Tech Science Press

Cluster Representation of the Structural Description of Images for Effective Classification

Yousef Ibrahim Daradkeh[1,*], Volodymyr Gorokhovatskyi[2], Iryna Tvoroshenko[2] and Medien Zeghid[3,4]

[1]Department of Computer Engineering and Networks, College of Engineering at Wadi Addawasir, Prince Sattam Bin Abdulaziz University, Wadi Addawasir, 11991, Saudi Arabia
[2]Department of Informatics, Kharkiv National University of Radio Electronics, Kharkiv, 61166, Ukraine
[3]Electrical Engineering Department, College of Engineering at Wadi Addawasir, Prince Sattam Bin Abdulaziz University, Wadi Addawasir, 11991, Saudi Arabia
[4]Electronics and Micro-Electronics Laboratory, Faculty of Sciences, University of Monastir, Monastir, 5000, Tunisia
*Corresponding Author: Yousef Ibrahim Daradkeh. Email: y.daradkeh@psau.edu.sa
Received: 22 March 2022; Accepted: 18 May 2022

Abstract: The problem of image recognition in the computer vision systems is being studied. The results of the development of efficient classification methods, given the figure of processing speed, based on the analysis of the segment representation of the structural description in the form of a set of descriptors are provided. We propose three versions of the classifier according to the following principles: "object–etalon", "object descriptor–etalon" and "vector description of the object–etalon", which are not similar in level of integration of researched data analysis. The options for constructing clusters over the whole set of descriptions of the etalon database, separately for each of the etalons, as well as the optimal method to compare sets of segment centers for the etalons and object, are implemented. An experimental rating of the efficiency of the created classifiers in terms of productivity, processing time, and classification quality has been realized of the applied. The proposed methods classify the set of etalons without error. We have formed the inference about the efficiency of classification approaches based on segment centers. The time of image processing according to the developed methods is hundreds of times less than according to the traditional one, without reducing the accuracy.

Keywords: Cluster representation; computer vision; description relevance; descriptor; image classification; keypoint; processing speed; vector space

1 Introduction

The key tasks of computer vision are related to the formation of useful conclusions and effective management decisions regarding physical objects of the real world based on the analysis of their images [1–9]. The advantage of the structural description for image analysis is the representation of visual

objects as the set of structural elements, which makes it possible to make effective decisions in the recognition process and provides the necessary resistance to obstacles on the analyzed image [10–16].

The development of structural recognition is associated with the perfecting of classification mechanisms in the description space in the kind of descriptors of image keypoints, as well as with the evaluation of the recognition efficiency on applied samples.

Achieving high-efficiency indicators for classification methods requires solving problems associated with the multidimensional data [17–20].

During implementation methods of structural classification, descriptions of objects are represented as high-dimensional numerical vectors [21,22].

Descriptions using Oriented FAST and Rotated BRIEF (ORB) [23–25] contain 256 bits, and descriptions using Binary Robust Invariant Scalable Keypoints (BRISK) [26] detectors have 512 bits. The number of descriptors reaches 1500 elements. The transformation of the feature space in the form of a cluster system significantly simplifies their implementation [18,21]. The main tool is the cluster data analysis, based on the metric relations on the set of descriptions and the definition of similar classes of the objects and etalons [18,21,27].

Also, the problem is setting the correlation between the parameters of various descriptions [27–29]. These parameters, for example, include the center of the cluster, which can be determined by some approaches. The approaches are the modifications of the mean value and median [30,31]. All similar research is related to the Content-Based Image Retrieval (CBIR) direction [1,32]. This direction is aimed at improving intelligent recognition algorithms in accordance with the hierarchy of representation of information [18].

The structuring of the data description and a cluster representation are mechanisms of generalization for the deep analysis of intellectual-level applied solutions [33–36].

2 Problem Statements

The advantage of grouping by the set of centers of clusters in comparison with the traditional grouping method directly by description descriptors has counted the fact that the relevance is calculated based on their optimal number of indicators [16,19,37]. When the set of image keypoints varies from 200 to 1000 elements, then the number of centers recommended is from 2 to 10 elements of the defined dimension.

The relevance is calculated based on researching all of the etalons, for example, in some tasks it reaches 1500 values; the proposed transformation 1000 times reduces the number of calculations [27].

Such success necessitates the grouping of the data of the researched object and a reduction in the level of fragmentation. This fact is due to the creation and application of main characteristics, namely the set of description centers. In this case, the segmenting of the etalon data is performed at the preliminary step and does not affect the performance of the grouping [18,20].

The basis of structural recognition is a priori set of descriptions of etalon images (class alphabet), represented by the set of structural features namely descriptors.

The image database is represented by the set of $E = \{E_i\}_{i=1}^N$ etalons descriptions.

Let's split the set E: $E = T = \{T_f\}_{f=1}^k$, $T_f \cap T_d = \varnothing$, $T_f \neq \varnothing$.

Let's assume that the elements of $e \in T_f$ are equivalent to each other.

Formally we have two models for the classification system on the existing set of features:

- $\{E_i\}$–for images.
- $\{T_f\}$–for features in the form of cluster representation.

Due to splitting $\{T_f\}_{f=1}^k$, each description $\{E_i\}$ of the etalon is transformed into the integer vector of the following form

$$h[E_i] = \langle h_{i,1}, h_{i,2}, \ldots, h_{i,j}, \ldots, h_{i,k} \rangle, \tag{1}$$

where $h_{i,f} = card\{e \,|\, e \in E_i \,\&\, e \in T_f\}$; k is number of clusters.

Representation Eq. (1) is the image of the etalon obtained as the result of the cluster representation of the whole base, and the finite set of vectors Eq. (1) of the base forms a matrix $H[E] = \{\{h_{if}\}_{i=1}^N\}_{f=1}^k$ of the base description using N samples.

The row of matrix H represents the cluster description of the individual etalon, and the column of matrix shows the content of the cluster as a part of the equivalent elements of different etalons. In the general case, for the most common clustering procedures which use vector data with non-integer components (k-mean, hierarchical classification, etc. [20,21,35]), Euclidean or Manhattan distances can be used.

The purpose of the research is to study and define level the effectiveness of the structural methods of image grouping. The image classification is based on the creation of segmental systems by implementing approaches to finding the relevance of the descriptions using the apparatus of optimal comparison and voting.

The tasks of the given research are:

- To process data representation models to define the relevance of descriptions according to the characteristics of the segment, which is important in terms of increment of data processing speed.
- To study questions on the effectiveness of developing these models using the results of the experimental evaluation of the proposed approaches for the applied image base.

3 Related Works

Modern structural methods for analyzing visual information are based on evaluating the similarity of descriptor descriptions of the analyzed and etalon images, represented as the set of descriptors of keypoints [10,21,24–26]. The value of similarity reflects the degree of response of the binary relation of relevance for two information units as descriptions of visual objects.

The set of structural features in the form of the set of numerical vectors is formed using special filters, namely detectors; in this work, the ORB and BRISK methods are used [25,26]. Detectors provide obtaining descriptor vectors that are invariant to geometric transformations of displacement, rotation, scale [10].

The task of the splitting for obtaining the model of etalons as groups of structural elements is solved in different ways [21,27].

One of the way under research is to use clustering on set E. Cluster splitting can be configured on the data structure that is determined by the base [27].

The modern deep learning methods using specialized processing networks aim to identify meaningful differences in the classified data to achieve the most significant performance indicators for fixed datasets [2–5,29,32]. At the same time, methods under development, which due to the combined

implementation of the principles of local and global data analysis make it possible to significantly simplify the classification process while ensuring the required level of efficiency [12,18–22].

Recently, methods of fuzzy data clustering [38] have started to be applied; including for image analysis tasks [5]. However, our research [15,18,27] is largely focused on the means of traditional clustering with non-overlapping clusters, because the available description data in the kind of multidimensional descriptor vectors are characterized by the significant similarity between each other for different etalons. And fuzzy clustering does not make it possible to correct this problem and improve the classification indicators.

Currently, the number of promising models for image analysis are being implemented, based on combining the advantages and capabilities of global and local data analysis; it contributes to the better identification of distinguishing features, for example, for complicated problems such as vehicle classification and person identification [39,40]. In these systems, specialized networks are used to generate features. It is clear that such approaches are related to the increase in the amount of calculations.

The indisputable advantage of using the vector model Eq. (1) when calculating the relevance of images in terms of calculations is that this method is much (tens of times) less consuming than comparing the image description with the descriptions of etalon sets E_i [21].

The obtained image Eq. (1) of the object can be considered as some combinatorial configuration [21,30,32,41], which is formed from the elements of the base set, using $b = \{b_1, b_2, \ldots, b_k\}$, $b_i \in M_i$ centers, formed in the process of the cluster structure formation. In this configuration, the visual object is represented in the form of the vector Eq. (1), which component is equal to the number of occurrences of the base element (from the feature dictionary) in the object description. The "occurrence rating" term is used for structural features [18]. The process of formation of such changes can be considered as self-formation. The base set b of cluster centers, obtained as a result of clustering, is determined directly by the set of samples $\{E_i\}$ which are the focuses in the unlimited space of images recognized by the computer system.

4 Model of Cluster Representation of Structural Description

Consider base E of descriptions of images of etalons with dimension N:E, $E = \{E_1, E_2, \ldots, E_N\}$. Etalon description E_i is the separate class in the computer vision problem and has the representation of the finite set of keypoints: $E_i = \{e_v(i)\}_{v=1}^s$, s is the number of descriptors [10,27].

Each descriptor $e_v(i)$ characterizes some neighborhood of the keypoint image and is the element of vector space R^n of finite dimension $n : e_v(i) \in R^n$ with integer, binary, or real values. We accept the power of descriptions of etalons to be equivalent to simplify the analysis:

$$card(E_1) = card(E_2) = card(E_3) = card(E_4) = \ldots = card(E_N) = s. \tag{2}$$

This condition can be achieved by using components from a larger set.

Let us apply the mechanism mapping data $E \rightarrow T$ from the set of descriptors to the set T of clusters created according to a definite algorithm [21,27].

The cluster is a definite subset of the image description. Images of E_i etalons are proceeding to M of their non-intersecting $T_k(E_i)$ subsets:

$$E_i = T(E_i) = \cup_{k=1}^M T_k(E_i), \ k = \overline{1, M}, \ T_k(E_i) \cap T_j(E_i) = \varnothing. \tag{3}$$

For set $T_k(E_i)$ of the elements of each cluster, we determine its center parameter $b_{k,i}$, $k = \overline{1, M}$, which is the main feature of the proposed system of clustering for the researched data. Centers and clustering $b_{k,i}$ can be calculated using a wide variety of methods [18,21,27,30,32,34,35,42].

Using the cluster structure, we create image E_i of the etalon as of M disjoint subsets of segments $T_k(E_i)$ with defined centers $b_{k,j}$, $M * N$, which is the number of formed segments and centers for the etalons. The recognizable object is described by the set $Z = \{z_v\}_{v=1}^s$, $z_v \in Z$ are descriptors of keypoints, $s = card\ Z$.

We apply the segment distribution of the set Z using reflection $Z \rightarrow T$; the description will be represented by M segments:

$$Z = T(Z) = \{T_k(Z)\}_{k=1}^M, \ \ T_k(Z) \cap T_j(Z) = \varnothing. \tag{4}$$

For segments $T_k(Z)$, we find the characteristics of centers $b_k(Z)$, which are used in the segmentation process. The number of M segments to be identical for etalons and the input image.

5 Classification Methods

Researched the classification of the model "object–etalon" based on the definition of the relevance values of descriptions [18,21,27]. We will use a sped-up segmentation algorithm based on the calculation of centers b_k of the system of segments an etalon and of the object.

The definition of the relevance of the "object–etalon" is formed as an interrelation of the sets of the center's etalons and of the object with the successive finding of the most relevant candidate.

Let us apply some distance ρ in vector space R^n. The centers of b_k clusters can also be submitted in binary form. In the binary space of vectors, the Hamming metric is more computationally efficient [18]:

$$\rho(b_k, z_v) = \sum_{a=1}^n 1\ (b_{k,a}, z_{v,a}), \tag{5}$$

where function $1\ (b_{k,a}, z_{v,a})$ for comparing bits numbered a for two vectors are equal to 1 if the bits do not match and 0 otherwise.

Let's calculate $M \times M$ of all distances $\rho(T_{k_1}(Z), T_{k_2}(E_i))$, $k_1, k_2 \in [1, 2, \ldots, M]$ between the segment systems of the researching object and of etalons by finding the distances $q = \rho(b_{k_1}(Z), b_{k_2}(E_i))$ between the components of the sets of centers the i-th etalon and of the object and accept them as $\{q_a\}$, $a = 1, 2, \ldots, M^2$.

Using the components of set $\{q_a\}$ based on classical methods, we can find the distances: farthest or nearest neighbor, Hausdorff, bond distance, or their improvement associated with the processing of components q_a or logical research [30,34]. For example, the average bond distance has the interpretation

$$\rho_{av}(T_k(Z), T_k(E_i)) = \frac{1}{M^2} \sum_{a=1}^{M^2} q_a. \tag{6}$$

The improvement for determining the distance is happened by choice of three minimum components of the in advance ranked sample $q^* = q_1 \leq q_2 \leq \ldots \leq q_{M^2}$ for the set of distances [27]:

$$\rho_{3m}(T_k(Z), T_k(E_i)) = \sum_{a=1}^3 q_a^*. \tag{7}$$

Distance Eq. (7) has the attributes of integral metrics and both differential. It is received because of the choice of three unique minima.

As an analog to Eq. (7), we will use the opportunity to add the nearest neighbor distances singly for all of the object centers

$$\rho_{NN}(T_k(Z), T_k(E_i)) = \sum_{a=1}^{M} \min_{j=1,\ldots,M} \rho(b_a(Z), b_j(E_i)). \tag{8}$$

Distance models Eqs. (7) and (8) do not assure that certain a center of the etalon will belong to a definite center of an object. The finding of a center with certain characteristics is using the Hungarian algorithm for the best definition of the optimal center of the etalon to all centers of the objects [27,32].

Let us use the Hungarian algorithm for the successful definition of correlation between different sets of cluster centers b_1, b_2. They were calculated for values Z_1, Z_2 taking into account the possible availability of noise [30,34], which causes deviations from desired parameters.

The result of applying the Hungarian approach is the creation of the optimal matching for the parameters researched sets with the reduction of all costs. It can be evaluated as the sum of the distances between two certain elements in b_1, b_2. The correlation mechanism usually leads to a minimization problem

$$R(x) = \sum_{i=1}^{M} \sum_{j=1}^{M} \rho(b_{1i}, b_{2j})x_{ij} \to \min, \tag{9}$$

where $b_{1i} \in b_1$; $b_{2j} \in b_2$; x_{ij} is a binary feature; $x_{ij} \in \{0, 1\}$, which is 1 if the i-th and j-th elements match.

The solution to task Eq. (9) with restrictions on the conformity of characteristics from the researched sets

$$\begin{cases} \sum_{j=1}^{M} x_{ij} = 1, & \forall i = \overline{1, M}, \\ \sum_{i=1}^{M} x_{ij} = 1, & \forall j = \overline{1, M} \end{cases} \tag{10}$$

optimizes the distance Eq. (9) between sets the researched of centers b_1, b_2.

The number of segment centers in a given task is small $(3 \ldots 5)$. Optimal approaches can be used with minimal claims for processing speed data.

The offered expressions Eqs. (7) and (8) and distances for researching the sets of vectors apply to formed descriptions E_i and Z, but with a greater quantity of calculations.

Segmentation of an object based on the defined distances Eqs. (6)–(8) between considered the centers of the data is performed usually by finding the minimum value among the values of the etalon

$$Z \to E_a : a = \arg \min_{i=1,\ldots,N} \rho(T_k(Z), T_k(E_i)). \tag{11}$$

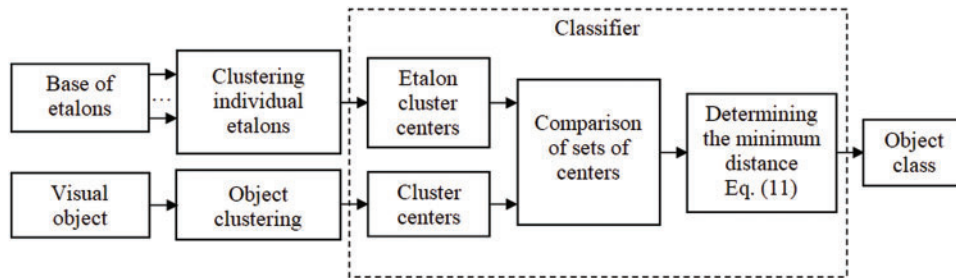We showed the scheme of such a classification method in Fig. 1.

Figure 1: Classification scheme No. 1

An independent segmentation realization for the etalon set simplifies the representation of classification algorithms of the "object descriptor–etalon" variety regarding separate characteristics of the description of the researched object. This method takes into account the possible effects of obstacles on the filtration process or image processing during segmenting and the sudden onset of false keypoints caused by obstacles. In this case, the segmenting of the researched object description is not performed (this considerably minimizes the computational resources).

All object description descriptors have an "individual etalon class" formed in comparison with the calculated set $\{b_j(E_i)\}$ of etalon centers. For all object descriptions $z_v \in Z$ we determine the closest of all etalon centers $\{b_j(E_i)\}$ according to the nearest neighbor procedure

$$d = \arg\min_{i} \rho(z_v, b_j(E_i)), \ d \in \{1, 2, \ldots, N\}, \tag{12}$$

where ρ is the calculated distance between the center and object descriptor $b_j(E_i)$; $i \in \{1, 2, \ldots, N\}$; $j \in \{1, 2, \ldots, M \times N\}$.

Eq. (12) embodies the multivalued function $d : R^n \to \{1, 2, \ldots, N\}$, which forms the etalon class based on an individual descriptor from the researched object description.

Considering the results Eq. (12) $\forall z_v \in Z$ the number of $r_1, r_2, \ldots, r_N$ votes of $z_v \in Z$, elements selected to a certain center $\{b_j(E_i)\}$ of the description E_i is determined:

$$r_i = \sum_{v=1}^{s} F[z_v \to \{b_j(E_i)\}], \tag{13}$$

where F is a logical function that provides the hit a certain of the element z_v to the determined center with the number j of the segment of the etalon E_i. The mechanism for realizing the function F to provide filtering of obstacles takes into account the threshold indicator δ_F for minimum the value of the distances defined for all of the existent centers' $\{b_j(E_i)\}$ segments for each etalon [10,18,27,34].

The researched object's image is processed considering the defined values of $r_1, r_2, \ldots, r_N$ votes:

$$Z \to E_j : j = \arg\max_{i} r_i. \tag{14}$$

According to the classification Eqs. (12)–(14), the researched image will be directed to the etalon that will have the maximum number of votes of keypoint descriptors. Classification scheme No. 2 is shown in Fig. 2.
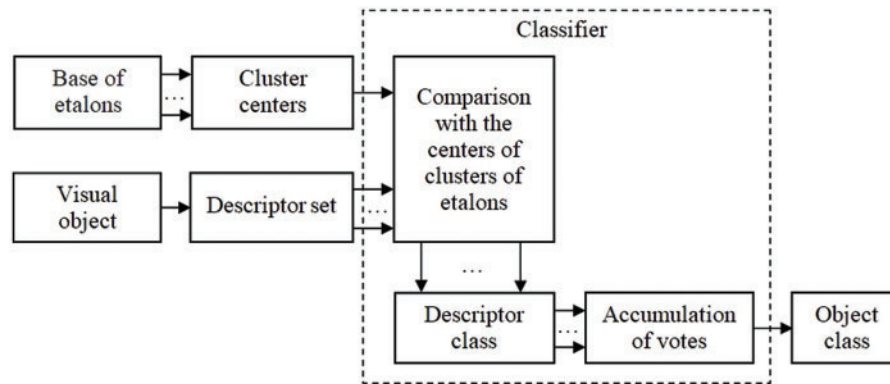
Figure 2: Classification scheme No. 2

Let us now consider model No. 3 of the classifier based on representation Eq. (1). The classification process differs from previous methods in that segmentation is implemented on the full-fledged set of etalon descriptions and centers b_f^* are determined that are common to the entire base of etalons [18,27].

We consider the set of object descriptors $z_v \in Z$ in turn; we assign the current descriptor to cluster T_d under the contention rule

$$z_v \to T_d \ \left| \ \arg\min_f \rho(z_v, b_f^*) = d. \right. \tag{15}$$

Because of the research of the entire set Z, a description of the object $Z = (h_1, h_2, \ldots, h_k)$ is obtained in the form Eq. (1). Further, under model Eq. (11), the class of the object is determined using the vector (for example, Manhattan) distance between the cluster representations of the object and the etalons in the space of integer vectors.

6 Experiments and Results

The proposed classification methods based on the cluster representation have undergone detailed experimental verification on various image bases [10,20,27].

The represented schemes of classifiers are used in the samples of images of Pokemon [27,43]. The software tools were applied–the Java programming language, IDLE with the means of the OpenCV library [44], and IntelliJ IDEA 2020 [45].

The researched images are represented in Fig. 3. ORB descriptors are applied ($n = 256$).

The idea of simulation modeling was implemented using the applied development of two species of classifiers:

- Realization voting descriptors considering the Eqs. (12)–(14).
- By the finding of the best correlation between the centers of the segments Eqs. (9)–(11).

The first algorithm embodies the mechanism of classification "object descriptor–etalon". The second algorithm embodies the mechanism of classification "object–etalon". They apply principles of an independent segment system developed for all etalons.

Figure 3: Examples of analyzed images

The classification was implemented by a method when a set of descriptors (algorithm No. 1) or of centers of the segmental view of certain etalon (algorithm No. 2) were correlated with the set of centers of segments for ten etalons [27,43]. Methods of descriptions of 300 keypoint, 500 keypoint, 1000 keypoint, and 1500 keypoint (with ORB descriptors) were implemented.

The definition of the normalized quantity of votes (in %) for the researched images Fig. 3 on the set of 10 etalons at $s = 500$ are represented in Tab. 1.

Table 1: Generalized classification results based on method No. 1

Input images	Etalons									
	1	2	3	4	5	6	7	8	9	10
3	7.4	8.8	**23.8**	9.6	12.6	7	6.8	9.4	6.6	8
5	11	9.8	10.2	7.2	**29.8**	9.2	3.8	5.8	3.2	10
6	9.6	9.8	8.2	12	9.4	**26**	6.6	4.8	5.6	8

The maximum values are painted with color in Tab. 1.

The first algorithm classifies the researched images, and the superiority of the biggest over the nearby value in the line is confirmed for all considered images. Test results have represented that for the available base of etalons, the quantity of applied descriptors almost does not influence the upshot. All researched variants of keypoints (300, 500, 1000, 1500) received analogous values. The integrated application of the descriptions provides stable sets of segment centers.

The conclusions of classification in the values of the minimal distance with the application the method No. 2 by the Hungarian algorithm to calculate the optimum compliance between the centers of the segments of the researched image Fig. 3 (3 etalons–No. 3, No. 5, No. 6) and an etalon (10 etalons) for 500 keypoints are represented in Tab. 2.

To successfully describe the researched image, segmenting was repeated. The minimum values are painted with color in Tab. 2.

By the modeling conclusions, we observe a successful classification for the researched etalons. The value of the minimum distance defined in keeping with the Hungarian algorithm is much less from all etalons. The etalon No. 6 (row No. 3 of Tab. 2) is detected as the best.

Table 2: Generalized classification results based on method No. 2

Input images	Etalons									
	1	2	3	4	5	6	7	8	9	10
3	212	230	**133**	242	201	197	224	193	179	198
5	248	182	227	276	**133**	221	238	193	199	214
6	232	196	227	248	211	**49**	212	213	183	106

Research of received data (Tabs. 1 and 2) for the processed base of etalons demonstrates much more successful classification based on algorithm No. 1 compared to algorithm No. 2. This fact is substantiated and explained by a particularized research of the data properties in algorithm No. 1. It grounds algorithm No. 2 exclusively on the integral properties of descriptions being the results of segmenting. Testing has proven the ability the successfully using of optimum scientific methods (Hungarian algorithm) for applied tasks where input data are filed in a consolidated segment form. Outright using these algorithms to research hundreds of descriptor vectors does not make sense we observe realization time limitations.

By results of performed modeling, it appreciated the necessary time for using the described approaches. The basic part of the program code implements the finding and description of keypoints based on the ORB detector (using the OpenCV library) [24,25,44].

The number of keypoints does not affect in any way the speed of algorithm No. 2 (its advantage); is taken into account only the number of segments [18,27]. With a magnification in the number of keypoints, the realization time of algorithm No. 1 is not significantly enlarged.

The realization time of classification for algorithms No. 1 and No. 2 fluctuated from 1.1 to 1.5 s for the keypoint scans, the segmenting method, the researched image data, and the equipment applied.

Let us analyze the features of the proposed method No. 3 based on model Eq. (15) using the example of the experimental base of marine mammals (Fig. 4) when presenting their structural descriptions within 10 clusters.

During the research, we identified keypoints for etalon images based on ORB and BRISK (OpenCV and JSFeat libraries), clustering (JavaScript, k-Means library), calculation of the distance matrix for the image database (Fig. 4).

Note that the examined images are similar to each other, and in natural conditions of observation, it is difficult to classify them even for a person. Another feature of this database is a significant visual similarity between the 1st and 4th images (whale and baby whale). Image sizes are 139×309 pixels. The database descriptions were clustered using the k-means method [20,35]. The coordinates of certain keypoints are shown in Fig. 5.

As a result, a description of the base (matrix H of the cluster representation) was obtained in Tab. 3. We showed the normalized matrix of weight coefficients in Tab. 4.

According to Tab. 4, we can build a variety of classification methods, which are based on the accumulation of weights as an analog of the membership function to the corresponding cluster [10,27,34].

Figure 4: Image of marine mammals

Figure 5: Image of keypoints coordinates

As seen, the representation obtained in Tab. 4 looks quite effective. In columns 2, 3, 8, 9, 10, there is a significant predominance of weight coefficients for a few etalons. We can calculate the similarity matrix for the etalon base by multiplying the cluster weight column by the number of structural features of the etalon found in the cluster (Tab. 3).

We presented the distance matrix in Tab. 5.

From Tab. 5, we see significant differences in the etalons in the cluster space, while etalons 1 and 4 remain the most similar.

Table 3: Number of KT etalons in the cluster view

H	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
Z^1	3	6	2	2	5	3	6	8	0	4
Z^2	11	12	0	15	8	14	9	0	0	0
Z^3	3	2	0	14	5	0	18	13	0	0
Z^4	4	1	2	3	4	5	6	1	1	1

Table 4: Weight matrix for marine animals' base

α^1	M_1	M_2	M_3	M_4	M_5	M_6	M_7	M_8	M_9	M_{10}
Z^1	0.18	0.38	0.42	0.08	0.27	0.17	0.19	0.43	0	0.74
Z^2	0.37	0.43	0	0.34	0.24	0.44	0.16	0	0	0
Z^3	0.13	0.09	0	0.40	0.19	0	0.40	0.50	0	0
Z^4	0.33	0.09	0.58	0.17	0.30	0.39	0.26	0.07	1.00	0.26

Table 5: Class distance matrix

Etalons	1	2	3	4
1	0	58	42	21
2	58	0	58	51
3	42	58	0	47
4	21	51	47	0

The key parameter of the cluster representation is the number of clusters related to the equivalence threshold for structural features.

The reduction in the size of computations during recognition takes into account the number of clusters and the composition of the considered set.

For comparison, we present the cluster representation matrix within 5 clusters based on the images of Fig. 4 in Tab. 6.

Table 6: Cluster view for 5 clusters

H	M_1	M_2	M_3	M_4	M_5
Z^1	10	11	3	7	8
Z^2	19	14	0	20	16
Z^3	7	12	4	26	5
Z^4	8	4	3	6	7

The histogram of the representation of the first image (Fig. 4) in the system of 5 and 10 clusters is shown in Fig. 6. The values of the columns (these are the first rows in Tabs. 3 and 6) indicate the number of image descriptors related to the corresponding cluster.
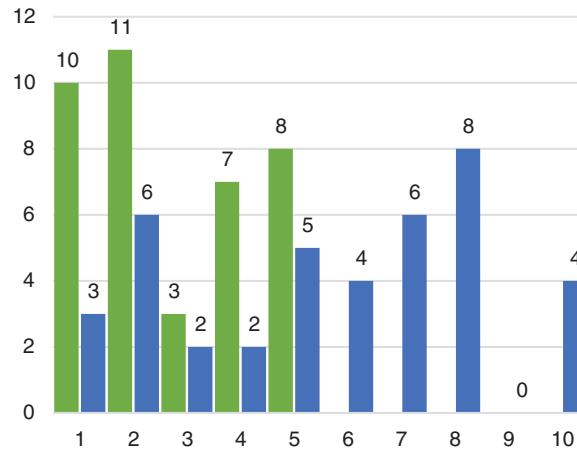

Figure 6: Histogram of image clusters in Fig. 4 (green–5 clusters, blue–10 clusters)

Conducted computer simulation showed error-free recognition of the analyzed visual objects with the number of clusters from 5 to 10, with an increase in clusters (over 8), the accuracy does not improve. The transition to the vector form increases the speed of recognition by simplifying processing. Comparison of classification schemes using different detectors (ORB and BRISK) shows that method No. 3 is more accurate and sensitive; it distinguishes even very similar images.

The described experiments confirmed the effectiveness and efficiency of the researched modifications of algorithms based on segments.

The developed approaches show a significant advantage in the analysis and classification of enormous sets of images with arbitrary sets of vectors, allowing you to speed up the total recognition time by hundreds of times.

7 Conclusions

The principal preference of the algorithm represented in this research is the essential reduction in computational costs for image classification while providing the sufficiently high efficiency indicator.

Application of the segment representation on the set of descriptors promotes perfecting the temporary indicators of the classification at the expense using of segment centers and provides a sufficient level of productivity.

Research and working of collected data allow the development of the hierarchical model with the variable element of detailing and to reveal characteristics of the image description that are meaning for classification.

The developed methods of classification based on the segment description provide a sufficient level of image resolution. Cluster representation reduces the time spent on recognition by dozens of times.

The effectiveness of cluster recognition significantly depends on the resulting system of clusters in the applied basis of the classes specified by the etalons. At the stage of system training, it is advisable to

achieve approximately an equal rendering for the set of structural characteristics of etalons, so that the obtained cluster representation results in an equivalent effect of all structural elements on the result. For similar images (Fig. 4, etalons 1 and 4), it is advisable to create the unified etalon description.

According to the experiments, the recommended number of clusters in applied tasks of compmuter vision should be between 5 . . . 10.

Method parameters such as the threshold for the equivalence of structural features, the threshold for the significance of relevance degree when determining the certain class, as well as the number of clusters are determined in the training process for specific application databases of images.

Perspectives of the further research include the implementation of developed methods for large image databases, and also to the study of their effectiveness by using modern means of fuzzy and multidimensional clustering.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present research.

References

[1] C. Celik and H. S. Bilge, "Content based image retrieval with sparse representations and local feature descriptors: A comparative study," *Pattern Recognition*, vol. 68, pp. 1–13, 2017.
[2] C. B. Murthy, M. F. Hashmi, G. Muhammad and S. A. AlQahtani, "YOLOv2PD: An efficient pedestrian detection algorithm using improved YOLOv2 model," *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3015–3031, 2021.
[3] M. E. Safi and E. I. Abbas, "Robust face recognition algorithm with a minimum datasets," *Diyala Journal of Engineering Sciences*, vol. 14, no. 2, pp. 120–128, 2021.
[4] H. Wu, Q. Liu and X. Liu, "A review on deep learning approaches to image classification and object segmentation," *Computers, Materials & Continua*, vol. 60, no. 2, pp. 575–597, 2019.
[5] Y. Tang, F. Ren and W. Pedrycz, "Fuzzy C-means clustering through SSIM and patch for image segmentation," *Applied Soft Computing*, vol. 87, pp. 105928, 2020.
[6] Y. Xu, Z. Jia, L.-B. Wang, Y. Ai, F. Zhang *et al.,* "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features," *BMC Bioinformatics*, vol. 18, no. 1, pp. 281, 2017.
[7] Y. Zheng, J. Fan, J. Zhang and X. Gao, "Exploiting related and unrelated tasks for hierarchical metric learning and image classification," *IEEE Transactions on Image Processing*, vol. 29, pp. 883–896, 2019.
[8] A. Nasirahmadi and S.-H. M. Ashtiani, "Bag-of-feature model for sweet and bitter almond classification," *Biosystems Engineering*, vol. 156, pp. 51–60, 2017.
[9] X. Tang, M. Zhang, P. Yu, W. Liu, N. Cao *et al.,* "A nonuniform clustering routing algorithm based on an improved k-means algorithm," *Computers, Materials & Continua*, vol. 64, no. 3, pp. 1725–1739, 2020.

[10] Y. I. Daradkeh, I. Tvoroshenko, V. Gorokhovatskyi, L. A. Latiff and N. Ahmad, "Development of effective methods for structural image recognition using the principles of data granulation and apparatus of fuzzy logic," *IEEE Access*, vol. 9, pp. 13417–13428, 2021.

[11] V. O. Gorokhovatskyi, I. S. Tvoroshenko and O. O. Peredrii, "Image classification method modification based on model of logic processing of bit description weights vector," *Telecommunications and Radio Engineering*, vol. 79, no. 1, pp. 59–69, 2020.

[12] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko and M. Al-Dhaifallah, "Classification of images based on a system of hierarchical features," *Computers," Materials & Continua*, vol. 72, no. 1, pp. 1785–1797, 2022.

[13] V. Gorokhovatskyi, N. Rusakova and I. Tvoroshenko, "The application of image analysis methods and predicate logic in applied problems of magnetic monitoring," *Telecommunications and Radio Engineering*, vol. 79, no. 20, pp. 1801–1811, 2020.

[14] Y. I. Daradkeh and I. Tvoroshenko, "Application of an improved formal model of the hybrid development of ontologies in complex information systems," *Applied Sciences*, vol. 10, no. 19, pp. 6777, 2020.

[15] O. Gorokhovatskyi, V. Gorokhovatskyi and O. Peredrii, "Analysis of application of cluster descriptions in space of characteristic image features," *Data*, vol. 3, no. 4, pp. 52, 2018.

[16] C. Dewi, R.-C. Chen, Y.-T. Liu, X. Jiang and K. D. Hartomo, "Yolo V4 for advanced traffic sign recognition with synthetic training data generated by various GAN," *IEEE Access*, vol. 9, pp. 97228–97242, 2021.

[17] I. S. Tvoroshenko and V. O. Gorokhovatsky, "Effective tuning of membership function parameters in fuzzy systems based on multi-valued interval logic," *Telecommunications and Radio Engineering*, vol. 79, no. 2, pp. 149–163, 2020.

[18] V. O. Gorokhovatskyi, I. S. Tvoroshenko and N. V. Vlasenko, "Using fuzzy clustering in structural methods of image classification," *Telecommunications and Radio Engineering*, vol. 79, no. 9, pp. 781–791, 2020.

[19] Q. Bai, S. Li, J. Yang, Q. Song, Z. Li *et al.,* "Object detection recognition and robot grasping based on machine learning: A survey," *IEEE Access*, vol. 8, pp. 181855–181879, 2020.

[20] V. Gorokhovatskyi and I. Tvoroshenko, "Image classification based on the kohonen network and the data space modification," in *CEUR Workshop Proc.: Computer Modeling and Intelligent Systems (CMIS-2020)*, Zaporizhzhia, Ukraine, vol. 2608, pp. 1013–1026, 2020, [Online]. Available: http://ceur-ws.org/Vol-2608/.

[21] Y. I. Daradkeh, V. Gorokhovatskyi, I. Tvoroshenko, S. Gadetska and M. Al-Dhaifallah, "Methods of classification of images on the basis of the values of statistical distributions for the composition of structural description components," *IEEE Access*, vol. 9, pp. 92964–92973, 2021.

[22] X. Shen, C. Wang, X. Li, Y. Peng, Z. He *et al.,* "Learning scale awareness in keypoint extraction and description," *Pattern Recognition*, vol. 121, pp. 108221, 2022.

[23] S. Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner *et al.,* "ORB feature detector and binary descriptor," (2022, Feb. 10). [Online]. Available: https://scikit-image.org/docs/dev/auto_examples/features_detection/plot_orb.html.

[24] E. Karami, S. Prasad and M. Shehata, "Image matching using SIFT, SURF, BRIEF and ORB: Performance comparison for distorted images," in *Proc. of the 2015 Newfoundland Electrical and Computer Engineering Conf.*, St. John's, Canada, arXiv preprint arXiv:1710.02726, 2017.

[25] E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *2011 Int. Conf. on Computer Vision*, Barcelona, Spain, pp. 2564–2571, 2011.

[26] S. Leutenegger, M. Chli and R. Y. Siegwart, "BRISK: Binary robust invariant scalable keypoints," in *2011 Int. Conf. on Computer Vision*, Barcelona, Spain, pp. 2548–2555, 2011.

[27] M. A. Ahmad, V. Gorokhovatskyi, I. Tvoroshenko, N. Vlasenko and S. K. Mustafa, "The research of image classification methods based on the introducing cluster representation parameters for the structural description," *International Journal of Engineering Trends and Technology*, vol. 69, no. 10, pp. 186–192, 2021.

[28] S. Zhang, J. Wang, X. Tao, Y. Gong and N. Zheng, "Constructing deep sparse coding network for image classification," *Pattern Recognition*, vol. 64, pp. 130–140, 2017.

[29] J. Leskovec, A. Rajaraman and J. D. Ullman, "Clustering," in *Mining of Massive Datasets*, New York, NY, USA: Cambridge University Press, pp. 241–276, 2014.

[30] W. Ma, J. Qin, X. Xiang, Y. Tan, Y. Luo *et al.,* "Adaptive median filtering algorithm based on divide and conquer and its application in CAPTCHA recognition," *Computers, Materials & Continua*, vol. 58, no. 3, pp. 665–677, 2019.

[31] J. Wang and Y. Zhang, "Median filtering forensics scheme for color images based on quaternion magnitude-phase CNN," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 99–112, 2020.

[32] P. Flach, "Features," in *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*, New York, NY, USA: Cambridge University Press, pp. 298–329, 2012.

[33] X. Zhang, F. X. Yu, S. Karaman and S.-F. Chang, "Learning discriminative and transformation covariant local feature detectors," in *2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 4923–4931, 2017.

[34] R. Szeliski, "Recognition," in *Computer Vision: Algorithms and Applications*, London, Great Britain: Springer-Verlag, pp. 655–718, 2010.

[35] L. Xiangxiao, O. Honglin and X. Lijuan, "Kernel-distance-based intuitionistic fuzzy c-means clustering algorithm and its application," *Pattern Recognition and Image Analysis*, vol. 29, no. 4, pp. 592–597, 2019.

[36] M. Ghahremani, Y. Liu and B. Tiddeman, "FFD: Fast feature detector," *IEEE Transactions on Image Processing*, vol. 30, pp. 1153–1168, 2021.

[37] Q. Wu, Y. Li, Y. Lin and R. Zhou, "Weighted sparse image classification based on low rank representation," *Computers, Materials & Continua*, vol. 56, no. 1, pp. 91–105, 2018.

[38] Y. Zhang, X. Bai, R. Fan and Z. Wang, "Deviation-sparse fuzzy c-means with neighbor information constraint," *IEEE Transactions on Fuzzy Systems*, vol. 27, no. 1, pp. 185–199, 2019.

[39] W. Sun, X. Chen, X. Zhang, G. Dai, P. Chang *et al.,* "A Multi-feature learning model with enhanced local attention for vehicle re-identification," *Computers, Materials & Continua*, vol. 69, no. 3, pp. 3549–3561, 2021.

[40] Z. Luo, "Review of GAN-based person re-identification," *Journal of New Media*, vol. 3, no. 1, pp. 11–17, 2021.

[41] D. Zheng, Z. Ran, Z. Liu, L. Li and L. Tian, "An efficient bar code image recognition algorithm for sorting system," *Computers, Materials & Continua*, vol. 64, no. 3, pp. 1885–1895, 2020.

[42] R. Chen, L. Pan, C. Li, Y. Zhou, A. Chen *et al.,* "An improved deep fusion CNN for image recognition," *Computers, Materials & Continua*, vol. 65, no. 2, pp. 1691–1706, 2020.

[43] M. Takahashi, "Pokemon image library," (2022, Feb. 10). [Online]. Available: https://www.pokemon.com/ru/pokedex/.

[44] G. Bradski and A. Kaehler, "OpenCV Open Source Computer Vision," (2022, Feb. 14). [Online]. Available: https://docs.opencv.org/master/index.html.

[45] T. Grin, B. Noll and N. Chisler, "IntelliJ IDEA," (2022, Feb. 12). [Online]. Available: https://www.jetbrains.com/idea.