Date of publication DEC-31, 2023, date of current version NOV-13, 2023.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.22.4.3355

Accelerating Image Classification based on a Model for Estimating Descriptor-to-Class Distance

VOLODYMYR GOROKHOVATSKYI[1], SVITLANA GADETSKA[2], NATALIA STIAHLYK[3]
[1]Kharkiv National University of Radio Electronics, Kharkiv, Prospect Nauki, 14, 61166, Kharkiv, Ukraine
[2]Kharkiv National Automobile and Highway University, Yaroslava Mudrogo str., 25, 61002, Kharkiv, Ukraine
[3]V. N. Karazin Kharkiv National University, Svobody Sq., 4, 61022, Kharkiv, Ukraine

Corresponding author: Svitlana Gadetska (e-mail: svgadetska@ukr.net).

ABSTRACT The article describes a method of image classification based on the estimation of the distance to the etalon class. The implementation of estimates gives a significant gain in classification speed compared to linear search while maintaining a decent level of accuracy. The methodology is based on the use of the triangle inequality for images given by a set of binary vectors as descriptors of the image key points. The evaluation is applied to the "object descriptor – etalon" classification method, which is based on the descriptor voting procedure. An analysis of evaluation options is carried out using the parameters of the etalon sets in the form of a medoid and the closest or farthest points from it. The gain in classification time compared to the traditional method proportionally depends on the number of descriptors in the etalon description. Software simulation of classifiers with the implementation of evaluation shows a gain in speed of 350-450 times for the description of 500 descriptors while maintaining one hundred percent classification accuracy on the training set of similar NFT images. A control sample experiment shows that the classifier with estimation can respond better to image details compared to the traditional method.

KEYWORDS image classification; keypoint descriptor; distance estimation; classification speed.

I. INTRODUCTION

NOW there is an intensive development and implementation of computer vision systems, which require the creation and research of new effective methods of intellectual analysis and processing of multidimensional data [1-8]. One of the urgent applied problems is the implementation of computer vision systems in conditions of limited resources: in unmanned aerial vehicles, mobile devices, robotic and satellite systems. Achieving the speed of real-time operation while maintaining satisfactory performance is a challenge currently faced by researchers [1, 6, 8-11].

Taking into account the significant volumes of multidimensional data analyzed and processed in computer vision, it becomes vitally necessary to use computing resources more efficiently by admissible cost reduction for class determination or recognition of visual objects [8, 12, 13]. Thus, in methods of image classification by description in the form of a set of multidimensional descriptors, the number of elements reaches 500 or more, and the dimension of each descriptor is up to 512 binary components [3, 6, 27]. As a rule, the latest recognition models have a combinatorial sense, as they include optimal metric or statistical procedures for searching for the relevance of input and reference data, which is often associated with the implementation of linear search approaches both on a set of classes and within the class description [4, 9, 25].

The application of approaches with accelerated processing is based on the use of complete a priori information about descriptions from the etalon database, on the basis of which training or estimation of classifier parameters is practically carried out, and some transformed data space is also determined, that helps to reduce the necessary computational costs [6, 11-13].

The implementation of the discussed transformations in the practice of classification (categorization) of images means the transition to a new modified data structure by transforming or quantizing the existing feature space in order to simplify processing and ensure the necessary speed. Such transformations directly affect the effectiveness of the classifier, so they require in-depth study for the existing feature systems.

Note that an important target requirement for the implementation of any new classifier models, including the use of accelerated search tools, is to ensure a sufficient level of classification effectiveness within a fixed base of categories (etalons), taking into account the influence of existing external factors that may interfere with the decision-making process [5, 6, 14, 15].

II. PROBLEM STATEMENT

In structural methods of image classification, the description Z of a visual object is presented in the form of a finite set $Z = \{Z_v\}_{v=1}^{s}$ of s key points of the descriptors (KP). The descriptor z_v is a numerical vector of dimension n [6, 10, 29]. The descriptions of the object and etalons are finite sets of multidimensional vectors.

Let the set of etalon base $E = \bigcup_{i=1}^{N} E_i$ be given as the union of N descriptions that construct the set of N recognized classes. In fact, E is the aggregate set of vectors-descriptors of the composition of all etalons, $E = \{E_i\}_{i=1}^{N} = \{\{e_v(i)\}_{v=1}^{s}\}_{i=1}^{N}$, where i is the class number, v is the current number of the element within the class, and s is the fixed number of descriptors in each of the etalons.

The traditional formulation of the problem of image classification by description in the form of a set of key point descriptors is reduced to determining the relevance degree of two sets of multidimensional vectors and optimizing this criterion on the available set of etalons [6, 20].

At the same time, one of the most effective practical approaches is the "object descriptor - etalon" classification method, which is based on the definition of class parameter for descriptors of key points [6]. For this, the distances $\rho(z, E_i)$ from each descriptor $z \in Z$ of the analyzed object Z to each of the sets E_i, $i \in \{1,2,\dots,N\}$ of descriptors of etalons are calculated. Based on the minimization of the obtained distance, the etalon class to which the descriptor most likely belongs is determined. The next stages are the calculation of the accumulated number of votes for the descriptors assigned to each of the etalons E_i, on the basis of which the number of votes is optimized for the set of classes $i \in \{1,2,\dots,N\}$ and the class of the analyzed object is determined [6, 23].

In this approach, the key from the point of view of the calculation speed criterion is the implementation of the rule $z \to [1,2,\dots,N]$ for classification of an arbitrary vector $z \in Z$ to one of N classes by calculating the distance ρ from the object element $z \in Z$ to the class E_i.

$$\rho(z, E_i) = \min_{v=1,\dots,s} \rho(z, e_v(i)). \qquad (1)$$

The model (1) implements the search according to the "nearest neighbor" principle and determines the object closest in distance to the query [13].

In models of type (1), as a rule, when performing the classification, a limit restriction δ_ρ is additionally used for the value of the minimum distance between the descriptors. If the condition $\rho(z, E_i) \le \delta_\rho$ is fulfilled for the calculated minimum $\rho(z, E_i)$, then the class of the evaluated descriptor is considered to be defined. Otherwise, the descriptor z is classified as false. Such logical filtering contributes to the removal of outliers [3, 10].

Taking into account the fact that the number s of descriptors in the description E_i is calculated by several hundred (s=500 or more), and the sequential analysis of the entire set of etalons multiplicatively increases the required amount of calculations, an important practical task is the introduction of means of reducing calculations in the implementation of (1), for example, by estimating the distance (1) with full use of the available information and classification conditions, under which the descriptions of etalons E_i are sets of numerical vectors and considered to be given a priori [6, 19, 27].

III. LITERATURE REVIEW

Today, in order to reduce the available volumes of calculations in image classifiers, hashing and clustering tools are being successfully developed. These tools group the analyzed data by identifying and using their common features or cluster centers [2, 5-8]. As a result of the development, high-speed hierarchical methods were obtained. At the first stage of processing they determine whether the query belongs to a certain segment of data, and at the second stage they perform a full linear search within the subset defined at the first stage [6-12].

If the descriptions for different classes are sufficiently naturally grouped around their own data centers, then the direct use of these centers as significant (integrated by description) parameters for classes can be effective in terms of processing speed [9, 20–23]. Such centers can be regarded as estimates for class descriptions. Despite some idealization of these assumptions, these centers, for example, in the form of a set medoid [3, 24], showed sufficient computational and classification efficiency [20, 23, 25].

But due to the more complete use of a priori knowledge and the reduction of the degree of integration, the introduction of approximate estimates of belonging to a class (as an etalon set of data) may be more effective on the basis of conducting an in-depth statistical analysis for the base of etalon descriptions, in particular, on the set of distance values between descriptors within the reference description [5, 12, 13]. Such methods are based on estimating the distance from an element to a set using the triangle inequality in a numerical metric space. They are described by the term "metric indexing" and are used to organize effective data search or classification in image and video databases [12-15]. The index here means a specially created data structure to speed up the search. Such structures are used in numerous search ways like method of nearest neighbor, method of k nearest neighbors, using the elements with similarity within a range of values, etc.

In particular, distance estimation was implemented in search methods based on the features of segmented images, which accelerated the search and classification due to the purposeful exclusion from the analysis of the sets that are distant (on the evaluation result) from the image query [5, 15].

Researchers note that with significant dimensions of the feature space (more than 20), even search systems based on the sufficiently developed structures - trees lose their effectiveness [12, 16]. At the same time, the means of establishing the equivalence of images on the basis of structural information, implemented on applied datasets, proved their effectiveness even when using a small subset of the description [20, 29].

The implementation of evaluation models provides a promising opportunity to simplify and accelerate the implementation of methods of applied classification of visual objects along with such already sufficiently developed approaches as hashing and clustering. The main common idea of such specialized data transformations (estimation, hashing, clustering) is to introduce approximate and granulation methods of analysis in the used data search procedures instead of full-scale linear search, which becomes practically impossible when processing multidimensional data arrays.

Applied studies of hashing and clustering applications for descriptions of images showed a tenfold increase in performance [6, 10, 14, 19].

All three analyzed approaches are within the framework of a single concept of data management tools in the recognition process [18], but they have their own specific features when performing intellectual analysis.

Other methods of accelerating classification, which use methods of forming a compressed volume of features based on the values of weighting criteria for the classifier, are also implemented at the stage of preliminary study of data and can be carried out independently of evaluation or transformation in the classification process [19-22].

The introduction of a hierarchical presentation of descriptions by granulating the values of lower-level features also speeds up the calculation due to some reduction in classification accuracy [23]. Another method of acceleration is the determination of the class based on the result of the analysis of the value of the measure within the given threshold of accuracy (equivalence). Such approaches belong to the group of random search tools and require research on determining the threshold [22, 23, 26].

It is important to study the effectiveness of the entire range of means of increasing computational efficiency in classification methods, since the choice of one or another of them may depend on the type of presentation of the analyzed data, the applied measures of relevance, and the requirements for the applied purpose of the classifier. The second important aspect is the study of applied features regarding the methods of accelerating calculations, given that in practice, as a rule, the simplest variants of processing models are used.

IV. METHODOLOGY

A. DISTANCE ESTIMATION FOR CLASSIFICATION

To speed up calculation according to (1) within the etalon description, we introduce a method based on determining the estimate $\tilde{\rho}$ for distance (1) in the form $\tilde{\rho}(\widetilde{z, E_\iota}) = F(z, E_i, \rho)$, where F is a rule for forming an estimate $\tilde{\rho}$, the result of which application depends parametrically on the values of the analyzed descriptor, the set of descriptors of the description, and the type of distance ρ.

Let us consider the formal representation of the rule F using the properties of the sides of a triangle on the plane [12, 13, 28]. Let a, b, c be the lengths of the sides (Fig. 1). Then for each of them, for example, for a, the following conditions are valid, resulting from the triangle inequality:

$$c - b \le a \le c + b. \tag{2}$$

We apply estimate (2) in the metric space of multidimensional vectors (KP descriptors). Let us denote $e^*(i) \in E_i$ as some fixed point of the set E_i and consider a point $d(i) \in E_i^*$ belonging to the set $E_i^* = E_i \backslash e^*(i)$ with the exception of $e^*(i)$. Then the distance $a_i = \rho(z, d(i))$ from the point z of the object to the point $d(i)$ of the set can be estimated using the calculated distance $b_i = \rho(z, e^*(i))$ and the predetermined distance $c_i = \rho(d(i), e^*(i))$ (Fig. 1) [12, 28]:

$$\rho\big(d(i), e^*(i)\big) - \rho\big(z, e^*(i)\big) \le \rho\big(z, d(i)\big) \le \rho\big(d(i), e^*(i)\big) + \rho\big(z, e^*(i)\big) \tag{3}$$

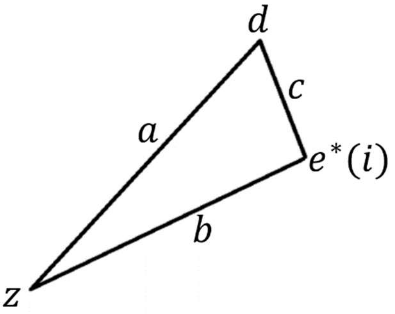

Figure 1. To the geometric interpretation of inequality (3)

Let us introduce the notation: $c_{i,min} = \min_{d(i)\in E_i^*} \rho\big(d(i), e^*(i)\big)$, $c_{i,max} = \max_{d(i)\in E_i^*} \rho\big(d(i), e^*(i)\big)$.

Considering the need to obtain the most accurate estimate, which is determined by the shortest interval size in inequality (3), we purposefully choose the largest value of the left-hand side of (3) among the points of the set and the smallest value of its right-hand side. As a result, we have the estimate:

$$c_{i,max} - b_i \le \rho(z, E_i) \le c_{i,min} + b_i. \tag{4}$$

Note that the result of estimates (3) and (4) parametrically depends on both the content of the set E_i and the selected point $e^*(i) \in E_i$ in the set E_i. The estimation result is some numerical value ρ_i for the distance (1) from the descriptor to the class with the number i.

It is appropriate to choose as a medoid parameter [20, 24] the point $e^*(i)$ of the set with the minimum total distance to the rest of the points. The medoid is calculated according to a simple scheme, it is an element of the set, and it can be universally applied with arbitrary dimensionality of the data. It is known that on the basis of medoid the new classification features for a set of points as a structural description of the object can be effectively built [20, 23].

Components $c_{i,min}$, $c_{i,max}$ of the estimates (4) can be obtained directly at the stage of preliminary analysis of available etalon data, therefore, the amount of calculations in the classification process is not affected by their determination. To obtain the current estimate ρ_i, it is only necessary to calculate the distance b_i from the query to the points $e^*(i) \in E_i, i = 1, \dots, N$.

Based on the result of applying more general estimate (3) for each descriptor of the object Z, the classifier performs an analysis procedure according to some model A

$$k = \underset{i=1,\dots,N}{opt} A\{c_{i,max} - b_i \le \rho(z, E_i) \le c_{i,min} + b_i\}, \tag{5}$$

which generalizes the value of the estimates obtained for the components of the database of the etalons $\{E_i\}$ and determines the etalon with the best estimate.

One of the schemes for building A is as follows. Let us implement the model A in the form of the optimal choice of the descriptor class as the determination of the minimum among the obtained estimates from above as:

$$k = arg \min_{i=1,\dots,N} \big(c_{i,min} + b_i\big). \tag{6}$$

In the case of an ambiguous determining of k according to formula (6), if the equivalent minimum is reached for several classes at the same time, we will make the final choice of the descriptor class by determining the maximum among the calculated estimates from below (the left-hand side of inequality (4)):

$$k = arg \max_{i=1,\dots,N} \left(c_{i,max} - b_i\right). \quad (7)$$

To determine the class of a whole object Z based on its composition of s components, we introduce a vector $\{h_i\}_{i=1}^{N}$ with integer values, where we will accumulate the received class numbers (votes) for each component $z \in Z$. Based on the introduction of a local classifier for each $z \in Z$ according to (5), we now determine the class number k and increment the accumulator $h_k = h_k + 1$ for the corresponding class number.

According to the result of processing the description of the object Z, we accumulate a vector $\{h_i\}_{i=1}^{N}$. The class of the object is traditionally defined as an argument from the maximum number of votes

$$k = arg \max_{i=1,\dots,N} h_i \;||\; h_k \geq \delta_h, \quad (8)$$

where δ_h is the threshold for the marginal minimum number of votes, which is set experimentally for the given database. If the inequality $h_k \geq \delta_h$ in (8) does not hold, then the class of the object is considered to be undefined (refusal of classification).

The gain from the introduced innovation (3), (5) instead of the linear search scheme on a set of s components is proportional to the increase in the value of s, which can reach 500 or more. The time cost is only in providing the estimate, and the accumulation of votes and the determination of the class according to (8) is the same for the traditional and modified approaches.

Estimation (1) and the result of its implementation for classification can be generalized by providing the following methods.

1. Using two or more points instead of one special point $e^*(i) \in E_i$ for the set E_i.
2. Using for one point $e^*(i) \in E_i$ two or more points of the set E_i which are closest or farthest from it.
3. Applying the logical rules to increase the reliability of the estimate when assigning the object descriptor to the etalon [6]. Such rules can be based on the values of set parameters, such as diameter, farthest point, etc. [2, 3].

The discussed approaches are aimed at expanding the ways based on the points of the set, by which the estimating can be carried out. However, their use in general complicates processing and reduces the gain in time compared to estimation (3), (5).

B. CRITERIA OF CLASSIFICATION EFFICIENCY

We evaluate the effectiveness of the classification method by the accuracy indicator pr, which is calculated by the ratio of the number of correctly classified objects to the total number of them used in the experiment [3].

$$pr = r_p / r. \quad (9)$$

Indicator (9) will be considered in two senses: both as a value pr_1 in relation to sets E_i of etalon descriptors, and also as a more important value pr_2 in relation to a complete description of the object with the assignment of class numbers $i = 1, \dots, N$. It is clear that these criteria are related to each other, since as pr_1 increases, then pr_2 increases, and vice versa. But in models of classification by a set of descriptors a quite high level of pr_2 is often observed even with insignificant levels of pr_1. A collective decision, as a rule, is more effective [23, 30].

One of the confidence criteria for classification using a voting apparatus is the relative value of the excess of the maximum $h_{max\,1}$ of number of the winning class votes over the nearest maximum $h_{max\,2}$ for another class

$$\Delta = [h_{max\,1} - h_{max\,2}] / h_{max\,1}, \quad (10)$$

which can be expressed as a fraction of the value $h_{max\,1}$ [3, 23]. Another option of (10) is normalization for the maximum possible number of votes, which is equal to the power of the description. The indicator Δ varies from 0 to 1 and shows the level of confidence for the classification decision.

V RESULTS AND EXPERIMENTS

A. COMPUTATION EXAMPLE

Let us analyze the effectiveness of the method on a demonstrative example. Let us consider simulated descriptions of three etalons represented by five binary 8-component descriptors (Table 1). The columns of Table 1 are taken as fragments from ORB descriptors for real images [6].

For each etalon, as a selected point $e^*(i)$, we take the medoid as a descriptor with the minimum sum of distances to the rest of the components of the description [24]. We choose Hamming distance (the number of non-matching bits) as a metric for descriptors.

Table 1. Input data for 3 etalons

Descriptors and medoid of the etalon E1					Descriptors and medoid of the etalon E2					Descriptors and medoid of the etalon E3				
$d_1(1)$	$d_2(1)$	$d_3(1)$	$e^*(1)$ $d_4(1)$	$d_5(1)$	$d_1(2)$	$d_2(2)$	$d_3(2)$	$d_4(2)$	$e^*(2)$ $d_5(2)$	$d_1(3)$	$e^*(3)$ $d_2(3)$	$d_3(3)$	$d_4(3)$	$d_5(3)$
1	1	1	1	1	0	0	0	0	0	0	0	1	0	0
0	0	0	0	1	0	1	0	1	1	0	1	1	1	1
1	1	0	0	1	0	0	1	1	0	0	0	0	0	1
1	0	1	0	0	1	1	1	0	1	1	1	1	0	1
0	0	1	1	1	1	1	0	1	1	0	0	0	0	0
1	1	0	1	1	1	1	0	1	1	0	1	1	1	0
1	1	0	1	1	1	0	1	0	1	1	0	1	0	0
1	0	0	1	1	0	0	0	0	0	1	1	0	1	0

As objects for classification, we successively take the sets of etalon vectors E1, E2, E3 as a training sample.

According to classifier A, we calculate the value $c_{i,min} + b_i, i = 1,2,3$, which is the right-hand side of the double inequality (5), taking into account that $c_{1,min} = 2$, $c_{2,min} = 1$, $c_{3,min} = 1$.

According to the results of calculations for 14 of 15 available descriptors, the application of the upper estimate (6) for the classifier A in (5) accurately determines the etalon number. And only for $d_1(1)$, the minimum value of the expression $c_{i,min} + b_i, i = 1,2,3$, is reached for two etalons at once. After applying the lower estimate, we have $c_{1,max} = 4$, $c_{3,max} = 3$ and $c_{1,max} - b_1 = 1$, $c_{3,max} - b_3 = 0$ therefore, according to (7), we have the final determination of belonging $d_1(1)$ to the etalon E_1.

The presented demo calculation for the training set confirms the effectiveness of the application of the proposed estimate when implementing the classification. The accuracy indicators pr_1 and pr_2 here are equal to 1.

B. ANALYSIS OF COMPUTER SIMULATION RESULTS

To perform modeling BRISK descriptors were applied based on the library OpenCV using NET 6 application and auxiliary package Emgu.CV [6, 30, 31].

For evaluation, medoids are defined as selected points for descriptions of etalons using a distance matrix. For the experiment, three etalons of NFT images with a size of 540x540 pixels were selected (Fig. 2), and the fourth and fifth images were used as recognized objects (Fig. 3). We can see that visually the etalons are significantly similar to each other, and individual fragments of the objects are quite similar to the etalons. It was done specifically to analyze the marginal capabilities of the estimate classifier.

Figure 2. Etalon images in the experiment

Figure 3. Images of recognized objects and KP coordinates

500 descriptors are selected in each description; the Hamming metric is chosen to calculate the relevance of a pair of BRISK descriptors. Descriptors are considered equivalent if the value of the metric for them is less than 512 *0.25 = 128. The experiment confirmed the significant similarity of the images in Fig. 2 by description in the form of a set of descriptors. The number of equivalent elements with the fixed threshold is 1:2 – 310, 1:3 – 347, 2:3 – 426 for pairs of the etalons. Value $\Delta = (500 - 426)/500 = 0.15$. Here, the parameter Δ characterizes the similarity of the etalon data among themselves. We can see that more than 60% of the number of the first etalon elements is equivalent to the description of the rest in the database. A special similarity is observed between etalons 2 and 3.

The value δ_ρ of the threshold for the equivalence of descriptors has a significant impact on the classification result. With the value $\delta_\rho = 64$, the experiment shows: for 1:2 – 140, 1:3 – 169, 2:3 – 141, and the value $\Delta = 0.66$, which indicates a much more reliable difference in the descriptions of the training sample. In general, the value δ_ρ of the threshold is determined using the existing base of the etalon descriptions and the possible impact of the outliers.

The procedure for voting the components of the first etalon with the other etalons by linear search and using estimate models (6), (7) was implemented. The spectrum of class votes when compared with the first etalon was: (500, 0, 0) - for linear search, (297, 155, 48) – for the estimate model (6) according to the minimum, (292, 158, 50) – for the estimate model (7) by the maximum. At the same time, the classification time estimate was 4566 ms for linear search and 13 ms for search based on estimation (three and a half hundred times less!).

The experiment shows that the high classification accuracy for the training set based on the estimate is kept, as the indicator $pr_2 = 1$ has been obtained, while pr_1 slightly decreases and reaches the value 0.58. However, with computer estimation, it is possible to significantly reduce the time of classification by 350 times! A similar situation was observed in experiments with other descriptions of etalons. Due to fluctuations in performance for a computer processor with built-in overclocking, the gain for the modified method reached more than 450 times in some experiments!

If we compare the efficiency of estimation according to models (6) and (7), then the processing time for them is almost the same, while the application of estimate (7) according to the maximum gives a slightly higher accuracy rate pr_1.

The experiment with object image descriptions (Fig. 3, control sample) shows the following. For the first image based on the linear search method with the minimum significance check, the votes of the elements are distributed uniformly (approximately) between the classes (148, 118, 167), that is, the class of the first object is considered uncertain due to the low value of the indicator Δ.

At the same time, the classifier using estimation (6), (7) shows a persuasive assignment of the object to the first class with indicators $pr_1 = 0.50$, $\Delta = 0.36$ for (6) and $pr_1 = 0.70$, $\Delta = 0.69$ for (7). Specifically, the range of votes was (249, 160, 91) for (6) and (347, 107, 46) for (7). As we can see, both evaluation methods independently confirmed the identification of the analyzed object as a first class image. The information extracted from the image description for estimation methods is found to be more powerful than for traditional search. At the same time, a gain in classification time of more than 400 times has been received.

VI. DISCUSSION

In fact, the model (4) is a case of the general model (3), where the side property is applied for two different triangles with sides a, b, c_{min} and a, b, c_{max}. Here, only the parameter b is calculated for the classification (Fig. 1), and the rest of the values are determined at the preparatory stage.

Note that the estimate obtained in accordance with (3) or its variants must also satisfy the boundary condition $\rho(z, E_i) \le \delta_\rho$ of classification.

Formally, in the classification process, it is possible to implement various options for using the branches of inequalities (3) or (4). The classifier can rely separately on only one of the branches, for example, when learning. At the same time, combined solutions are permissible when one of the branches has priority.

As practical experience shows, for a spatial vector massive of data, which is a description of components, it is unlikely that estimates (6) and even estimate (7) will coincide with each other. But in this case, it is formally possible to apply more complex analysis options using a set of special points $e^*(i)$ for E_i or a set $e^*(i)$ of the closest (farthest) points to the points in E_i. However, these analysis options require additional study.

Another method of compatible use of estimation (6), (7) is to organize a two-criterion decision, which is accepted only in the case of determining the same class at the same time for both estimates. As the calculations and experiments have shown, compatible estimation helps to reveal more detailed properties of the image description. Such approaches increase the accuracy of estimation due to the speed reduction. In addition, as in the situation of ambiguous estimate (6), it is possible to use a simpler practical way of assigning the descriptor to one of the ambiguously estimated classes or simultaneously to all classes with the same minimum.

Note that the universal model (3) can be used to estimate the distance to the set in any other case. One of them is using parameters $c_{i,max}$ or $c_{i,min}$ in both branches. Then the sides of only one triangle are analyzed.

At the same time, in our opinion, the option with the parameter $c_{i,max}$ is more informative and productive, as it is based on information about the point of the set that is farthest from the point $e^*(i)$. Here, information of the type "inside or outside the set" can be effectively taken into account for classification. At the same time, this is also a disadvantage, as the estimation and classification results become dependent on this one point. The estimate with $c_{i,min}$ more uses $e^*(i)$ and its nearest neighbor in the set. The result of its implementation is more strongly influenced by the special point $e^*(i)$ as the chosen "center" of the set.

Theoretically, an arbitrary point of the set can be chosen as the special estimation point $e^*(i) \in E_i$. But a thorough analysis shows that in order to ensure acceptable estimation accuracy, it must be some "inside" point. Points with such properties include the measurement of some "center of the set" in the form of a medoid [20, 24], geometric center, center of gravity (average value of the components), midpoint from the diameter of the set, the center of the described multidimensional layer around the points of the set, etc. [3, 11, 25, 28]. Based on our studies with sets of descriptors, we can recommend choosing the point $e^*(i)$ as the medoid of the set.

In the trivial case, if the set of descriptors consists of the same points (vectors), the distances from the descriptors to the etalons are estimated directly by the distances to an arbitrary point of the set.

Thus, according to the results of the experiments, the implementation of the estimation in the form (3) using the models (4)-(7) makes it possible to avoid the complex procedure of spatial linear search in the process of classification and due to this significantly reduce computational costs (in proportion to the volume of the description). To determine the chosen point of the set, we can use the information from the matrix of internal distances, which contains all pairwise distances between the elements of the set [5, 11].

Experiments for the second image of the control sample (Fig. 3) show that all three methods (traditional, using (6), (7)) confidently assigned it to the second class with confidence indicators $\Delta = 0.53$, $\Delta = 0.90$, $\Delta = 0.80$. Specifically, for estimation (6), the number of class votes was (17, 438, 45). At the same time, if the classification was done by a human, visually, then this image would be clearly identified as the first etalon. Human vision would pay more attention to the face, not to the hair! As it can be seen, computer systems using the features of descriptor set analyze images according to completely different principles. An approach to human vision could be achieved here by introducing weighting coefficients of the descriptor importance.

According to the modeling, the use of estimation in the form of models (6), (7) makes it possible not only to speed up the classification by hundreds of times, but also to identify and estimate in more detail the similarity of individual fragments of images on set of classes. Classification decision clearly contributes to the reliability of the classification result by several estimation methods independently.

VII. CONCLUSIONS

The considered estimation method can be universally used in any applications to speed up the calculation of metric values or other measures of relevance (based on the metric) for arbitrary data sets. At the same time, unlike other methods, in practice etalon sets of descriptors can have different power, since the process and result of estimation do not depend on the size of the set.

The proposed innovation in classifiers in the applied aspect can be more effective than clustering, hashing or matching by centers due to more flexible use of a priori knowledge about the values of distances within the set, parameters or spatial boundaries of the set, the ability to control the degree of integration during estimation.

The gain in classification time increases proportionally with the increase in the number of components in the description. Experimentally, for the description of 500 elements, a speed gain of more than 400 times has been achieved.

It is clear that despite the experimental speedup during hashing by tens of times, and using the proposed approach by hundreds of times, the choice of one or another classifier or speedup model should be made based on the results of the analysis of the nature and content of the available image descriptions.

The novelty of the investigation is the method of building and accelerating the functioning of image classifiers based on a description of key points using a set of descriptors on the basis of the implementation of tools and models of metric estimation.

The implementation of estimation significantly simplifies and accelerates the process of determining the class of the image without a significant decrease in the efficiency indicator. The application of descriptor-to-class distance estimation models has given the new opportunities for in-depth detection of image details and multi-criteria decision-making in the classification process.

The practical significance of the research lies in the construction of applied classification models using estimation, confirmation of the workability, high speed and classification effectiveness of the proposed modifications on examples of images, creation of software applications for the implementation of the developed classifiers in computer vision systems.

Prospects may be related to the development of evaluation schemes on a large-scale set of classes, where it is possible to perform a preliminary metric analysis of the data for descriptions of the image database.

VIII. ACKNOWLEDGEMENTS

The authors thank the students of Kharkiv National University of Radio Electronics for their participation in calculations and computer modeling.

References

[1] R. Tymchyshyn, O. Volkov, O. Gospodarchuk, Yu. Bogachuk, "Modern approaches to computer vision," *Control Systems and Computers*, no. 6, pp. 46-73, 2018. https://doi.org/10.15407/usim.2018.06.046.

[2] J. Leskovec, A. Rajaraman, J. D. Ullman, *Mining of Massive Datasets*, Cambridge Univ. Press, New York, NY, USA, 2020, 511 p. https://doi.org/10.1017/9781108684163.

[3] P. Flach, *Machine learning. The Art and Science of Algorithms that Make Sense of Data,* Cambridge University Press, New York, USA, 2012, 409 p. https://doi.org/10.1017/CBO9780511973000.

[4] C. D. Manning, P. Raghavan, and H. Schutze, *Introduction to Information Retrieval*, Cambridge University Press, 2008, 528 p. [Online]. Avai;ab;e at: https://nlp.stanford.edu/IR-book/pdf/irbookonlinereading.pdf

[5] S. Mashtalir, V. Mashtalir, "Spatio-temporal video segmentation," In: V. Mashtalir, I. Ruban, V. Levashenko (Eds), *Advances in Spatio-Temporal Segmentation of Visual Data. Studies in Computational Intelligence*, vol 876. Springer, Cham, 2020, pp. 161-210. https://doi.org/10.1007/978-3-030-35480-0_4.

[6] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, M. Zeghid, "Tools for fast metric data search in structural methods for image classification," *IEEE Access*, vol. 10, pp. 124738-124746, 2022. https://doi.org/10.1109/ACCESS.2022.3225077.

[7] W. Hu, Y. Chen, L. Wu, G. Shi, M. Jian, "Boundary-aware hashing for hamming space retrieval," *Appl. Sci.,* vol. 12, issue 1, p. 508, 2022. https://doi.org/10.3390/app12010508.

[8] A. Babenko, A. Slesarev, A. Chigorin, V. Lempitsky, "Neural codes for image retrieval," *Conference Paper. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8689 LNCS(PART 1), pp. 584-599, 2014. https://doi.org/10.1007/978-3-319-10590-1_38.

[9] B. Ramesh, C. Xiang, T. H. Lee, "Shape classification using invariant features and contextual information in the bag-of-words model," *Pattern Recognition*, vol. 48, pp. 894-906, 2015. https://doi.org/10.1016/j.patcog.2014.09.019.

[10] P. Dhar, S. Guha, "Fish image classification by XgBoost based on Gist and GLCM features," *International Journal of Information Technology and Computer Science (IJITCS)*, vol. 13, no. 4, pp. 17-23, 2021. https://doi.org/10.5815/ijitcs.2021.04.02.

[11] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3nd ed., Cambridge University Press, 2007, 1262 p.

[12] E. Chavez, G. Navarro, "A compact space decomposition for effective metric indexing," *Pattern Recognition Letters*, vol. 26, no. 9, pp. 1363−1376, 2005. https://doi.org/10.1016/j.patrec.2004.11.014.

[13] G. Hjaltason, H. Samet, "Index-driven similarity search in metric spaces," *ACM Trans. on Database Systems (TODS)*, vol. 28, no. 4, pp. 517−580, 2003. https://doi.org/10.1145/958942.958948.

[14] A. Gionis, P. Indyk, R. Motwani, "Similarity search in high dimensions via hashing," *Proceedings of the Intl. Conf. on Very Large Databases*, 1999, pp. 518–529.

[15] D. Kinoshenko, V. Mashtalir, E. Yegorova, V. Vinarsky. "Hierarchical partitions for content image retrieval from large-scale database," *Machine Learning and Data Mining in Pattern Recognition*, Lecture Notes in Artificial Intelligence, Springer-Verlag, vol. 3587, pp. 445−455, 2005. https://doi.org/10.1007/11510888_44.

[16] Y. Liu, D. Zhanga, G. Lua, W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262–282, 2007. https://doi.org/10.1016/j.patcog.2006.04.045.

[17] A. Berman, L. Shapiro, "A flexible image database system for content-based retrieval," *Computer Vision and Image Understanding*, vol. 75, no. 1/2, pp. 175–195, 1999. https://doi.org/10.1006/cviu.1999.0772.

[18] C. Celik, H. Sakir, "Content based image retrieval with sparse representations and local feature descriptors: A comparative study," *Pattern Recognit*, vol. 68, pp. 1–13, 2017. https://doi.org/10.1016/j.patcog.2017.03.006.

[19] M. Ghahremani, Y. Liu, and B. Tiddeman, "FFD: Fast Feature Detector," *IEEE Trans. Image Process*., vol. 30, pp. 1153–1168, 2021. https://doi.org/10.1109/TIP.2020.3042057.

[20] V. Gorokhovatskiy, "Compression of descriptions in the structural image recognition," *Telecommunications and Radio Engineering,* vol. 70, no.

15, pp. 1363–1371, 2011. https://doi.org/10.1615/TelecomRadEng.v70.i15.60.

[21] A. Oliinyk, S. Subbotin, V. Lovkin, O. Blagodariov, T. Zaiko, “The system of criteria for feature informativeness estimation in pattern recognition,” *Radio Electronics, Computer Science, Control*, no. 4, pp. 85–96, 2017. https://doi.org/10.15588/1607-3274-2017-4-10.

[22] H. Kuchuk, A. Podorozhniak, N. Liubchenko, and D. Onishchenko, “System of license plate recognition considering large camera shooting angles,” *Radioelectronic and Computer Systems*, vol. 4, no. 100, pp. 82–91, 2021. https://doi.org/10.32620/reks.2021.4.07.

[23] S. Gadetska, V. Gorokhovatskyi, N. Stiahlyk, N. Vlasenko, “Aggregate parametric representation of image structural description in statistical classification methods,” *CEUR Workshop Proceedings: Computer Modeling and Intelligent Systems (CMIS-2022)*, vol. 3137, pp. 68-77, 2022. https://doi.org/10.32782/cmis/3137-6.

[24] T. Kohonen, *Self-Organizing Maps*, Springer-Verlag, Berlin Heidelberg, 1995. https://doi.org/10.1007/978-3-642-97610-0.

[25] R. Duda, P. Hart, D. Stork, *Pattern Classification*, Wiley, 2000, 738p.

[26] A. Aho, J. Hopcroft, and J. Ullman, *Data Structures and Algorithms*, Delhi, India: Pearson, 2003, 479 p.

[27] S. Leutenegger, M. Chli and R. Y. Siegwart, “BRISK: Binary Robust invariant scalable keypoints,” *Proceedings of the International Conference on Computer Vision*, Barcelona, Spain, November 3-16, 2011, pp. 2548-2555. https://doi.org/10.1109/ICCV.2011.6126542.

[28] M. Khamsi, and W. Kirk, *An Introduction to Metric Spaces and Fixed Point Theory”*, John Wiley & Sons, 2001. https://doi.org/10.1002/9781118033074.ch1.

[29] O. Gorokhovatskyi, O. Peredrii, “Image pair comparison for near-duplicates detection,” *International Journal of Computing*, vol. 22, issue 1, pp. 51-57, 2023. https://doi.org/10.47839/ijc.22.1.2879.

[30] K. Laskhmaiah, S. Murali Krishna, B. Eswara Reddy, “An optimized k-means with density and distance-based clustering algorithm for multidimensional spatial databases,” *International Journal of Computer Network and Information Security (IJCNIS)*, vol. 13, no. 6, pp. 70-82, 2021. https://doi.org/10.5815/ijcnis.2021.06.06.

[31] OpenCV Open Source Computer Vision, [Online]. Available at: https://docs.opencv.org/master/index.html.

***VOLODYMYR GOROKHOVATSKYI** is a Professor at the Department of Informatics, Kharkiv National University of Radio Electronics (NURE), Dr.Sc. degree in artificial intelligence systems and tools. Research interests include artificial intelligence, computer vision, data science, features extraction, pattern and object recognition.*

***SVITLANA GADETSKA** is an Associate Professor at the Department of Higher Mathematics, Kharkiv National Automobile and Highway University, PhD in Physics and Mathematics. Research interests include computer vision systems and mathematical modeling.*

***NATALIA STIAHLYK** is a Head at the Department of Information Technology and Mathematical Modeling, Educational and Research Institute “Karazin Banking Institute” of V. N. Karazin Kharkiv National University, PhD in Pedagogical Sciences. Research interests include programming and software development.*