UDC 004.7 **doi:** https://doi.org/10.20998/2522-9052.2021.3.02

Pavlo Pustovoitov, Maxim Okhrimenko, Vitalii Voronets, Dmitry Udalov

National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine

THE SPEED CALCULATING INCREASING METHOD OF THE MARKOV MODEL NETWORK NODE

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

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

Introduction

A network node solves a set of tasks for servicing packet flows, which in terms of Markov, queuing systems are called queries. Queries are generated by various sources in a random way, at random time moments and arrive at the network node interface. Next, the packets pass through the node interface to a free intelligent node to initiate the service process. If all intelligent nodes are busy at the moment, then the newly arrived query from the interface enters the queue to wait for service (Fig. 1).

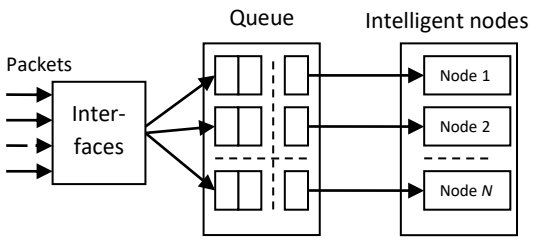

Fig. 1. Scheme of the work of network node f or packets service

The network node functioning for packet flows service can be described by using models of multidimensional continuous random processes with a discrete states set.

Analysis of the literature

Queuing systems (QS) are the most common mathematical apparatus used to analyze processes in networks. Classical works in the field of the theory of QS are [1, 2, 3]. Sources [2, 3] generalize the theory of QS, considering the subject area objects as a set of "queues" and "service channels" at the input of which queries for service are received. Single and multi-link, single and multichannel QS with a normal uniform and exponential queries flow are considered in these sources. The general laws of the query delay time calculating, the length of the queue, etc., depending on the type of QS and the flow of queries, are substantiated and deduced. The source [2] considers the QS theory for substantiating the structure and topology of ARPONET networks. In this work, the main assumptions that allowed to obtain analytical dependencies for the network quality assessment are substantiated. Kleinrock-Jackson Model represents a network as a set of independent networks, each of them is a single-channel QS with an infinite queue and flow of packets (queries). Kleinrock Independence Approximation is the assumption about the independence of the flow at each node, the independence of the query arrival time and its length from the previous queries, while it is assumed that the query flow is stationary and the query arrival time and service time are distributed exponentially. In [2] the main network quality indicators (QoS) are formulated as:

- average time of query service;
- effective bandwidth with constraints on the denial of service probability, on the network topology, on the network cost (directly proportional to the length and maximum network bandwidth);
- denial of service probability.

The constraints and assumptions of [2] will subsequently form the basis of many studies on the information computing network (ICN) quality assessment. Further works set the goal of various types network modeling on the QS basis. It were substantiated the ALOHA models for satellite channels [4, 2], multiple access models for air radio communication networks [2, 12], models of wired (including fiber-optic) "point-to-point" type networks [4, 2, 5, 6]. Based on the QS, two-level models were built, which include backbone network model, serving traffic between computer complexes and terminal networks, describing terminal access to computer complexes [7, 4, 8, 9, 11]. However, the traffic redistribution, the new data transmission technologies emergence [10, 11], the formation of glob-

al networks associated with the information and computing support growth in the 80-90s lead to the needs of new adequate ICN models forming [12, 13]. For adequately describe processes in ICN it starts to widely use models with constraints on buffer devices, adaptive traffic distribution models, models with priority dispatching of incoming queries, multichannel service models [9, 12]. However, all of the above models, describing real traffic in network nodes, have an ultra-high dimension, which makes the problem solving as very time-consuming or impossible. The paper proposes a technology for accelerating the problem solution using the decomposition approach.

The problem formulation

Let's postulate that the total flow of queries at the node interface, which resulting from the superposition of several flows from different sources, is Markov flow. This assumption is natural and has the following substantiation.

Let's assume that each of the queries flows entering the computer network is a stationary, ordinary random process, and the number of generate query sources is quite large. If the query sources have the same priority, then they have an approximately equal influence on the overall process. Then, similar to the central limit theorem, we can prove that summing (overlapping) of a large number of ordinary stationary flows with practically any aftereffect leads to a flow arbitrarily close to the simplest one. In practice, it is enough to add 4-5 flows to get a flow that can be operated as the simplest one [1].

Let at random time moments from different sources queries arrive to the network input with intensity λ_i, and the queries service intensity by intelligent nodes is equal to μ_j. Let's create a graph of states and transitions for a quite simple system with three servicing nodes, a superposition of two flows comes to the input of it (Fig. 2).

In the resulting graph, the vertices correspond to the states of the system. Each state corresponds to N-dimensional vector, the components of which reflect the queries amount in the queue in front of the corresponding node.

When a new query arrives or an already received one is processing, the system switches to another state corresponding to the queries amount in the queue in front of each node.

To find the final probability distribution of the system, the Kolmogorov differential equations system can be used, based of which a system of linear algebraic equations relative to the sought states probabilities of the system forms. In this case it is significantly that for a real network node and real traffic the problem dimension is very large and its solution is associated with serious computational problems.

The states and transitions graph for the simplest network (Fig. 2) illustrates this remark well.

Due to this, the work goal is to substantiate a practical methodology for solving the problem of analyzing service quality indicators for network nodes, the model of which has a large dimension.

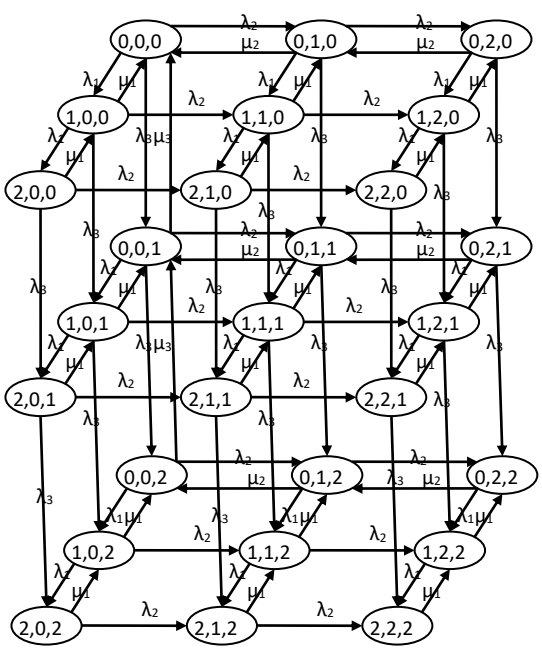

Fig. 2. The states and transitions graph of the network functioning Markov model with three independent incoming flows of queries, when the maximum number of each type queries in the system is two

Analysis method for high dimensional Markov systems

To solve the large dimension problems, the phase consolidation technology of the Markov chain states, which implements the decomposition approach, can be applied [2]. The essence of the method is as follows. The problem solution is cyclical. Each cycle contains a number of iterations. Let's consider the first cycle content. The set of all system states is divided into subsets, and all subsets, except one, are upsized. At the first iteration the first subset remains non-upsized. As a result, the problem dimension is significantly reduced, which allows to calculate the states probabilities for such a system (Fig. 3, a). At the next iteration the second subset is downsized and all others are upsized. The states probabilities are calculated already for such a system (Fig. 3, b). The probabilities of the system stay in each upsized states are equal to the sum of the probabilities of being in the states of the corresponding subset. Thus, if the upsized state I_q contains states with numbers $i_{q1}, i_{q2}, ..., i_{qn_4}$, then the probability of the system being in this upsized state $\pi_{I_4}^{(k)}$, calculated at the next step, is calculated by the formula

$$\pi_{I_4}^{(k)} = \sum_{s=1}^{n_4} \pi_{qs}^{(k)}, \; q = 1, 2, 3, ..., \ell - 1, \ell + 1, ..., m. \quad (1)$$

In this case we obtain a system with a set of possible states

$$\{I_1, I_2, ..., I_{\ell-1}, i_{\ell 1}, i_{\ell 2}, ..., i_{\ell n_1}, I_{\ell+1}, ..., I_m\}$$

and the probability distribution of the system being in these states

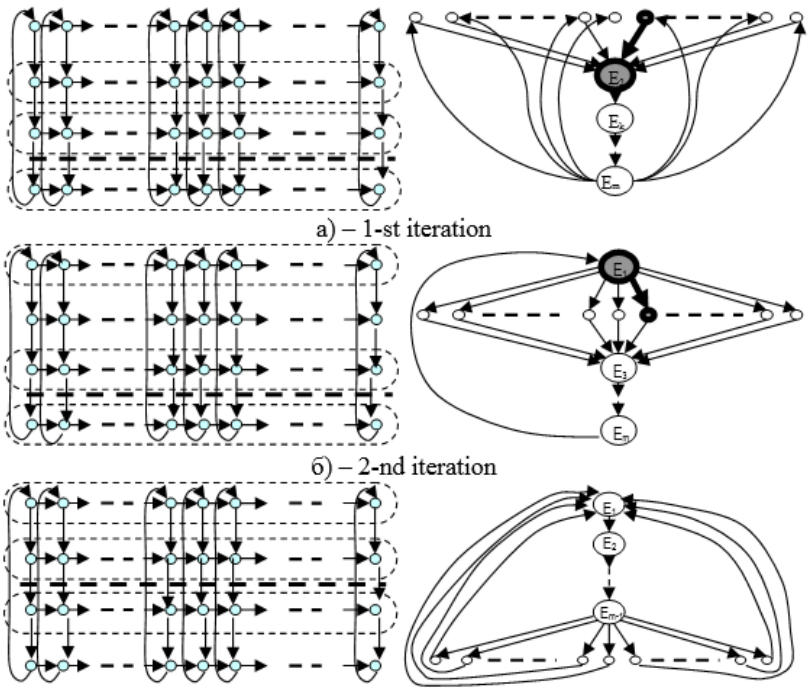

Fig. 3. Phase coarsening of states method

$$\tilde{\pi}^{(k)} = \left(\pi_{I_1}^{(k)}, \pi_{I_2}^{(k)}, ..., \pi_{I_{\ell-1}}^{(k)}, \pi_{\ell 1}^{(k)}, \pi_{\ell 2}^{(k)}, ..., \pi_{\ell n_\ell}^{(k)}, \pi_{I_{\ell+1}}^{(k)}, ..., \pi_{I_m}^{(k)}\right).$$

This distribution will be called group distribution.

Let's calculate the transition probabilities in this new system. The numerical value of these probabilities depends on the type of the corresponding states.

The transition probabilities from the states of the selected subset ℓ to the states of the same subset remain the same as they were in the original matrix W, that is

$$\tilde{p}_{\ell s_1, \ell s_2}^{(k)} = p_{\ell s_1, \ell s_2}, \quad (\ell s_1, \ell s_2) \in E_\ell .$$

The transition probability from any non-upsized state $i_{\ell s}$ to any upsized state I_v is equal to the sum of the true transitions probabilities from this non-upsized state $i_{\ell s}$ to all states of the subset E_v, that is

$$\tilde{p}_{\ell s, I_V}^{(k)} = \sum_{q=1}^{n_V} p_{\ell s, vq}, \quad \ell s \in E_\ell, \quad v = 1, 2, ..., \ell-1, \ell+1, ..., m. \tag{2}$$

The transition probability from the upsized state I_v to some non-upsized state $i_{\ell s}$ is equal to the sum of the products of the true transition probabilities $p_{vq,\ell s}$ from the states of subsets E_v to the state $i_{\ell s}$ by conditional probabilities $W_{vq}^{(k)}$ of the system stay in these states of the subset E_v. In turn of, this conditional probability $W_{vq}^{(k)}$ of the system stay in the state i_{vq} of subset E_v is calculated as the ratio of the probability of being in this state, corresponding to the next obtained complete distribution, to the sum of such probabilities for all states of the subset. Thus

$$\tilde{p}_{Iv,\ell s}^{(k)} = \sum_{q=1}^{n_v} p_{vq,\ell s} \cdot W_{Vq}^{(k)} = \sum_{q=1}^{n_v} p_{vq,\ell s} \cdot \pi_{vq}^{(k)} \Big/ \sum_{q=1}^{n_v} \pi_{vq}^{(k)} . \tag{3}$$

Finally, the transition probability from any upsized state I_{v_1} to another upsized state I_{v_2} is calculated as the sum of the system transition probabilities from the upsized state I_{v_1} to all states included in the upsized state I_{v_2}, that is

$$\tilde{p}_{Iv_1 Iv_2}^{(k)} = \sum_{q_2=1}^{n_{v_2}} \sum_{q_1=1}^{n_{v_1}} p_{v_1 q_1, v_2 q_2} \cdot \pi_{v_1 q_1}^{(k)} \Big/ \sum_{q_1=1}^{n_{v_1}} \pi_{v_1 q_1}^{(k)} . \tag{4}$$

Thus, using formulas (1) - (4), the next group probability distribution of the system stay on the set of states $\tilde{\pi}^{(k)}$, obtained after the upsizing, and the transition probability matrix corresponding to the next upsizing stage are calculated as

$$\tilde{W}^{(k)} = \left(\tilde{p}_{ij}^{(k)}\right),$$

$$(i, j) \in \left\{I_1, I_2, ...I_{\ell-1}, i_{\ell_1}, i_{\ell_2}, ..., i_{\ell n_\ell}, I_{\ell+1}, ..., I_m\right\}. \tag{5}$$

The preparatory calculations have been completed. Now the new group probability distribution is calculated using the formula

$$\tilde{\pi}^{(k+1)} = \left(\pi_{I_1}^{(k+1)}, \pi_{I_2}^{(k+1)}, ..., \pi_{I_{\ell-1}}^{(k+1)}, \pi_{\ell 1}^{(k+1)}, \pi_{\ell 2}^{(k+1)}, ...,\right.$$

$$, \pi_{\ell n_\ell}^{(k+1)}, \pi_{I_{\ell+1}}^{(k+1)}, ..., \pi_{I_m}^{(k+1)} \Big) = \tilde{\pi}^{(k)} \tilde{W}^{(k)} . \quad (6)$$

The resulting group distribution is used to obtain a new whole distribution

$$\pi^{(k+1)} = \Big(\Big(\pi_{11}^{(k+1)}, \pi_{12}^{(k+1)}, ..., \pi_{1n_1}^{(k+1)} \Big), ..,$$

$$\Big(\pi_{\ell 1}^{(k+1)}, \pi_{\ell 2}^{(k+1)}, ..., \pi_{\ell n_\ell}^{(k+1)} \Big), ..., \Big(\pi_{m1}^{(k+1)}, \pi_{m2}^{(k+1)}, ..., \pi_{mn_m}^{(k+1)} \Big) \Big).$$

The components of this distribution are calculated by the formulas

$$\pi_{\ell s}^{(k+1)} = \tilde{\pi}_{\ell s}^{(k+1)}, s = 1, 2, ..., n_\ell , \quad (7)$$

$$\pi_{vq}^{(k+1)} = \tilde{\pi}_{v}^{(k+1)} \cdot \pi_{vq}^{(k)} \Big/ \sum_{q=1}^{n_v} \pi_{vq}^{(k)} , \quad (8)$$

$$v = 1, 2, ..., \ell - 1, \ell + 1, ..., m, \; q = 1, 2, ..., n_v .$$

The iterations are repeated until the last subset is downsized (Fig. 3, c).

The resulting probability distribution is compared with that one obtained at the end of the previous cycle. The solving process continues until the difference between the probabilities of the system being in states with the same indices in two adjacent cycles is lower than a predetermined small value ε.

The fundamental advantage of the described method is that it allows transforming the original complex high dimension problem to a sequence of significantly lower dimension problems, the solving of each of them no longer creates serious computational problems.

Conclusions

The paper proposes a simple practical method for analyzing multiinput and multichannel computer networks. The method realizes the technology of complex problem decomposition to set of simpler ones. The described method can be used to analyze any high dimensional Markov systems. Further research can be aimed at constructing hierarchical procedures for phase consolidation, the needs for which will arise in the study of ultrahigh dimensional systems.

Список літератури

1. Клейнрок Л. Теория массового обслуживания: пер. с англ. / под ред. В.И. Неймана. Москва: Машиностроение, 1979. 330 с.
2. Клейнрок Л. Вычислительные сети с очередями: пер. с англ. Москва: Мир, 1979. 406 с.
3. Кофман А., Крюон Р. Массовое обслуживание (теория и приложения): пер. с фр. под ред. И.Н. Коваленко. Москва: Мир, 1965. 302 с.
4. Авен О.И., Гурин Н.Н., Коган Я.А. Оценка качества и оптимизация вычислительных систем. Москва: Наука, Главная редакция физико-математической литературы, 1982. 132 с.
5. Пороцкий С. Моделирование алгоритма маршрутизации транспортной АТМ сети. *Электросвязь*. 2000. №10. С. 16-19.
6. Турко С.А., Фомин Л.А., Будко П.А., Зданевич С.Н., Гахова Н.Н. Оптимизация пропускной способности звеньев Ш-ЦСИС при ограниченных сетевых ресурсах. *Электросвязь*. 2002. №2. С. 17-19.
7. Ябных Г.А., Столяров Б.А. Оптимизация информационно вычислительных сетей. Москва: Радио и связь, 1987. 230 с.
8. Якубайтис Э.А. Информационно вычислительные сети. Москва: Финансы и статистика, 1984. 310 с.
9. Будко П.А. Выбор пропускных способностей каналов при синтезе сети связи в условиях изменяющейся нагрузки. *Физика волновых процессов и радиотехнические системы*. 2000. Т. 3, № 3-4. С. 68-72.
10. Мизин И.А., Богатырев В.А., Кулешов А.П. Сети коммутации пакетов / под ред. В.И. Семинихина. Москва: Радио и связь, 1986. 408 с.
11. Морозов В.К., Долганов А.В. Основы теории информационных сетей: учебник. Москва: Высшая школа, 1987. 271 с.
12. Будко П.А., Федоренко В.В. Управление в сетях связи. Математические модели и методы оптимизации: монография. Москва: Изд. физико-математической литературы, 2003. 228 с.
13. Пасечников И.И. Методология анализа и синтеза предельно нагруженных информационных сетей. Монография. Москва: Машиностроение-1, 2004. 216 с.

References

1. Kleinrock, L. (1979), *Queuing Theory*, Mechanical Engineering, Moscow, 330 p.
2. Kleinrock, L. (1979), *Computing networks with queues*, Mir, Moscow, 406 pp.
3. Kofman, A. and Kruon, R. (1965), *Mass Service (Theory and Applications)*, Mir, Moscow, 302 pp.
4. Aven, O.I., Gurin, N.N. and Kogan, Ya.A. (1982), *Assessment of the quality and optimization of computing systems*, Nauka, Moscow, 132 pp.
5. Porotsky, S. (2000), “Modeling the routing algorithm of the transport ATM network”, *Electrosvyaz*, no. 10, pp. 16-19.
6. Turko, S.A., Fomin, L.A., Budko, P.A., Zdanevich, S.N. and Gakhova, N.N. (2002), “Optimization of the bandwidth of SH-ISDN links with limited network resources”, *Electrosvyaz*, No. 2, pp. 17-19.
7. Yabnykh, G.A. and Stolyarov, B.A. (1987), *Optimization of information computer networks*, Radio i svyaz, Moscow, 230 p.
8. Yakubaitis, E.A. (1984), *Information computing networks*, Finance and statistics, Moscow, 310 p.
9. Budko, P.A. (2000), “The choice of channel capacities in the synthesis of a communication network under changing load conditions”, *Physics of wave processes and radio engineering systems*, Vol. 3, No. 3-4, pp. 68-72.
10. Mizin, I.A., Bogatyrev, V.A. and Kuleshov, A.P. (1986), *Packet Switching Networks*, Radio i svyaz, Moscow, 408 p.
11. Morozov, V.K. and Dolganov, A.V. (1987), *Fundamentals of the theory of information networks*, Vysshaya Shkola, Moscow, 271 p.
12. Budko, P.A. and Fedorenko, V.V. (2003), *Management in communication networks. Mathematical models and optimization methods*,:Izd. physico-mathematical literature, Moscow, 228 p.
13. Pasechnikov, I.I. (2004), *Methodology for analysis and synthesis of extremely loaded information networks*, Machine building-1, Moscow, 216 p.

Received (Надійшла) 24.04.2021
Accepted for publication (Прийнята до друку) 28.07.2021

ABOUT THE AUTHORS / ВІДОМОСТІ ПРО АВТОРІВ

Пустовойтов Павло Євгенович – доктор технічних наук, професор, завідувач кафедри "Системи інформації" імені В. О. Кравця, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Pavlo Pustovoitov – Doctor of Technical Sciences, Professor, Head at V. O. Kravets System Information Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine.
e-mail: p.pustovoitov@gmail.com; ORCID ID: https://orcid.org/0000-0003-3884-0200.

Охрименко Максим Юрійович – старший викладач кафедри "Системи інформації" імені В. О. Кравця, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Maxim Okhrimenko - Senior Lecturer at V. O. Kravets System Information Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine.
e-mail: Maksym.Okhrymenko@khpi.edu.ua; ORCID ID: https://orcid.org/0000-0003-1292-748X

Воронець Віталій Миколайович – аспірант кафедри "Системи інформації" імені В.О. Кравця, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Vitalii Voronets– postgraduate student at V. O. Kravets System Information Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine.
e-mail: voronets.vitaliy@gmail.com; ORCID ID: https://orcid.org/0000-0002-7793-3824.

Удалов Дмитро Володимирович – асистент кафедри "Системи інформації" імені В.О. Кравця, Національний технічний університет «Харківський політехнічний інститут», Харків, Україна;
Dmitry Udalov – assistant at V. O. Kravets System Information Department, National Technical University "Kharkiv Polytechnic Institute", Kharkiv, Ukraine.
e-mail: dmytro.v.udalov@gmail.com; ORCID ID: https://orcid.org/0000-0003-2586-5513.

Метод підвищення швидкості модельного вузла Маркова

П. Є. Пустовойтов, М. Ю. Охрименко, В. М. Воронець, Д. В. Удалов

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

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

Метод повышения скорости модельного узла Маркова

П. Е. Пустовойтов, М. Ю. Охрименко, В. М. Воронец, Д. В. Удалов

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

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