Determining of Parameters in the Construction of Recognition Operators in Conditions of Features Correlations

Shavkat Kh. Fazilov, Nomaz M. Mirzaev,
Sobirjon S. Radjabov, and Olimjon N. Mirzaev

Scientific and Innovation Center of Information and Communication Technologies,
Tashkent, Republic of Uzbekistan
sh.fazilov@mail.ru,nomazmirza@rambler.ru,
s_radjabov@yahoo.com,mirzaevon@mail.ru
http://www.insoft.uz

Abstract. The problems of constructing an extreme recogniton operator in the framework of the model of recognition operators based on radial functions are considered in the context of features correlations. To solve the problem of finding the optimal value of parameters, a heuristic approach based on the successive application of local procedures for calculating parameter values at each stage is proposed.
In order to assess the efficiency of the proposed method, an experimental study was carried out to solve a model problem generated from specified distribution parameters and the problem of recognizing a person by photo portraits.
The analysis of the results of the conducted experimental research showed that the proposed method of determining parameters in the construction of recognition operators in conditions of features correlations allows to improve recognition accuracy in solving applied problems. In this case, the constructed extreme model of recognition operators signicantly reduces the number of computational operations when recognizing an unknown object.

Keywords: Extremal recognition operators · Features correlation · Parameters of recognition models

1 Introduction

Pattern recognition is one of the most intensively developing areas in the field of computer science. This is due to the fact that the methods and algorithms for pattern recognition have become increasingly used in science, technology,

In: S. Belim et al. (eds.): OPTA-SCL 2018, Omsk, Russia, published at http://ceur-ws.org

production and everyday life in recent years. Therefore, an increasing number of specialists pay attention to the problem of pattern recognition and the number of scientific publications on this subject is constantly growing.

An analysis of existing literature sources on pattern recognition shows that the development of this direction is divided into two stages. The first stage of development consisted of the projects of various technical devices or algorithms for solving specific applications. The value of the developed methods is determined, first of all, by the achieved experimental results. At the second stage, the transition from individual algorithms to the construction of models - families of algorithms for a single description of methods for solving classification problems was carried out. At this stage, the recognition problem was formulated as an optimization problem, which allowed using the known optimization methods and stimulated the appearance of new methods, in particular [1–3].

Several well-known models of recognition algorithms have been constructed and studied up to now [1, 4–8]: models based on the principle of separation; statistical models; models based on the potential principle; models based on mathematical logic; models based on the calculation of estimates. However, analysis of these models shows that currently the models of recognition algorithms, oriented to solving problems, where objects are described in the space of independent (or weakly dependent) features, are mainly being developed. In practice, often there are applied problems of recognizing of objects defined in the large dimensional feature space. When solving such problems, the assumption of the independence of features is often not fulfilled [9]. Consequently, there remains an insufficiently resolved issue related to the creation of recognizing algorithms that can be applied to solving applied problems of diagnosing, forecasting and classifying objects in conditions of large dimensions of the feature space and correlated features.

In [10], a model of recognition operators based on the evaluation of the features correlations in conditions of features correlations was described. The main idea of this model is to find correlations between the features characterizing objects belonging to the same class. We note that in [10] the problem of constructing an extreme recognition operator is not considered.

For completeness of the presentation, we briefly consider the above-mentioned model, determined by specifying seven steps:

1) extracting the subset of strongly correlated features:

$$\mathfrak{G}_B = \left\{\mathfrak{T}_1, \mathfrak{T}_2, \dots, \mathfrak{T}_{n'}\right\}; \tag{1}$$

2) forming of the set of representative features:

$$\mathfrak{R}_X = \left\{x_{i_1}, \dots, x_{i_q}, \dots, x_{i_{n'}}\right\}; \tag{2}$$

3) determining the models of correlations F_q in each subset $\mathfrak{T}_q$ $(q = \overline{1, n'})$ for K_j $(j = \overline{1, l})$:

$$\mathfrak{M} = \left\{F_{i_1}, \dots, F_{i_q}, \dots, F_{i_{n'}}\right\}; \tag{3}$$

4) extracting the set of preferred correlation models:

$$\mathfrak{R}_{\mathfrak{M}} = \Big\{F_{j_1}, \dots, F_{j_q}, \dots, F_{j_{n''}}\Big\}; \tag{4}$$

5) determining the difference function between the object S and the objects of the class K_j $(j = \overline{1, l})$;

6) determining the proximity function between the object S and the objects of the class K_j $(j = \overline{1, l})$.

We have defined a model of recognition operators of the type of potential functions based on the evaluation of the features correlations. An arbitrary operator B from this model is completely determined by specifying the set of parameters $\tilde{\pi}$ [11]:

$$\tilde{\pi} = \Big(n', \{\tilde{w}\}, \{\overline{c}\}, \{\tilde{\omega}\}, \{\lambda_i\}, \{\xi_i\}, \{\gamma_u\}\Big). \tag{5}$$

It is known [4] that an arbitrary recognition algorithm A can be represented as a sequential execution of the operators B (recognition operator) and C (decision rule):

$$A = B \cdot C. \tag{6}$$

It follows from (6) that the problem of finding the optimal algorithm A can be considered as a search for the optimal recognition operator B for a fixed decision rule $C(c_1, c_2)$.

The aim of this paper is to solve the problems related to calculating the values of the parameters (5) of the model for constructing extremal recognition operators based on the potentials principle in condition of features correlations. This uses a heuristic approach based on the consistent application of local optimization procedures at each stage.

The formal description of the problem of determining the parameters $\tilde{\pi}$ is as follows.

2 Statement of the Problem

A model of recognition operators based on the potential principle is given. Any operator B from this model is completely determined by specifying the set of parameters $\tilde{\pi} = \Big(n', \{\tilde{w}\}, \{\overline{c}\}, \{\tilde{\omega}\}, \{\lambda_i\}, \{\xi_i\}, \{\gamma_u\}\Big)$. We denote the set of all recognition operators from the proposed model by $B(\tilde{\pi}, S)$. Then the problem of constructing extreme recognizing operators based on the potential principle can be formulated as the problem of finding the extreme operator $B(\tilde{\pi}^*, S)$ among the recognizing operators $B(\tilde{\pi}, S)$. Here $\tilde{\pi}$ is a vector of configurable parameters; $\tilde{\pi}^*$ is the vector of optimal parameters. The recognition quality criterion is given in the form

$$\varphi_A(\tilde{\pi}) = \frac{1}{m} \sum_{j=1}^{m} \mathfrak{K}\Big(\|\tilde{\alpha}(S_j) - A(\tilde{\pi}, S_j)\|_B\Big), \tag{7}$$

$$\mathfrak{K}(x) = \begin{cases} 0, \text{ for } x = 0; \\ 1, \text{ otherwise}, \end{cases}$$

where m - remount of training set; $\|\cdot\|_B$ norm of a Boolean vector.

Then the problem of constructing extreme recognition operators consists in finding the optimal value of the components of the vector-parameter $\tilde{\pi}^*$ for a given model of recognizing operators ensuring the fulfillment of the following condition:

$$\tilde{\pi}^* = \arg\min_{\tilde{\pi}} \varphi_A(\tilde{\pi}). \tag{8}$$

Thus, the problem of determining parameters in the construction of extreme recognition operators is reduced to the optimization problem (8).

3 Proposed Method

For solving the formulated problem, it is reduced to finding the optimal values of the parameters at each stage. After each iteration, the value of the quality functional (7) is calculated. If it is less than the specified threshold or the number of iterations is greater than the specified one, the search procedure stops. We consider procedures for determining the values of the parameters of each stage separately.

3.1 The Procedure for Determining Subsets of Strongly Correlated Features

Let $\mathfrak{T}_q$ $(q = \overline{1, n'})$ be subset of strongly correlated features. The proximity measure $\mathfrak{L}\Big(\mathfrak{T}_p, \mathfrak{T}_q\Big)$ between the subsets $\mathfrak{T}_p$ and $\mathfrak{T}_q$ can be given in various ways, for example:

$$\mathfrak{L}\Big(\mathfrak{T}_p, \mathfrak{T}_q\Big) = \frac{1}{(N_p - 1)(N_q - 1)} \sum_{x_i \in \mathfrak{T}_p} \sum_{x_i \in \mathfrak{T}_p} \sum_{k=1}^{m} \upsilon_k d_k(x_i, x_j), \tag{9}$$

$$N_p = \operatorname{card}\Big(\mathfrak{T}_p\Big), N_q = \operatorname{card}\Big(\mathfrak{T}_q\Big),$$

where $d_k(x_i, x_j)$ a proximity measure between the features x_i and x_j over the k-th object.

Determining of $\mathfrak{G}_B = \Big\{\mathfrak{T}_1, \mathfrak{T}_2, \ldots, \mathfrak{T}_{n'}\Big\}$ is carried out as follows.

Step 1. The first step assumes that each subset contains only one element. In this case we have the following n subsets:

$$\mathfrak{T}_1 = \{x_1\}, \mathfrak{T}_2 = \{x_2\}, \ldots, \mathfrak{T}_n = \{x_n\} \quad (N_1 = N_2 = \ldots = N_n = 1). \tag{10}$$

We define the initial link matrix $\|\mathfrak{L}_{ij}^1\|$ as $\mathfrak{L}_{ij}^1 = b_{ij}$ [12]. Next, we consider the execution of an arbitrary u-th step $(k > 1)$.

Step u. We suppose that u' subsets $\mathfrak{T}_1, \ldots, \mathfrak{T}_{n'}$ are defined and the link matrix $\|\mathfrak{L}_{ij}^{(u-1)}\|_{u' \times u'}$ is constructed at the step $(u-1)$, where $u' = n - u + 1$.

Then at the u-th step the following operations are performed:

1) combining of $\mathfrak{T}_p$ $\mathfrak{T}_q$ into one subset, if

$$\mathfrak{L}(\mathfrak{T}_p, \mathfrak{T}_q) = \max\|\mathfrak{L}_{ij}^{(u-1)}\|_{u' \times u'}, \text{where } i, j \in [1, 2, \ldots, u'] \text{ and } i \neq j; \tag{11}$$

2) formation of a new u-th order communication matrix: $\|\mathfrak{L}_{ij}^u\|$.

The process of combining features continues until n' subsets (n' - some given number), i.e. n' independent subsets of the features $\mathfrak{T}_1, \mathfrak{T}_2, \ldots, \mathfrak{T}_{n'}$, where each feature is strongly correlated in its subset, are obtained.

3.2 The Procedure for Determining a Representative Feature in each Subset of Strongly Correlated Features

At this stage, different methods can be used to select uncorrelated (representative) features from a subset of strongly correlated features. The main idea of choosing is to allocate the most "independent" (or weakly dependent) set of features.

Let $\mathfrak{T}_q$ $(q = \overline{1, n'})$ be subsets of strongly correlated features. It is assumed that N_q is the number of elements (power) of these subsets:

$$N_q = \text{card}(\mathfrak{T}_q), \quad (q = \overline{1, n'}). \tag{12}$$

Then the procedure of this stage can be described as follows. In the beginning it is assumed that $u = 0$.

Step 1. Selection as a representative feature of the isolated elements of the subset $\mathfrak{T}_q$, which differ sharply from other features. At this step, the following actions are performed:

- the value of u increases by one and the condition $u > n'$ is checked. If this condition is met, then the algorithm stops;

- if $N_q = 1$, then the element belonging to the subset $\mathfrak{T}_q$ refers to the number of representative features, and the transition to the previous action is performed.

Step 2. Selecting a representative feature when a subset of strongly correlated features contains more than two elements, i.e. under the condition $N_q > 2$, the following sequence of operations is performed for all elements of $\mathfrak{T}_q$, except for the element under consideration: - for each element $\mathfrak{T}_q$, the proximity of each element to other elements of a given subset of features is calculated

$$\mu_i = \sum_{j=1}^{i-1} \rho(x_i, x_j) + \sum_{j=i+1}^{N_q} \rho(x_i, x_j), \tag{13}$$

$$\rho(x_i, x_j) = \sum_{k=1}^{m} \upsilon_k d_k(x_i, x_j); \tag{14}$$

- an element of the subset $\mathfrak{T}_q$, that is closest to other elements is determined

$$\mu_j = \max_{i \in [1,\ldots,N_q]} \mu_i; \tag{15}$$

- the feature x_j is selected as a representative feature.

Step 3. When $N_q = 2$, the following actions are performed:

- for each element $\mathfrak{T}_q$, the proximity estimate for the representative elements of the other subsets that were selected at the previous selection stages is calculated:

$$\mu_{i_\tau} = \sum_{j=1}^{N_0} \rho(x_i, x_j); \quad i_\tau = 1, 2, \ldots, \kappa; \quad \tau \in [1, 2], \tag{16}$$

where κ–number of subsets that consist of two elements. N_0 is the number of isolated elements and elements selected from subsets with a power of more than two;

- element of the subset $\mathfrak{T}_q$, which is significantly different from other selected representative features, is determined:

$$\mu_j = \min_{\tau \in [1,2]} \mu_{i_\tau}; \tag{17}$$

- the feature x_j is selected as a representative feature;
- go to step 1.

As a result of this procedure, an n'-dimensional space $\mathbb{X}$ of features, each of which is representative of the distinguished subset of strongly correlated features, is formed:

$$\mathbb{X} = \left(x_{i_1}, \ldots, x_{i_q}, \ldots, x_{i_{n'}}\right). \tag{18}$$

3.3 The Procedure for Determining the Correlation Models in each Subset of Features for a Class

Let x_{i_q} be a representative feature belonging to the set $\mathfrak{T}_q$. The correlation models between the features x_{i_q} and x_i ($x_{i_q} \in \mathfrak{T}_q$, $x_i \in \mathfrak{T}_q \backslash x_{i_q}$) are defined for each class K_j in the form

$$x_i = F_j\left(\bar{c}, x_{i_q}\right), \quad x_i \in \Omega_q \backslash x_{i_q}, \tag{19}$$

where $\bar{c}$ is a vector of unknown parameters; F_j is some correlation model that belongs to some given class $\{F\}$. It is assumed that the parametric form and the number of parameters are known.

For the sake of simplicity, it is assumed that the set $\{F\}$ consists of only the linear models. It is assumed that the feature x_{i_q} ($x_{i_q} \in \mathfrak{I}_\mathfrak{q}$ is an independent variable, and the feature x_i $\left(x_i \in \mathfrak{T}_q \backslash x_{i_q}\right)$ is a dependent variable. Then the model of correlation takes the form

$$x_i = c_1 x_{i_q} + c_0, \tag{20}$$

where c_1, c_0 are unknown parameters of the correlation model.

To determine the numerical values of these parameters, we use the least square method [13].

3.4 The Procedure for Identifying Preferred Correlation Models

Let $\mathfrak{I}_0$ – be training set, E_1 – set of objects belonging to the class K_j : $E_1 = \mathfrak{I}_0 \cap K_j$, and E_2 – set of objects not belonging to the class K_j : $E_2 = \mathfrak{I}_0 \backslash E_1$. We consider the procedure for identifying an important (presumed) model of correlation on the basis of estimating the dominance of the models under consideration [10]:

$$T_i = \frac{D_i}{\Delta_i}, \tag{21}$$

Here D_i – a selective error calculated for objects not belonging to the subset $\tilde{K}_j$:

$$D_i = \frac{1}{card(E_2)} \sum_{S \in E_2} |\zeta_i(S)|, \tag{22}$$

Δ_i – the selective variance of the error calculated for objects belonging to the subset $\tilde{K}_i$:

$$\Delta_i = \frac{1}{card(E_1)} \sum_{S \in E_1} |\zeta_i(S)|, \tag{23}$$

$$\zeta_i(S) = y_i(S) - F_j\Big(\bar{c}, x_{i_q}(S)\Big), \tag{24}$$

where $x_{i_q}(S)$ – representative feature q-th subset ($x_{i_q} \in \mathfrak{I}_\mathfrak{q}$) of the object S; y_i – an arbitrary feature of the same object, except for the representative one ($y_i \in \mathfrak{T}_q \backslash x_{i_q}$).

On the basis of formula (21), we compute the dominance estimate for all the characteristics that belong to the subset $\mathfrak{I}_\mathfrak{q} \backslash x_{i_q}$. As a result, we obtain $(N_q - 1)$ values of $T_i\Big(i = \overline{1, (N_q - 1)}\Big)$. The choosing of the important correlation is conducted as follows:

$$T_q^* = max\{T_1, \ldots, T_i, \ldots, T_{(N_q - 1)}\} \tag{25}$$

Repeating this procedure for all n'' subsets, we get a set of preferred correlation models for the class K_j. It is assumed that $N_q > 1$. Otherwise, the number of preferred models decreases as much as the number of subsets has only one element.

3.5 The Procedure for Determining the Difference Function between the Object S_u and the Objects of the Class K_j

With the help of this procedure, difference function that characterizes the quantitative measure of the remoteness of the object S from the source of the potential is defined. In this case, the source of the potential is given as the model of the correlation (19) in each subset $\mathfrak{R}_q$ $(q = \overline{1, n''})$ for each class K_j.

The difference function $d_q(K_j, S)$ between objects of class K_j and object S in R_q can be defined as follows:

$$d_q(K_j, S) = \sum_{q=1}^{n''} |a_i - F_j(\bar{c}, a_q)|, \tag{26}$$

where a_i, a_q the value of the i-th and i_q-th feature, corresponding to the object S. It is assumed that the feature x_i $(i \neq i_q)$ belongs to a subset of strongly correlated features – R_q; a_q – the value of the basic (representative) feature $x_{i_q}(x_{i_q} \in \mathfrak{R}_q)$.

We consider the problem of constructing a function characterizing the quantitative estimation of the difference between the objects S_u and S_v in the subspace of strongly correlated features.

Let an admissible object in the subspace of strongly correlated features $\mathfrak{J}_q$ $(\mathfrak{J}_q = (\chi_{i_q}, \chi_1, \ldots, \chi_{q'}),\ q' = card(\mathfrak{J}_q) - 1)$ be given:

$$S = (a_{i_q}, a_1, \ldots, a_{q'}). \tag{27}$$

The difference between this object and the class K_j is defined as follows:

$$d(K_j, S) = \sum_{q=1}^{n''} \mathfrak{g}_q d_q(K_j, S), \tag{28}$$

where $\mathfrak{g}_q$ recognition operator parameter $(q = \overline{1, n''})$. The set of these parameters forms a vector $\tilde{\mathfrak{g}} = (\mathfrak{g}_1, \ldots, \mathfrak{g}_q, \ldots, \mathfrak{g}_{n''})$.

The problem is to determine the values of the unknown parameters $\{\mathfrak{g}_q\}$ $(q = \overline{1, n''})$ of the difference function (26) over a given set of objects $\tilde{S}^m$.

For this problem, we introduce a functional that characterizes the importance of $d_q(K_j, S)$:

$$R(\tilde{\mathfrak{g}}) = \frac{\sum\limits_{q=1}^{n''} \mathfrak{g}_q \mathfrak{R}_q^1}{\sum\limits_{q=1}^{n''} \mathfrak{g}_q \mathfrak{R}_q^2} \tag{29}$$

$$\mathfrak{R}_q^1 = \sum_{j=1}^{l} \left(\frac{1}{m_j} \sum_{S \in \tilde{K}_j} d_q(K_j, S) \right), \quad \mathfrak{R}_q^2 = \sum_{j=1}^{l} \left(\frac{1}{m - m_j} \sum_{S \in C\tilde{K}_j} d_q(K_j, S) \right) \tag{30}$$

Without loss of generality, we introduce restrictions on the coefficients ($\mathfrak{g}_1, \ldots, \mathfrak{g}_q, \ldots, \mathfrak{g}_{n''}$) in the form

$$\sum_{q=1}^{n''} \mathfrak{g}_q = 1. \tag{31}$$

Taking into account the imposed restrictions on the components of the vector, we can formulate the problem of determining $\overline{\mathfrak{g}}$ as follows

$$\overline{\mathfrak{g}} = \arg \min_{S_u \in \tilde{S}^m} R(\overline{\mathfrak{g}}) \tag{32}$$

To find the values of the vector $\overline{\mathfrak{g}}$ we construct a Lagrange function of the form:

$$L(\tilde{\mathfrak{g}}, \lambda) = \frac{\sum\limits_{q=1}^{n''} \mathfrak{g}_q \mathfrak{R}_q^1}{\sum\limits_{q=1}^{n''} \mathfrak{g}_q \mathfrak{R}_q^2} - \lambda \left(\sum_{q=1}^{n''} \mathfrak{g}_q \right), \tag{33}$$

where λ Lagrange multiplier.

We differentiate (33) with respect to ($i = 1, n''$) and, by equating it to zero, we obtain the following system of equations:

$$\frac{\partial L}{\partial \mathfrak{g}_i} = \frac{\mathfrak{R}_i^1 - \mathfrak{R}_i^2}{\left(\sum\limits_{i=1}^{n''} \mathfrak{g}_i \mathfrak{R}_i^2 \right)^2} + \lambda \mathfrak{g}_i. \tag{34}$$

By summing (34) over each derivative of $\mathfrak{g}_i$, we find the value λ

$$\lambda = \frac{\sum\limits_{i=1}^{n''} \left(\mathfrak{R}_i^1 - \mathfrak{R}_i^2 \right)}{\left(\sum\limits_{i=1}^{n''} \mathfrak{g}_i \mathfrak{R}_i^2 \right)^2}. \tag{35}$$

By substituting λ in (34), we compute the value g_i ($i = \overline{1, n''}$:

$$\mathfrak{g}_i = \frac{\mathfrak{R}_i^1 - \mathfrak{R}_i^2}{\sum\limits_{i=1}^{n''} \left(\mathfrak{R}_i^2 - \mathfrak{R}_i^2 \right)}. \tag{36}$$

Thus, by repeating the calculation process for all features, the difference function is determined.

3.6 The Procedure for Determining the Proximity Function between the Object S_u and the Objects of the Class K_j

Let the proximity function between the objects S_u and S be given as potential functions of the second type [14]:

$$f(\xi, d(K_j, S)) = \frac{1}{1 + \xi d(K_j, S)}. \tag{37}$$

The main tasks in constructing the proximity function between objects on the basis of a potential function is to determine the value of the smoothing parameter ξ. However, the search for the optimal value of this parameter in the creation of a potential function has not been studied enough, and for each specific case, elements of search and creativity are contained in finding ξ [15, 16]. The authors of [15] assert that, depending on ξ, the resolving power of the potential function is located. So, as ξ increases, fast decay of the potential function occurs by moving away from objects of the given class, which leads to the creation of a "relief" peaked above the "top" of this image. A decrease in ξ results not only to smoothing out the peaks, but also to the leveling the "heights" of images of different classes, which makes recognition difficult and leads to a large number of vague and erroneous solutions.

To calculate the smoothing parameter of a potential function, the approach, the idea of which is borrowed from [16], is used.

For the sake of simplicity, we introduce some restrictions (ie. we consider the problem for only two classes) and the notations:

$$\tilde{S}^m = \tilde{K}_1 \cup \tilde{K}_2, \tilde{K}_1 \cap \tilde{K}_2 = \varnothing, \tilde{K}_1 \neq \varnothing, \tilde{K}_2 \neq \varnothing; \tag{38}$$

$$d(\tilde{K}_j, S) = \begin{cases} a_u, \text{ if } P_j(S) = 1; \\ b_u, \text{ if } P_j(S) = 0, \end{cases} \tag{39}$$

where a_u - measure of the difference between the object S and the objects of the class $\tilde{K}_j$, when S belongs to the class $\tilde{K}_j$ ($a_u \geqslant 0$); b_u is the measure of the difference between the object S and the objects of the class $\tilde{K}_j$, when S does not belong to the class $\tilde{K}_j$ ($b_u \geqslant 0$).

Then the problem of determining the smoothing parameter is formulated as follows:

$$\xi = \arg \max_{S_u \in \tilde{S}^m} \mathfrak{R}(\xi), \tag{40}$$

$$\mathfrak{R}(\xi) = \left(\frac{1}{M_0} \sum_{u=1}^{M_0} \frac{1}{1 + \xi a_u} + \frac{1}{N_0} \sum_{u=1}^{N_0} \frac{1}{1 + \xi b_u} \right)^2, \tag{41}$$

$$\xi \in (0, \xi_{max}), \xi_{max} = max(a_{max}, b_{max}), \tag{42}$$

$$a_{max} = \max_{a_u \in [1, M_0]} \{a_u\}, b_{max} = \max_{b_u \in [1, N_0]} \{b_u\}, \tag{43}$$

$$M_0 = 0.5(m_1(m_1 - 1) + m_2(m_2 - 1)), N_0 = m_1 \times m_2, \tag{44}$$

$$m_1 = \text{card}(\tilde{K}_j), m_2 = m - m_1. \tag{45}$$

To solve problem (10) by taking into account the unimodality of function (11) (as confirmed by experimental studies), we use the Fibonacci Search method [17].

Thus, with the application of the proposed procedures, the values of all parameters of the considered model of recognition operators are determined. To assess the operability of the heuristic approach examined, experimental studies were conducted.

4 Experiments

An experimental study of the efficiency of the proposed approach in the construction of recognition operators was carried out on the example of solving a number of problems, in particular, the model problem, the problem of person identification by the geometrical features of a photo portrait and the problem of diagnosing cotton diseases from leaf images.

As test models of recognition operators, the following were chosen: the classical model of recognition operators of the type of potential functions (B_1), the model of recognition operators based on the calculation of estimates (B_2), and the model of recognition operators based on the evaluation of the features correlations (B_3) [10].

A comparative analysis of the above mentioned models of recognition operators in solving problems was carried out according to the following criteria: accuracy of recognition of test sample objects; time spent on training; time spent on recognizing objects from the test sample.

To calculate the above criteria in order to exclude the successful (or unsuccessful) partitioning of the initial sample $\mathfrak{B}$ to $\mathfrak{B}_o$ and $\mathfrak{B}_k$ sets ($\mathfrak{B} = \mathfrak{B}_0 \cup \mathfrak{B}_k$, $\mathfrak{B}_0$ is the training sample, $\mathfrak{B}_k$ is the test sample), the cross-validation method was used [18]. The experiments were carried out on a Pentium IV Dual Core 2.2 GHz computer with 1 Gb of RAM.

4.1 Model Problem

The initial data of the recognized objects for the model problem are generated in the space of the correlated features. The number of classes in this experiment is equal to two. The volume of the initial sample is 1000 implementations (500 implementations for objects of each class). The number of features in the model example is 200. The number of subsets of strongly correlated features is 5.

4.2 The Problem of Person Identification

The initial data used for the problem of person identification by the geometric features of a photo portrait consist of 360 portraits. The number of classes in this experiment is equal to six. Each photo portrait is characterized by the corresponding parameters (see Fig. 1): the distance between the center of the retina of the left eye and the center of the tip of the nose; distance between the center of the retina of the left eye and the center of the oral opening; distance between the center of the retina of the right eye and the point of the tip of the nose; distance between the centers of the retina of the eyes, etc.

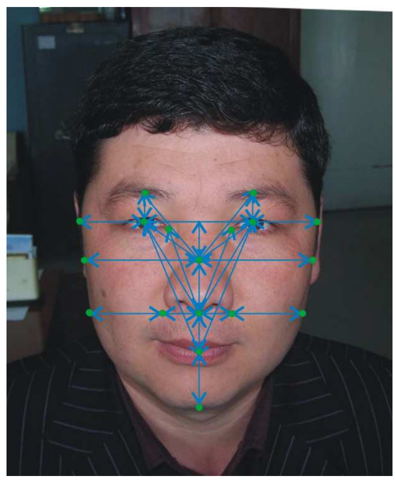

Fig. 1. Examples of some distances between anthropometric points.

There are 60 different one-person portraits photographed at different times, but with the same shooting conditions in each class. To distinguish these parameters, we used the algorithm for searching for characteristic features of the face, described in [19–21].

4.3 The Problem of Diagnosis of Cotton Diseases

The initial data used in the problem of diagnosing cotton diseases from leaf images consist of 200 images. The number of classes in this experiment is equal to two:

- images of cotton leaves, sick with wilt (K_1);
- images of cotton leaves without wilt (K_2).

The number of images in each class is the same and is equal to 100. To extract the features characterizing the phytosanitary state of cotton by the original image of the leaves, we used the algorithm described in [22].

5 Results

As mentioned earlier, all tasks were solved using the recognition operators B_1, B_2 and B_3.

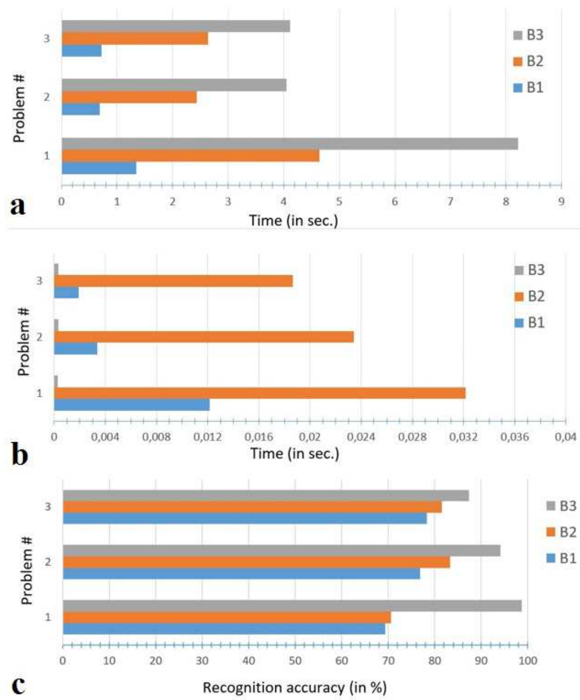

Fig. 2. Indicators of the speed of training (a) and recognition (b) and recogtion accuracy (c).

Fig. 2,a shows the training speed of the recognition model on the training sample of the problems under consideration, and Fig. 2,b shows the speed of recognition of objects. The results of solving the problems under consideration with the use of B_1, B_2 and B_3 during the test are shown in Fig. 2,c.

A comparison of these results shows (see Fig.2) that the model of recognition operators B_3 allowed to increase the accuracy of recognition of objects described in the space of correlated features (more than 6-10% higher than B_1 and B_2). This is because the models B_1 and B_2 do not take into account the features correlations. However, for model B_3, there is some increase in training time, which requires further investigation.

6 Discussion

The developed procedures are oriented to determine unknown parameters within the model of recognition operators, which differ from traditional recognition operators such as potential functions in that they are based on the evaluation of the features correlations. Therefore, it is advisable to use these procedures in those cases when there is some correlation between the features. Undoubtedly, this correlation should be different for objects of each class. This allows to describe the objects of each class with an individual model. If the relationship between the features is weak, then the classical model of recognition operators is used (for example, the model considered in [4, 1, 5]). Consequently, the models of recognition operators considered in [10] are not an alternative to models of recognition operators of the type of potential functions, but only complement them.

In the case when a sufficiently strong correlation is found between the features of all the objects under consideration, then in the process of forming a set of representative features (described in the first and second stages of specifying the model), the features repeating the same information are excluded, which ensures the selection of features that are sufficiently representative of all those features that are not contained in the given set [10].

The results of the conducted experimental research show that the proposed optimization procedures for constructing extreme recognition operators allows to solve the problem of pattern recognition more accurately in conditions of features correlations.

7 Conclusions

Procedures for constructing an extreme recognition operator based on potential functions are proposed in the context of features correlations. These procedures allow to expand the scope of application of recognition operators based on potential functions.

The results of solving a number of problems have shown that the proposed procedures for determining parameters in the construction of an extreme recognition operator improve accuracy and significantly reduce the number of computational operations by recognizing an unknown object given in the space of correlated features.

Acknowledgement. The work was carried out with partial financial support of the grant BV-M-F4-003 of the State Committee for the Coordination of the Development of Science and Technology of the Republic of Uzbekistan.

References

1. Zhuravlev, Yu.I., Ryazanov, V.V., Senko, O.V.: Recognition. Mathematical methods. Software system. Practical applications. Fazis, Moscow (2006), (in Russian)

2. Eremeev, A.V.: Methods of discrete optimization in evolutionary algorithms. In: 5th All-Russian conference on Optimization Problems and Economic Applications. pp. 17–21. Omsk (2012) (in Russian)
3. Kelmanov, A.V.: On some hard-to-solve problems of cluster analysis. In: 5th All-Russian conference on Optimization Problems and Economic Applications. pp. 33–37. Omsk (2012) (in Russian)
4. Zhuravlev, Yu.I.: Selected Scientific Works. Magister, Moscow (1998) (in Russian)
5. Merkov, A.: Pattern Recognition: An Introduction to Statistical Learning Methods. Editorial URSS, Moscow (2011) (in Russian)
6. Murty, M.N., Devi, D.V.S.: Introduction to Pattern Recognition and Machine Learning. World Scientific, New Jersey (2015)
7. Cuevas, E., Zaldivar, D., Perez-Cisneros, M.: Applications of Evolutionary Computation in Image Processing and Pattern Recognition. Springer, New York (2016)
8. Pattern Recognition Techniques: Technology and Applications. Edited by Yin, P.-Y. lTexLi, New York (2016)
9. Fazilov, Sh.Kh., Mirzaev, O.N., Radjabov, S.S.: State of the art of the problems of pattern recognition. J. Problems of computational and applied mathematics 2, 99–112 (2015), (in Russian)
10. Fazilov, Sh.Kh., Mirzaev, N.M., Mirzaev, O.N.: Building of recognition operators in condition of features correlations J. Radio Electronics, Computer Science, Control 1, 58–63 (2016), (in Russian)
11. Mirzaev, N.M., Radjabov, S.S., Jumayev, T.S.: On the parametrization of models of recognition algorithms based on the evaluation of the features correlations. J. Problems of Informatics and Energetics 2-3, 23–27 (2016) (in Russian)
12. Mirzaev, O.N.: Extraction of subsets of strongly correlated features in the construction of extreme recognition algorithms. J. Bulletin of Tashkent University of Information Technologies 3, 145–151 (2015) (in Russian)
13. Draper, N.R., Smith, H.: Applied Regression Analysis. Wiley, New York (1998)
14. Tou, J.T., Gonzalez, R.C.: Principles of Pattern Recognition. Addison Wesley, Reading (1974)
15. Ogibalov, P.M., Mirzadjanzade, A.Kh.: Mechanics of Physical Processes. MSU, Moscow (1976) (in Russian)
16. Fazilov, Sh.Kh., Rakhmanov, A.T., Mirzaev, O.N.: Determination of the smoothing parameter of a potential function in the construction of recognition algorithms. In: 12th International conference on Informatics. pp. 412–414. Voronezh (2012), (in Russian)
17. Cottle, R.W., Thapa, M.N.: Linear and Nonlinear Optimization. Springer, New York (2017)
18. Braga-Neto, U.M., Dougherty, E.R.: Error Estimation for Pattern Recognition. Springer, New York (2016)
19. Kuharev, G.A., Kamenskaya, E.I., Matveev, Yu.N., Shegoleva N.L.: Methods for Processing and Recognizing Faces in Biometrics. Politexnika, St. Petersburg (2013), (in Russian)
20. Starovoitov, V., Samal, D.: A geometric approach to face recognition. In: The IEEE EURASIP Workshop on Nonlinear Signal and Image Processing. pp. 210–213 (1999)
21. Wang, N., Ke, L., Du, Q., Liang, L.: Face recognition based on geometric features. In: International Conference on Computational Science and Engineering Applications. pp. 641–647. Sanya (2015)

22. Mirzaev, N.M.: Feature extraction in the problem of diagnosing of plants phytocondition by leaves images. J. Vestnik of Ryazan state radioengineering 3 (41), 21–25 (2012), (in Russian)