Recognition Algorithms Based on the Selection of 2D Representative Pseudo-objects

Olimjon N. Mirzaev[1] , Sobirjon S. Radjabov[1] , Gulmira R. Mirzaeva[2] ,
Kuvvat T. Usmanov[3] , and Nomaz M. Mirzaev[2(✉)]

[1] Research Institute for the Development of Digital Technologies and Artificial Intelligence,
17A, Buz-2, Mirzo Ulugbek, 100125 Tashkent, Republic of Uzbekistan
[2] Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, 108,
Amir Temur Avenue, 100200 Tashkent, Republic of Uzbekistan
[3] Tashkent Institute of Architecture and Civil Engineering, 13 Navoi Avenue, 100011 Tashkent,
Republic of Uzbekistan

Abstract. Building of a recognition algorithms (RA) based on the selection of representative pseudo-objects and providing a solution to the problem of recognition of objects represented in a big-dimensionality feature space (BDFS) are described in this article. The proposed approach is based on the formation of a set of 2D basic pseudo-objects and the determination of a suitable set of 2D proximity functions (PF) when designing an extreme RA. The article contains a parametric description of the proposed RA. It is presented in the form of sequence of computational procedures. And the main ones are procedures for determining: the functions of differences among objects in a 2D subspace of representative features (TSRF); groups of interconnectedness pseudo-objects (GIPO) in the same subspace; a set of basic pseudo-objects; functions of differences between the basic pseudo-object in a TSRF. There are also groups of interconnectedness and basic PF; the integral recognizing operator with respect to basic PF. The results of a comparative analysis of the proposed and known RA are presented. The main conclusion is that the implementation of the approach proposed in this paper makes it possible to switch from the original BDFS to the space of representative features (RF), the dimension of which is significantly lower.

Keywords: Recognition algorithms · Representative feature · 2D pseudo-objects · Proximity function · Basic proximity functions

1 Introduction

A detailed study of the literature, in particular [1–5], shows that pattern recognition (PR) issues are one of the main issues for specialists in computer science and applied mathematics. The reason for this is that lately the PR methods and algorithms are increasingly used in various fields of science and technology.

Currently, a number of recognition algorithms (RA) models have been developed and deeply studied [1, 4]. These include the following RA models:

- models based on discriminant functions [1, 6–10];
- models developed on the basis of mathematical statistics [8, 11–14];
- models based on the potential principle [8, 15–18];
- models developed on the basis of mathematical logic [19–24];
- models based on evaluation [1, 4, 25–28].

It is shown in [29] that these AR models were developed mainly for solving such PR problems, in the space of independent features (or the relationship between them is weak). Also, in real life there are often applied PR-tasks associated with the classification of objects described in BDFS. Under these conditions, when solving such problems, the assumption of the independence of features is not always true [29–32]. The study of available publications on PR, as well as the accumulated experience in solving a number of applied and model problems, shows that in conditions of violation of this assumption, many known RAs do not work correctly [29]. This circumstance indicates the relevance of the issues of modifying the existing and developing new AR models designed to solve the problems of object recognition presented in the BDFS.

The purpose of this work is the development of AR models based on 2D PF and providing a solution to the problems of object recognition presented in the BDFS.

To solve such problems, an approach is proposed, the key point of which is the formation of groups of interconnectedness elements in each pair of RF and the selection of representative pseudo-objects.

The concepts and designations used in this article are taken from works [1, 4, 5, 27, 28, 29].

2 Basic Concepts and Notation

Consider the case when, given a set $\mathbb{M}$ of admissible objects, in the n-dimensional feature space $\mathbb{X}$. Each admissible object $\mathfrak{O}$ ($\mathfrak{O} \in \mathbb{M}$) in the space $\mathbb{X}$ corresponds to an n-dimensional numerical characteristic (n- dimensional description vector) of the object $\mathcal{I}(\mathfrak{O})$, where $\mathcal{I}(\mathfrak{O}) = (\mathfrak{o}_1, .., \mathfrak{o}_u, .., \mathfrak{o}_n)$[1, 4]. Supposed that the set $\mathbb{M}$ consists of $\mathcal{k}$ disjoint subsets (classes) $\mathcal{C}_1, .., \mathcal{C}_j, .., \mathcal{C}_\mathcal{k}$:

$$\mathbb{M} = \cup_{j=1}^{\mathcal{k}} \mathcal{C}_j, \qquad \mathcal{C}_i \cap \mathcal{C}_j \neq \emptyset, \ i \neq j, \ i,j \in \{1, \dots, \mathcal{k}\}. \tag{1}$$

The partition (1) is not completely defined, but only some initial information $\mathbb{I}_0$ about the subsets $\mathcal{C}_1, .., \mathcal{C}_j, .., \mathcal{C}_\mathcal{k}$ is given. For clarify the concept of initial information, select m objects from the set of possible objects $\mathbb{M}$ and denote by $\widetilde{\mathfrak{O}}^m$:

$$\widetilde{\mathfrak{O}}^m = \{\mathfrak{O}_1, \dots, \mathfrak{O}_u, \dots, \mathfrak{O}_m\}.$$

Let us introduce the following notations:

$$\widetilde{\mathcal{C}}_j = \widetilde{\mathfrak{O}}^m \cap \mathcal{C}_j, \ \ \widetilde{\mathcal{D}}_j = \widetilde{\mathfrak{O}}^m / \widetilde{\mathcal{C}}_j.$$

In this case, the initial information can be described in the form of some set $\mathbb{I}_0$, elements of which are the pairs $< \mathfrak{O}_i, \mathbb{i}(\mathfrak{O}_i) > \ (\forall \mathfrak{O}_i \in \mathbb{M}, i = \overline{1, m})$::

$$\mathbb{I}_0 = \{< \mathfrak{O}_1, \mathbb{i}(\mathfrak{O}_1) >, \dots, < \mathfrak{O}_i, \mathbb{i}(\mathfrak{O}_i) >, \dots, < \mathfrak{O}_m, \mathbb{i}(\mathfrak{O}_m) >\},$$

$$\mathbb{i}(\mathfrak{O}_i) = (\xi_{i1}, \dots, \xi_{ij}, \dots, \xi_{i\mathcal{k}}), \tag{2}$$

where $\mathfrak{D}_i$ is an admissible object, and $\mathbb{i}(\mathfrak{D}_i)$ is the information vector of the object $\mathfrak{D}_i$.

Every element of the information vector $\mathbb{i}(\mathfrak{D}_i)$ is given in the form:

$$\xi_{ij} = \begin{cases} 1, \text{if } \mathfrak{D}_i \in \widetilde{\mathcal{C}}_j; \\ 0, \text{if } \mathfrak{D}_i \notin \widetilde{\mathcal{C}}_j. \end{cases}$$

3   Statement of the Problem

Let us consider the problem of PR in the standard setting [1, 26]. Let the set of objects $\widetilde{\mathfrak{S}}^q \left(\widetilde{\mathfrak{S}}^q = \left\{ \mathfrak{D}'_1, \ldots, \mathfrak{D}'_i, \ldots, \mathfrak{D}'_q \right\}, \; \widetilde{\mathfrak{S}}^q \subset \mathbb{M} \right)$, described in the space of features $\mathbb{X}$. The problem is to develope such a modek of RA $\mathfrak{A}$, which, using the initial information $\mathbb{I}_0$ calculates the values of the predicate $P_j \left(\mathfrak{D}'_i \right) \left(P_j(\mathfrak{D}'_i) = "\mathfrak{D}'_i \in \mathcal{C}_j" \right)$ for all objects $\widetilde{\mathfrak{S}}^q$:

$$\mathfrak{A} \left(I_0, \widetilde{\mathfrak{S}}^q \right) = \left\| \beta_{ij} \right\|_{m \times \mathscr{k}}, \; \beta_{ij} = P_j(\mathfrak{D}'_i).$$

$$\beta_{ij} = \begin{cases} 1, \text{if the object } \mathfrak{D}'_i \text{ belongs to the class } \mathcal{C}_j; \\ 0, \text{if the object } \mathfrak{D}'_i \text{ does not belong to the class } \mathcal{C}_j; \\ \Delta, \text{if the model has not computed the values of the predicate } P_j(\mathfrak{D}'_i). \end{cases}$$

4   Method of Solution

The article proposes an approach based on the choice of a set of representative pseudo-objects in each RF pair when constructing the RA model. The essence of the proposed RA model is based on the formation of a proximity function in 2D subspaces of the RF. Let us consider an arbitrary admissible object $\mathfrak{D}$ ($\mathfrak{D} \in \mathbb{M}$) given in an n-dimensional feature space $\mathbb{X}$. Let $\widetilde{\omega}\mathfrak{D}$ mean that the object $\mathfrak{D}$ is described in the n-dimensional feature space. Moreover, to each n-dimensional space there corresponds an n-dimensional Boolean vector $\widetilde{\omega}(\widetilde{\omega} = (\omega_1, \ldots, \omega_i, \ldots, \omega_n), \; \eth = \widetilde{\omega}_B)$, the components of which take on values 1 or 0 depending on whether or not the corresponding features are included in the description of some $\widetilde{\omega}$-part of the object $\mathfrak{D}$. Thus, $\widetilde{\omega}\mathfrak{D}$ is not an object, but consists of its $\widetilde{\omega}$-part. Next, let us call $\widetilde{\omega}$-parts of the object as a pseudo-object.

Defining the proposed RA model includes the following main stages.

1. Formation of groups of tightly coupled features. A set of n' ($n' < n$) "independent" groups of tightly coupled features (TCF) are formed. Issues related to the formation of TCFs are considered in more detail in [31, 32].

2. Selection of a set of RF. One representative from each TCF is determined, which is a typical element of its group of features. As a result, n' RF is determined, which is isolated by an n-dimensional Boolean vector $\mathbb{r}$ ($\mathbb{r} = (r_1, \ldots, r_i, \ldots, r_n), n' = \|\mathbb{r}\|_B)$,, where $r_i = 1$, if the feature x_i is an RF, or $r_i = 0$ otherwise. The issues of RF isolation are considered in works [31–33].

The generated RF space is denoted by $\mathbb{X}' \left(\dim \left(\mathbb{X}' \right) = n' \right)$.

3. Determination of the diversity function in the TSRF. The diversity function (DF), which characterizes the remoteness of the objects $\mathfrak{O}_u$ and $\mathfrak{O}$ in the TSRF, is given. Let in the space $\mathbb{X}'$ a set $\mathfrak{C}$, consisting of $\mathfrak{n}$ elements, that is $\mathfrak{C} = (\mathcal{D}_1, \dots, \mathcal{D}_\tau, \dots, \mathcal{D}_\mathfrak{n})$, be given. And each $\mathfrak{C}$ element forms a 2D RF subspace: $\mathcal{D}_\tau = (x_{\tau_1}, x_{\tau_2}), x_{\tau_1}, x_{\tau_2} \in \mathbb{X}')$. Then the difference between the objects $\mathfrak{O}_u$ and $\mathfrak{O}$ in the subspace $\mathcal{D}_\tau$ is determined as follows [8]:

$$d_\tau(\mathfrak{O}_u, \mathfrak{O}) = \sum_{i=1}^{2} \gamma_{\tau_i}(a_{u,\tau_i} - a_{\tau_i})^2, \tag{3}$$

where $\gamma_{\tau_i}, \gamma_{\tau_2}$ are the parameters of the algorithm, which we denote by $\widetilde{\gamma}_\tau = (\gamma_{\tau_1}, \gamma_{\tau_2})$.

4. Selection of GIPO in the TSRF. m' "independent" GIPO $\mathbb{V}_A$ are determined on the basis of pairwise comparison of these objects in order to assess their proximity in the 2D RF subspace. After the implementation of this stage, disjoint m' groups are separated out:

$$\mathbb{V}_A = \{\mathfrak{V}_1, \dots, \mathfrak{V}_q, \dots, \mathfrak{V}_{\mathrm{m}'}\}; \tag{4}$$

In doing so, GIPOs (4) are formed using the function (3).

5. Determination of basic pseudo-objects in the TSRF. A set of basic pseudo-objects E_q^τ, which are typical representatives of each GIPO in the TSRF, is determined. The components of the basic pseudo-object E_q^τ in each group of interconnectedness objects can be calculated as the average values for all elements of the group $\mathfrak{V}_q\left(q = \overline{1, m'}\right)$:

$$b_{qi}^\tau = \frac{1}{|\mathfrak{V}_q|} \sum_{a_{ui} \in \mathfrak{V}_q} a_{ui}, \ i = \overline{1, 2}. \tag{5}$$

As a result, we obtain m' basic pseudo-objects for each set of 2D RF subspace:

$$\mathbb{E}^\tau = \left\{E_1^\tau, \dots, E_q^\tau, \dots, E_{m'}^\tau\right\}.$$

Each of them is specified as a 2D vector, i.e., $E_q^\tau = \left(b_{qi_1}^\tau, b_{qi_2}^\tau\right)$.

6. Determination of the diversity function $d\left(E_q^\tau, \mathfrak{O}\right)$ between the basic pseudo-object E_q^τ and the object $\mathfrak{O}$ in the TSRF. At this stage, the DF is determined between the base pseudo-object E_q^τ and the object $\mathfrak{O}$ in the TSRF. The difference between them is determined by the formula (3):

$$d_\tau\left(E_q^\tau, \mathfrak{O}\right) = \sum_{i=1}^{2} \varrho_{\tau_i} \rho\left(b_{q,\tau_i}, a_{\tau_i}\right), \tag{6}$$

where $\rho\left(b_{q,\tau_i}, a_{\tau_i}\right)$ is the estimate of the difference between the base pseudo-object E_q^τ and the object $\mathfrak{O}$, calculated from x_{τ_i}; ϱ_{τ_i} is a parameter of the algorithm $\left(\widetilde{\varrho}_\tau = (\varrho_{\tau_i}, \varrho_{\tau_i})\right)$.

7. Specifying the proximity function $\mathfrak{H}(E_q^\tau, \mathfrak{O})$ between the E_q^τ and the $\mathfrak{O}$ in the TSRF. At this stage, based on the radial functions, the PF between the pseudo-object E_q^τ and the object $\mathfrak{O}$ in the TSRF is determined, for example, in the following form [15]:

$$\mathfrak{H}_\tau\left(E_q^\tau, \mathfrak{O}\right) = 1/\left(1 + \xi_\tau d_\tau\left(E_q^\tau, \mathfrak{O}\right)\right). \tag{7}$$

We obtain $\mathfrak{n}$ PFs. Each PF given in the form (7) is determined by the parameter ξ_τ.

8. Selection of groups of interconnectedness PF. At this stage, the systems $\mathfrak{W}_A$ of 'independent' groups of PF are specified. Let each proximity function $\mathfrak{H}_u$ correspond to a numerical matrix:

$$\mathfrak{H}_\tau\left(\mathbb{E}_2, \tilde{\mathfrak{O}}^m\right) = \|\mathfrak{h}_{ij}^{(\tau)}\|_{m \times m'}.$$

Let us consider the PF set $\{\mathfrak{H}_1, \ldots, \mathfrak{H}_{\mathfrak{n}}\}$. Let us introduce the function $\eta(\mathfrak{H}_u, \mathfrak{H}_v)$, which characterizes the strength of the pairwise connection between the numerical matrices $\|\mathfrak{h}_{ij}^{(u)}\|_{m \times m'}$ and $\|\mathfrak{h}_{ij}^{(v)}\|_{m \times m'}$. Let $\mathcal{W}_q$ ($q = \overline{1, \mathfrak{n}'}$) be a group of interconnectedness PF. The measure of proximity $\mathfrak{K}(\mathcal{W}_p, \mathcal{W}_q)$between the groups $\mathcal{W}_p$ and $\mathcal{W}_q$ can be specified in different ways, for example:

$$\mathfrak{K}(\mathcal{W}_p, \mathcal{W}_q) = \frac{1}{N_p \cdot N_q} \sum_{\mathfrak{H}_i \in \mathcal{W}_p} \sum_{\mathfrak{H}_i \in \mathcal{W}_q} \eta\left(\mathfrak{H}_i, \mathfrak{H}_j\right), \tag{8}$$

where N_p, N_q is the number of elements included, respectively, in the sets $\mathcal{W}_p, \mathcal{W}_q$; $\eta(\mathfrak{H}_i, \mathfrak{H}_j)$ is a function that characterizes the assessment of the pairwise relationship between $\mathfrak{H}_i$ and $\mathfrak{H}_j$.

We obtain $\mathfrak{n}'$ groups of DF in each TSRF.

9. Determination of basic PF in each group of interconnectedness PF. At this stage, basic PFs are selected and the set $\mathfrak{B}$, consisting of $\mathfrak{n}'$ basic PFs, is formed. The choice of basic PF is based on the removal $N_p - 1$ $\left(p = \overline{1, \mathfrak{n}'}\right)$ PFs, giving almost the same results when evaluating the membership, from the selected group of basic PFs $\mathcal{W}_p$. Moreover, each allocated PF should be a typical representative of the selected group of tightly coupled PFs. We obtain $\mathfrak{n}'$ PFs, which is much less than the initial one, i.e., $\mathfrak{n}' < \mathfrak{n}$.

10. Synthesis of an integral recognition operator based on basic PF. At this stage, the integral recognition operator B is determined by the selected basic PFs:

$$B\left(\mathcal{C}_j, \mathfrak{O}\right) = \sum_{u=1}^{\mathfrak{n}'} \nu_u \mathfrak{H}_u\left(\mathcal{C}_j, \mathfrak{O}\right), \tag{9}$$

where $\nu_u$is the parameter of the integral recognition operator B; $\mathfrak{n}'$ is the number of basic PF.

11. Decision rule. The decision is made element by element [1, 4], i.e.

$$\beta_{ij} = C\left(B\left(\mathcal{C}_j, S\right)\right) = \begin{cases} 0, & if\ B\left(\mathcal{C}_j, \mathfrak{O}\right) < c_1, \\ 1, & if\ B\left(\mathcal{C}_j, \mathfrak{O}\right) > c_2, \\ \Delta, & if\ c_1 \leq B\left(\mathcal{C}_j, \mathfrak{O}\right) \leq c_2 \end{cases}$$

where c_1, c_2 – are algorithm parameters.

Thus, we have defined the AR model $\mathfrak{A}(\widetilde{\pi}, \mathfrak{O})$, based on a two-dimensional PF of the intensional type. Any AR from the model $\mathfrak{A}(\widetilde{\pi}, \mathfrak{O})$ is one-to-one in the parameter space $\widetilde{\pi}$:

$$\widetilde{\pi} = \left(n', \mathbb{r}, \{\widetilde{\gamma}_\tau\}, m', \{\mathfrak{E}_q\}, \{\widetilde{\varrho}_\tau\}, \{\xi_\tau\}, \mathfrak{n}', \{\mathfrak{H}_u\}, \{v_u\}, c_1, c_2\right).$$

The search for the best RA within the framework of the considered model is carried out in the parameter space $\widetilde{\pi}$ [34, 35].

5   Experiments and Results

In order to conduct experimental studies on the assessment of the considered RA model, functional diagrams of the created software complex and the corresponding procedures were developed. The software for these procedures is developed in the C++ programming language using the OpenCV library.

An experimental study of the performance of the proposed RA model was carried out when solving a number of problems. The following were chosen as the tested RA models: 1) the classical RA model based on potential functions ($\mathfrak{A}_1$-models) [15]; 2) the proposed model ($\mathfrak{A}_2$-models). The selection of the $\mathfrak{A}_1$-model for comparison is explained by the fact that $\mathfrak{A}_1$ and $\mathfrak{A}_2$ belong to the same category of RA models.

When solving the problems under consideration (see Sects. 5.1 and 5.2), the comparative analysis of the above-mentioned RA models based on three indicators was carried out: recognition accuracy of objects in the control sample (in%); the time spent by the algorithm for training (in seconds); the time spent by the algorithm to recognize objects in the control sample (in seconds).

Let the initial sample T, consisting of m objects $\{\mathfrak{O}_1, \ldots, \mathfrak{O}_u, \ldots, \mathfrak{O}_m\}$, be given. To calculate the quality assessment by the accuracy of the tested models of recognition operators, the set T ($T = \{\mathfrak{O}_1, \ldots, \mathfrak{O}_u, \ldots, \mathfrak{O}_m\}$) is divided into two parts – V_t and V_c ($T = V_t \cup V_c$, V_t – the training sample, V_c – the control sample). To exclude a successful (or unsuccessful) partition of the set T into two parts, let us use the cross-validation method [36], the essence of which is as follows. The initial sample of objects T is divided by random selection into 10 subsets. As a result, we obtain the sets $\mathbb{T}$:

$$\mathbb{T} = \{\mathcal{T}_1, \ldots, \mathcal{T}_u, \ldots, \mathcal{T}_{10}\}.$$

In this case, the elements of $\mathbb{T}$ are required to meet the following simple conditions (at $u, v \in \{1, \ldots 10\}$):

1. $\mathcal{T}_u \cap \mathcal{T}_v = \emptyset, u \neq v$;
2. $|\mathcal{T}_1| = \cdots = |\mathcal{T}_u| = \cdots = |\mathcal{T}_{10}|$;
3. $\left|\mathcal{T}_1 \cap \mathcal{C}_j\right| \approx \ldots \approx \left|\mathcal{T}_u \cap \mathcal{C}_j\right| \approx \ldots \approx \left|\mathcal{T}_{10} \cap \mathcal{C}_j\right|$;
4. $|\mathcal{T}_u| = \sum_{j=1}^{l}\left|\mathcal{T}_u \cap \mathcal{C}_j\right|$.

The cross-validation process for these subsets consists of the following cyclically performed steps ($\mathfrak{v} = 0$; $\mathfrak{u} = 1$):

– the condition of completion of the process of forming the set $\mathbb{T}$. If $(\mathfrak{v} < \mathfrak{h})$, then the following actions are performed: a) initial data is generated for the given distributions. The initial sample consists of m implementations (for objects of each class, m_j implementations). In this case, the number of features is equal to n. The number of groups of strongly connected features - n'; b) the original sample is split into $\mathfrak{h}$ random subsets of objects. As a result, we get $\mathfrak{h}$ subsets of objects;
– choose $(\mathfrak{h} - 0.1\mathfrak{h})$ from $\mathfrak{h}$ blocks as V_t and on this sample AR are trained with the given parameters;
– the trained RA is compared with (V_c). As a result of the implementation of this stage, the RA accuracy is estimated at V_c;
– the condition $\mathfrak{u} \leq \mathfrak{h}$ checked; if it is true, then $\mathfrak{u} := \mathfrak{u} + 1$ and go to step 5; otherwise, go to step 1;
– from V_t one subset is selected as V_c. In this case, the subset used as V_c is fixed accordingly, and it does not participate in the procedure for selecting candidates when forming the next sample of control objects. one subset is selected from the training samples as a control sample. In this case, the subset used as a control sample is fixed accordingly, and it does not participate in the procedure for selecting candidates when forming the next sample of control objects.

5.1 Model Problem

The main characteristics of the initial data generated for the model problem in this experiment are as follows: the size of the original sample $m = 1000$, the number of classes $\mathit{k} = 2$, the number of features $n = 500$, the strongly coupled features groups $n' = 6$, the amount of training and control sample, respectively $|V_t| = 900$, $|V_c| = 100$.

The model problem was solved using the RA models $\mathfrak{A}_1$ and $\mathfrak{A}_2$. Accuracy of recognition in the training for $\mathfrak{A}_1$ is 96,8%, for $\mathfrak{A}_2$ is 97,5%. The results of solving the problem under consideration with the use of these RAs in the verification process are 82.7% and 94.6%, respectively (see Table 1).

Table 1. The results of the considered models

RA models	Time, s		Recognition accuracy, %
	Training	Recognition	
$\mathfrak{A}_1$-model	1.3869	0.00083	82.7
$\mathfrak{A}_2$-model	4.1218	0.00027	94.6

Analysis of the results shows that the proposed $\mathfrak{A}_2$ improves the accuracy of recognition described in the space of interrelated features (more than 10% higher than $\mathfrak{A}_1$). This is due to the fact that in the $\mathfrak{A}_2$ model, in contrast to the $\mathfrak{A}_1$, a number of procedures are used to improve the recognition accuracy, for example, the determination of representative representatives of each class; selecting representative and preferred combinations of features. Thus, the CV-dimension of the developed algorithm is much less than the

same indicator of the original recognition algorithm. Consequently, the accuracy of the proposed RA should be higher (on the control sample with the same size of the training sample), which was shown by the results of this experimental study.

5.2 Practical Tasks

With the development of information technology capabilities, systems with the ability to recognize a person using biometric characteristics are widely spread [37–40]. This is due to the fact that the introduction of biometric recognition (BR) methods based on unique biological characteristics that uniquely identify each person is relatively inexpensive and convenient. In addition, the use of BRs for solving various applied problems is constantly expanding [37].

Among BRs, a special place is occupied by recognition systems based on images of auricles. The advantages of such BRs are unobtrusiveness, inactivity and relatively low cost [37].

In this task, the initial sample T included 500 images of the auricles (250 images of the left and right auricles). The number of classes is five. Each class included 50 pairs of auricle images of one person. The images of the auricle were described by 147 features. Table 2 shows the results of solving this problem.

Table 2. The results of the proposed models for the recognition problem from the images of the auricle

RA models	Time, s		Recognition accuracy, %
	Training	Recognition	
$\mathfrak{A}_1$-model	0.23155	0.00230	83.07
$\mathfrak{A}_2$-model	0.82436	0.00019	97.84

The calculation results show that when using the $\mathfrak{A}_2$ model it allows to improve the recognition accuracy than RA $\mathfrak{A}_1$.

Also, the results of the experiment showed the high accuracy of the developed model when solving recognition problems from the images of the auricle.

The time spent on training the proposed model is more than the time spent on training RA $\mathfrak{A}_1$. It should be noted that the time spent by the RA $\mathfrak{A}_2$ model on recognizing objects from the control sample is less than the same indicator for the RA $\mathfrak{A}_1$ model.

Considering the results of the experimental study, we can say that the proposed RA model more accurately solves the PR problem under conditions when the size of the training sample and the dimension of the feature space are large enough.

6 Conclusions

A new approach based on formation of a set of 2D basic pseudo-objects within the training set is proposed. The implementation of the proposed approach makes it possible

to move from the BDFS to the RF space, which has a lower dimension. Based on this approach, an RA model was developed taking into account the structure of the original data. The essence of the proposed model is to identify independent groups of interrelated features and the corresponding set of RF. A distinctive feature of this RA model is the determination of the preferred PF when constructing a base of 2D functions.

The proposed new approach allows: expanding models of recognition operators based on potential functions; improving the recognition accuracy of objects described in BDFS; increasing the area of application of the RA model based on potential functions when solving applied problems.

The results of experiment showed that the proposed RA model (see Sect. 4) improves accuracy and significantly reduces the number of computational operations in the process of recognizing an unknown object specified in the BDFS. At the same time, the time spent on training the model increased. This circumstance is explained by the fact that rather complex optimization procedures are used to train it than to train the traditional RA model.

In the process of solving considered in Sect. 5, it was determined that the stages of the formation of groups of "independent" features, namely, the issues of determining the number of these groups, isolating basic PFs and constructing an integral recognizing operator based on basic PFs, are most important in determining the extreme RA within the proposed model. Therefore, it is necessary to continue research towards the development of algorithms that refine these parameters of the RA model.

References

1. Zhuravlev, Y.: An algebraic approach to recognition or classifications problems. Pattern Recogn. Image Anal. **8**(1), 59–100 (1998)
2. Homenda, W., Pedrycz, W.: Pattern Recognition: A Quality of Data Perspective. Wiley, New York (2018)
3. Beyere, M., Richter, M., Nagel, M.: Pattern Recognition: Introduction, Features, Classifiers and Principles. De Gruyter Oldenbourg, Boston (2018)
4. Zhuravlev, Y.I.: Selected Scientific Works. (Izbrannye Nauchnye Trudy). Magister, Moscow (1998). (in Russian)
5. Kamilov, M., Fazilov, S., Mirzaeva, G., Gulyamova, D., Mirzaev, N.: Building a model of recognizing operators based on the definition of basic reference objects. J. Phys. Conf. Ser. **1441**(1), 012142 (2020). https://doi.org/10.1088/1742-6596/1441/1/012142
6. McLachlan, G.J.: Discriminant Analysis and Statistical Pattern Recognition. Wiley, New York (2004)
7. Zhuravlev, Y.I., Dyusembaev, A.E.: A neural network construction for recognition problems with standard information on the basis of a model of algorithms with piecewise linear surfaces and parameters. Rep. Acad. Sci. **488**(1), 11–15 (2019). (in Russian). (Zhuravlev Yu.V. I., Dyusembaev A.E. Postroenie nejroseti dlya zadach raspoznavaniya so standartnoj informaciej na osnove modeli algoritmov s kusochno-linejnymi poverhnostyami i parametrami. Doklady Akademii Nauk). https://doi.org/10.31857/S0869-5652488111-15
8. Tou, J., Gonzalez, R.: Pattern Recognition Principles. Addison-Wesley, Boston (1974)
9. Li, Y., Liu, B., Yu, Y., Li, H., Sun, J., Cui, J.: 3E-LDA: three enhancements to linear discriminant analysis. ACM Trans. Knowl. Disc. Data **15**(4), 1–20 (2021). https://doi.org/10.1145/3442347

10. Li, C.-N., Shao, Y.-H., Yin, W., Liu, M.-Z.: Robust and sparse linear discriminant analysis via an alternating direction method of multipliers. IEEE Trans. Neural Netw. Learn. Syst. **31**(3), 915–926 (2020). https://doi.org/10.1109/TNNLS.2019.2910991

11. Duda, R., Hart, P., Stork, D.: Pattern Classification. Wiley, New York (2001)

12. Webb, A.R., Copsey, K.D.: Statistical Pattern Recognition. Wiley, New York (2011)

13. Jain, A.K., Duin, P.W., Mao, J.: Statistical pattern recognition: a review. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1), 4–37 (2000). https://doi.org/10.1109/34.824819

14. Merkov, A.B.: Pattern Recognition: An Introduction to Statistical Learning Methods (Raspoznavanie Obrazov: Vvedenie v Metody Statisticheskogo Obuchenija). URSS, Moscow (2019). (in Russian)

15. Ayzerman, M.A., Braverman, E.M., Rozonoer, L.I.: Method of Potential Functions in the Theory of Machine Learning (Metod potencialnyh funkcij v teorii mashinnogo obucheniya). Nauka, Moscow (1970).(in Russian)

16. Dubrovin, V.I., Koretsky, N.K., Subbotin, S.A.: Modified method of potential functions. Complex systems and processes (Modificirovannyj metod potencialnyh funkcij. Slozhnye sistemy i processy), vol. 1, pp. 12–19 (2002). (in Russian)

17. Oliveri, P.: Potential function methods: efficient probabilistic approaches to model complex data distributions. SAGE J. **28**(4), 14–15 (2017). https://doi.org/10.1177/0960336017703253

18. Sulewski, P.: Potential function method approach to pattern recognition applications. In: Environment, Technology, Resources. Proceedings of the 11th International Scientific and Practical Conference, Rezekne, Latvia, pp. 30–35 (2017). https://doi.org/10.17770/10.17770/etr2017vol2.2512

19. Kudryavtsev, V.B., Andreev, A.E., Hasanov, E.E.: Test Recognition Theory (Teoriya raspoznavaniya testov). Fizmatlit, Moscow (2007).(in Russian)

20. Lbov, G.S., Startseva, N.G.: Logical Decision Functions and Questions of Statistical Stability of Decisions (Logicheskie reshayushie funkcii i voprosy statisticheskoj ustojchivosti reshenij). IM SB RAS, Novosibirsk (1999).(in Russian)

21. Djukova, E.V., Masliakov, G.O., Prokofyev, P.A.: Logical classification of partially ordered data. In: Kuznetsov, S.O., Panov, A.I. (eds.) RCAI 2019. CCIS, vol. 1093, pp. 115–126. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30763-9_10

22. Fazilov, Sh., Khamdamov, R., Mirzaeva, G., Gulyamova, D., Mirzaev, N.: Models of recognition algorithms based on linear threshold functions. J. Phys. Conf. Ser. **1441**(1), 012138 (2020). https://doi.org/10.1088/1742-6596/1441/1/012138

23. Povhan, I.F.: Logical recognition tree construction on the basis of a step-to-step elementary attribute selection. Radio Electron. Comput. Sci. Control **2**, 95–105 (2020). https://doi.org/10.15588/1607-3274-2020-2-10

24. Povhan, I.: Logical classification trees in recognition problems. Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska **10**(2), 12–15 (2020). https://doi.org/10.35784/iapgos.927

25. Ignat'ev, O.A.: Construction of a correct combination of estimation algorithms adjusted using the cross validation technique. Comput. Math. Math. Phys. **55**(12), 2094–2099 (2015)

26. Nishanov, A.K., Djurayev, G.P., Khasanova, M.A.: Improved algorithms for calculating evaluations in processing medical data. COMPUSOFT Int. J. Adv. Comput. Technol. **8**(6), 3158–3165 (2019)

27. Zhuravlev, Y.I., Ryazanov V.V., Senko O.V.: "Recognition". Mathematical methods. Software system. Practical applications (Raspoznavaniye. Matematicheskiye metody. Programmnaya sistema. Prakticheskoye primeneniye), Fazis, Moscow (2006). (in Russian)

28. Kamilov, M.M., Fazilov, S.K., Mirzaev, N.M., Radjabov S.S.: Algorithm of calculation of estimates in condition of features' correlations. In: 3rd International Conference on Problems of Cybernetics and Informatics, PCI 2010, Baku, Azerbaijan, pp. 278–281. ANAS (2010)

29. Kamilov, M.M., Fazilov, S.K., Mirzaev, N.M., Radjabov S.S.: Models of recognition algorithms based on the assessment of the interconnectedness of features (Modeli algoritmov raspoznavaniya na osnove otsenki vzaimosvyazannosti priznakov). Science and Technology, Tashkent (2020). (in Russian)
30. Lantz, B.: Machine Learning with R: Expert Techniques for Predictive Modeling. Packt Publishing, Birmingham (2019)
31. Fazilov, S.K., Mirzaev, N.M., Mirzaeva, G.R., Tashmetov, S.E.: Construction of recognition algorithms based on the two-dimensional functions. In: Santosh, K.C., Hegadi, R.S. (eds.) RTIP2R 2018. CCIS, vol. 1035, pp. 474–483. Springer, Singapore (2019). https://doi.org/10.1007/978-981-13-9181-1_42
32. Fazilov, S., Mirzaev, N., Mirzaeva, G.: Modified recognition algorithms based on the construction of models of elementary transformations. Procedia Comput. Sci. **150**, 671–678 (2019). https://doi.org/10.1016/j.procs.2019.02.037
33. Fazilov, S.K., Lutfullaev, R.A, Mirzaev, N.M., Mukhamadiev, A.S.: Statistical approach to building a model of recognition operators under conditions of high dimensionality of a feature space. J. Phys. Conf. Ser. **1333**, 032017 (2019). https://doi.org/10.1088/1742-6596/1333/3/032017
34. Fazilov, S., Mirzaev, N., Radjabov, S., Mirzaev, O.: Determining of parameters in the construction of recognition operators in conditions of features correlations. CEUR Workshop Proc. **2098**, 10 (2018)
35. Fazilov, S., Mirzaev, N., Radjabov, S., Mirzaeva, G.: Determination of representative features when building an extreme recognition algorithm. J. Phys. Conf. Ser. **1260**, 102003 (2019). https://doi.org/10.1088/1742-6596/1260/10/102003
36. Braga-Neto, U.M., Dougherty, E.R.: Error Estimation for Pattern Recognition. Springer, New York (2016)
37. Fazilov S., Mirzaev O., Saliev E., Khaydarova M., Ibragimova S., Mirzaev N.: Model of recognition algorithms for objects specified as images. In: Proceedings of the 9th International Conference Advanced Computer Information Technologies, ACIT 2019, Ceske Budejovice, Czech Republic, 5–7 June 2019 (2019). https://doi.org/10.1109/ACITT.2019.8779943
38. Bolle, R.M., Connell, J.H., Pankanti, S., Ratha, N.K., Senior, A.W.: Guide to Biometrics. Springer, New York (2004). https://doi.org/10.1007/978-1-4757-4036-3
39. Benzaoui, A., Kheider, A., Boukrouche, A.: Ear description and recognition using ELBP and wavelets. In: International Conference on Applied Research in Computer Science and Engineering, Beirut, Lebanon, pp. 1–6. IEEE (2015). https://doi.org/10.1109/ARCSE.2015.7338146
40. Pflug, A., Busch, C.: Ear biometrics: a survey of detection, feature extraction and recognition methods. IET Biometrics J. **1**(2), 114–129 (2012). https://doi.org/10.1049/iet-bmt.2011.0003