telecommunications and Radio Engineering. – 2016, Vol. 75, No 14. – pp. 1271–1283.
4. Stefan Leutenegger, Margarita Chli, Roland Y. Siegwart. BRISK: Binary Robust Invariant Scalable Keypoints. – Computer Vision (ICCV), pp. 2548 – 2555, 2011.
5. Гороховатський В.О. Статистичні розподіли та ланцюжкове подання даних при визначенні релевантності структурних описів візуальних об'єктів / В.О. Гороховатський, С.В. Гадецька, Р.П. Пономаренко // Системи управління, навігації та зв'язку. – 2018. – №6 (52). – С. 87–92.

6. Sivaram, M., Porkodi, V., Mohammed, A.S., Manikandan V. Detection of Accurate Facial Detection Using Hybrid Deep Convolutional Recurrent Neural Network. *ICTACT Journal on Soft Computing*. 2019. Vol. 09, Issue 02. pp. 1844-1850. DOI: 10.21917/ijsc.2019.0256
7. Yogesh Awasthi, R P Agarwal, B K Sharma, "Intellectual property right protection of browser based software through watermarking technique", *International Journal of Computer Applications*, vol. 97, no. 12, 2014, pp. 32-36.
8. Yogesh Awasthi, R P Agarwal, B K Sharma, "Two Phase Watermarking for Security in Database", *International Journal of Computing*, vol. 4, no. 4, 2014, pp. 821-824.
9. Rublee, E., Rabaud, V., Konolige, K., and Bradski, G., (2011) ORB: an efficient alternative to SIFT or SURF, *IEEE International Conference on Computer Vision (ICCV)*, Proceedings, pp. 2564-2571.
10. Шапиро Л. Компьютерное зрение/ Л. Шапиро, Дж. Стокман.; пер. с англ. – М.: БИНОМ. Лаборатория знаний, 2006. – 752 с.
11. Прохоренко Н.А. OpenCV и Java. Обработка изображений и компьютерное зрение / Н. Прохоренко. – СПб.: БХВ-Петербург, 2018. – 320 с.
12. Adelson E.H. Pyramid methods in image processing [Электронный ресурс] / E. Adelson, C. Anderson, J. Bergen, P. Burt, J. Ogden // *RCA Engineer*. – Vol. 29(6), pp. 33-41. Режим доступа: http://persci.mit.edu/pub_pdfs/RCA84.pdf

Рецензент: д-р техн. наук, проф. Є. П. Путятін,
Харківський національний університет радіоелектроніки, м. Харків
Received (Надійшла) 30.01.2019
Accepted for publication (Прийнята до друку) 20.03.2019

Исследование системы иерархических признаков при блочном представлении описания в составе множества ключевых точек изображения

В.А. Гороховатский, Д.А. Руденко, Т.А. Сирьк

Предметом исследования статьи являются иерархические модели для определения степени релевантности описаний изображений при распознавании визуальных объектов в системах компьютерного зрения. **Целью** является разработка модификации метода структурного распознавания на основе внедрения блочных моделей данных с интегрированием вероятностных распределений. **Задание:** разработка математических и программных моделей для эффективной по быстродействию иерархической обработки данных при определении релевантности структурных описаний, изучение свойств этих моделей, оценивание результативности при обработке изображений. **Использованными методами** являются: детектор ORB для формирования дескрипторов ключевых точек, интеллектуальный анализ данных, методы построения распределения битовых данных, аппарат метрического определения релевантности, программное моделирование. Получены такие **результаты**. Переход от описания как множеств дескрипторов к распределениям фрагментов, построение иерархических признаков обеспечивают необходимую результативность распознавания. Обработка и анализ данных выполняется в несколько раз быстрее, чем на основании распределений. **Выводы.** Научная новизна исследования состоит в усовершенствовании метода структурного распознавания изображений на основе внедрения блочной структуры описания с использованием интегрированных значений распределения для фрагментов множества дескрипторов. Практическая значимость – достижение существенного уровня повышения быстродействия при вычислении релевантности, подтверждение результативности иерархических признаков на примерах изображений, получение прикладных программных моделей для исследования и внедрения методов классификации в системах компьютерного зрения.

Ключевые слова: структурные методы распознавания изображений, ключевая точка, детектор ORB, дескриптор, распределение данных фрагмента, релевантность описаний, манхэттенская метрика, быстродействия определения релевантности.

Investigation of the hierarchical features system in the block feed of the description in the composition of the image key points

V. Gorokhovatskyi, D. Rudenko, T. Siryk

The subjects of the paper are the hierarchical models for estimation of the image descriptions relevance when recognizing visual objects in computer vision systems. **The goal** is to modify an image structural recognition method based on the implementation of block data models with the integration of probability distributions. **The tasks are:** include the development of mathematical and software models of efficient hierarchical data processing for determining the relevance of structural descriptions, investigation of the properties of these models, evaluation of the effectiveness of image processing. **The methods** are used: an ORB detector to form the key point descriptors, data mining, methods for construction of the bit-data distribution, a method of metric relevance estimation, software modeling. **The following results** were obtained. The transition from the sets of descriptors to distributions of fragments, the construction of hierarchical features provide the necessary recognition performance. Data processing and analysis are performed several times faster compared to ones based on distributions. **Conclusions.** The contribution of the paper is the improvement of the structural image recognition method with the introduction of a block description structure using integrated distribution values for fragments of the set of key point descriptors. The practical significance of the paper is the achievement of an increase of image relevance calculation speed, confirmation of the effectiveness of hierarchical features using image examples, obtaining of an application software models for research and implementation of classification methods in computer vision systems

Keywords: structural image recognition methods, key point, ORB detector, descriptor, fragment data distribution, descriptive relevance, Manhattan metric, speed of relevancy estimation.