Estimates Calculations Algorithms in Condition of Huge Dimensions of Features' Space

Mirzoyan Kamilov[1], Shavkat Fazilov[2], Nomaz Mirzaev[3], Sobirjon Radjabov[4]
Institute of Mathematics and Information Technologies, Academy of Sciences of the Republic of Uzbekistan
[1]*komilov@yandex.ru*, [2]*sh.fazilov@mail.ru*, [3]*mnm2005@rambler.ru*, [4]*s_radjabov@yahoo.com*

***Abstract*— The problem of pattern recognition in condition of huge dimensions of features' space is considered. Extended model of recognition algorithms on the base of estimates' calculations algorithm is proposed.**

***Keywords*— recognition algorithm; features correlations; elementary threshold rule; representative feature; correlation model**

I. Introduction

Analysis of pattern recognition literature [1-9] shows that currently the problem of building of recognition algorithms in condition of huge dimensions of features' space is not studied enough. The experience on solving of the problem of pattern recognition in condition of huge dimensions of features' space shows that objects' features are often correlated. That is why the problem of building of recognition algorithms in condition of huge dimensions of features' space is urgent problem.

The goal of this work is to improve estimates calculations based recognition algorithms for classification of objects, given in the features' space of huge dimensions. For this the approach on the base of the method of algebraic extension of recognition algorithms is used. This method was proposed by academician of Russian Academy of sciences Yu.I. Juravlev.

II. Statement of the Problem

We introduce following terms and definitions based on [2]. Let Γ be set of possible objects, which consist of l disjoint subsets

$$K_1, K_2, ..., K_l,\ \ K_i \cap K_j = \varnothing, i \neq j,\ i, j \in \{1,...,l\}.$$

Knowledge about the division of Γ is not complete, only some initial information J_0 about classes is given.

Let objects $S_1,...,S_i,...,S_m$ $(\forall S_i \in \Gamma, i = \overline{1,m})$ are given in the given features' space X ($X = (x_1,...,x_j,...,x_n)$): $S_1 = (a_{11},...,a_{1j},...,a_{1n})$, … , $S_i = (a_{i1},...,a_{ij},...,a_{in})$, … , $S_m = (a_{m1},...,a_{mj},...,a_{mn})$. We introduce following definitions:

$$\widetilde{S}^m = \{S_1,...,S_i,...,S_m\},\ \widetilde{K}_j = \widetilde{S}^m \cap K_j,\ C\widetilde{K}_j = \widetilde{S}^m \setminus \widetilde{K}_j.$$

Then initial information J_0 could be given as

$$J_0 = \{S_1,...,S_i,...,S_m;\ \widetilde{\alpha}(S_1),...,\widetilde{\alpha}(S_i),...,\widetilde{\alpha}(S_m)\},$$

where $\widetilde{\alpha}(S_i)$ is information vector of the object S_i, it is given as

$$\widetilde{\alpha}(S_i) = (\alpha_{i1},...,\alpha_{ij},...,\alpha_{il}),\ \alpha_{ij} = \begin{cases} 1, \text{ if } S_i \in \widetilde{K}_j; \\ 0, \text{ if } S_i \notin \widetilde{K}_j. \end{cases}$$

The set of information vectors corresponding to the objects $\widetilde{S}^m$ forms information matrix $\|\alpha_{ij}\|_{m\times l}$.

It is assumed that for any object $S \in \Gamma$ the numerical characteristics $J(S) = (a_1,...,a_i,...,a_n)$ of the object could be matched in the feature space $X = (x_1,...x_i,...,x_n)$.

We consider random set of objects $\widetilde{S}^q = \{S_1^{'},...,S_q^{'}\}$ ($\widetilde{S}^q \subset \Gamma$). Let objects $\widetilde{S}^q$ is given in the space X. Here the dimensions of the features' space n is huge (for example, $n > 200$). The problem is to build the algorithm A, which calculates the value of the predicate $P_j(S_i^{'})$ by the initial information J_0. In other words, required algorithm A transforms the set $(J_0, \widetilde{S}^q)$ to the matrix $\|\beta_{ij}\|_{qxl}$ ($\beta_{ij} = P_j(S_i^{'})$, $P_j(S_i^{'})=\text{"}S_i^{'} \in K_j\text{"}$) [2]:

$$A(J_0, \widetilde{S}^q) = \|\beta\|_{qxl},\quad \beta \in \{0,1,\Delta\}.$$

Here β_{ij} is interpreted as follows. If $\beta_{ij} = \Delta$, then the algorithm is not able to calculate the value of the predicate $P_j(S_i^{'})$. If $\beta \in \{0,1\}$ then β_{ij} is the value of the predicate $P_j(S_i^{'})$, calculated by the algorithm A for the object $S_i^{'}$ by using its numerical characteristics.

III. Solving Method

In this paper the research results on building of recognition algorithms on the base of estimates calculations considering the features' correlations are presented. It is assumed that because of the huge dimension of feature space the likelihood of correlations between them increases. Proposed algorithms can be divided into three types based on the features' correlations: 1) algorithms given in the set of uncorrelated features (the fist type of recognition algorithms); 2) algorithms on the base of estimates' of features' correlations (the second type of recognition algorithms); 3) generalized algorithms, which are built on the base of integration of the first and the second type of recognition algorithms (the third type recognition algorithms).

Next the detailed description of these algorithms is given.

A. *The first type of recognition algorithms*

These algorithms consist of following nine stages:

1. Forming of subset of strongly correlated features. Let Ξ_q $(q=\overline{1,n'})$ - subset of strongly correlated features. Distance measure $L(\Xi_p, \Xi_q)$ between subsets Ξ_p and Ξ_q could be given by using different methods, for example [5]:

$$L(\Xi_p, \Xi_q) = \frac{1}{N_p \cdot N_q} \sum_{x_i \in \Xi_p} \sum_{x_j \in \Xi_q} \eta(x_i, x_j),$$

Where N_p, N_q are numbers of features of the subsets Ξ_p, Ξ_q respectively; $\eta(x_i, x_j)$ – a function that characterizes the strength of mutual correlations between the features x_i and x_j [10].

As the result of this stage we get the set of "uncorrelated" subsets of strongly correlated features $W_A = \{\Xi_1, \Xi_2, ..., \Xi_{n'}\}$.

2. Defining of representative features in each subset of strongly correlated features. In this stage the set of representative features $y_1, y_2, ..., y_{n'}$ is formed. Each representative feature is typical representative of the extracted subset of strongly correlated features [10].

3. Defining preferred features. Selecting of preferred features from representative features $\{y_1, ..., y_i, ..., y_{n'}\}$, defined in the previous stage is carried out on the base of dominancy of each feature, which divides objects of the set $\widetilde{S}^m$ into two subsets $\widetilde{K}_j$ and $C\widetilde{K}_j$ [11]:

$$\Re_{ij} = \frac{\hat{N}_2 \sum_{j=1}^{2} \sum_{S \in \widetilde{K}_j} \sum_{S_u \in \widetilde{K}_j} (a_i - a_{iu})^2}{\hat{N}_1 \sum_{S \in \widetilde{K}_j} \sum_{S_u \in C\widetilde{K}_j} (a_i - a_{iu})^2},$$

$\hat{N}_1 = (m_1(m_1-1) + m_2(m_2-1))/2$, $\hat{N}_2 = m_1 \times m_2$, $m_1 = |\widetilde{K}_j|$, $m_2 = |C\widetilde{K}_j|$.

The smaller the value $\Re_{ij}$, the greater the preference gets the appropriate feature in separation of objects belonging to $\widetilde{K}_j$. In calculating $\Re_{ij}$ it is suggested that S and S_u – different objects (i.e. $S \neq S_u$).

Preferred feature, which is denoted by $\overline{\chi}_j$ ($\overline{\chi}_j = (\chi_1, \chi_2, ..., \chi_{n'})$), is determined in this stage for each subset $\widetilde{K}_j$. Next, we consider only the preferred features.

4. Defining the system of supporting sets. Let $H_{\widetilde{\omega}}$ - be all the possible subsets of the set $\{y_1, y_2, ..., y_{n'}\}$. We define collection of such subsets by Ω. Fourth stage of the first type recognition algorithms is defining the system of supporting sets Ω_A $(\Omega_A \subseteq \Omega)$.

5. Defining the distance function between objects. Let's consider possible objects S and S_u. In the fifth stage of the first type recognition algorithms distance function $\mu_{\widetilde{\omega}}(S, S_u)$ between objects S and S_u in the $\widetilde{\omega}$-part of the features' space is defined.

6. Calculating of estimates on the objects of fixed supporting set. In the sixth stage of the first type recognition algorithms numerical characteristic called estimate $\Gamma_{\widetilde{\omega}}(S, S_u)$ is calculated:

$$\Gamma_{\widetilde{\omega}}(S, S_u) = \lambda_u \mu_{\widetilde{\omega}}(S, S_u),$$

where λ_u is given parameter of the algorithm.

7. Calculating of estimates for the class on the fixed supporting set. Let's assume that values $\Gamma_{\widetilde{\omega}}(S, S_u)$ ($S_u \in \widetilde{K}_j$) were calculated. Estimate for the class is determined as:

$$\Gamma_{\widetilde{\omega}}(S, K_j) = \sum_{S_u \in \widetilde{K}_j} \gamma_u \Gamma_{\widetilde{\omega}}(S, S_u),$$

where γ_u is given parameter of the algorithm.

8. Estimate for the class K_j on the system of supporting sets. Let numerical parameter $\tau(\widetilde{\omega})$ corresponds for each vector $\widetilde{\omega}$. Estimate on the system of supporting sets Ω_A is defined as

$$\Gamma(S, K_j) = \sum_{\widetilde{\omega} \in \Omega_A} \tau(\widetilde{\omega}) \Gamma_{\widetilde{\omega}}(S, K_j).$$

9. Decision rule. In the last stage of the first type recognition algorithms decision rule is defined as [2]:

$$\beta_{ij} = C(\Gamma(S_i, K_j)) = \begin{cases} 0, \text{ if } \Gamma(S_i, K_j) < c_1, \\ 1, \text{ if } \Gamma(S_i, K_j) > c_2, \\ \Delta, \text{ if } c_1 \leq \Gamma(S_i, K_j) \leq c_2. \end{cases}$$

where c_1, c_2 are parameters of the algorithm.

B. *Second type recognition algorithms*

This algorithm consists of eight stages. The first and the second stages of the second type recognition algorithms correspond to the first and the second stages of the abovementioned algorithm. That is here we consider remaining stages of the algorithm.

3. Determining the models of correlations in each subset of features for the class K_j $(j=\overline{1,l})$. Let x_i - be any feature belonging to the subset Ξ_q and $x_{i_0} = \arg\max_{x_j \in \Xi_q} \sum_{x_i \in \Xi_q \setminus x_j} \eta(x_i, x_j)$. Then model of correlations in Ξ_q could be defined as

$$x_i = F(\bar{c}, x_{i_0}), \quad x_i \in \Xi_q \setminus x_{i_0},$$

where $\bar{c}$ - vector of unknown parameters, F- function from some given class $\{F\}$.

4. Extraction of preferred correlation models. Let N_q – be power of the subset Ξ_q of strongly correlated features. It is assumed that $(N_q - 1)$ models of correlations is defined in Ξ_q for the class $\widetilde{K}_j$:

$$\chi_q = F(\bar{c}, x_i), \quad x_i \in \Xi_q \setminus \chi_q, \; i = \overline{1,(N_q - 1)},$$

where χ_q is representative feature ($\chi_q \in \Xi_q$).

The search for the preferred models of correlations in Ξ_q is carried out on the base of estimating dominancy of the considering models for the objects belonging to the set $\widetilde{S}^m$:

$$T_i = \frac{m_j \sum_{S \in C\widetilde{K}_j} \left|\chi_q - F(\bar{c}, x_i)\right|}{\widetilde{m}_j \sum_{S \in \widetilde{K}_j} \left|\chi_q - F(\bar{c}, x_i)\right|}, \; m_j = \left|\widetilde{K}_j\right|, \; \widetilde{m}_j = \left|C\widetilde{K}_j\right|.$$

After completing this stage preferred model of correlations for the features' subset Ξ_q is determined, and it is defined as $\chi_q = F(\bar{c}, x_{i_0})$ Hereinafter only these models of correlations are considered.

5. Determining elementary threshold rules of decision making. We denote all correlation models in the features' set Ξ_q by $\Im_q$. We define elementary threshold rules of decision making δ_i ($i = 1,...,N_q - 1$) on the base of correlation models $x_i = F_j(\bar{c}, x_{i_0})$ ($F_j(\bar{c}, x_{i_0}) \in \Im_q, \; j = \overline{1,l}$) in Ξ_q:

$$\delta_i(K_j, S) = \begin{cases} 1, & \text{if } \left|a_i - F(\bar{c}, a_{i_0})\right| < \Delta_i; \\ 0, & \text{else}, \end{cases}$$

where a_{i_0}, a_i are the values of i_0-th and i-th feature of the object S; Δ_i is given threshold.

6. Estimate $B_q(K_j, S)$ for the class K_j by the set of strongly correlated features Ξ_q. Distance function $B_q(K_j, S)$ for the class K_j ($j = \overline{1,l}$) could be given by using different methods, for example:

$$B_q(K_j, S) = b_j \delta_q(K_j, S) - \sum_{\substack{i=1 \\ i \neq j}}^{l} \widetilde{b}_i \delta_q(K_j, S),$$

where $b_j, \widetilde{b}_i$ are parameters of the algorithm.

7. Estimate for the class by the set of strongly correlated features. The estimate for belonging of an object to the class K_j ($j = \overline{1,l}$) is calculated as:

$$B(S) = (\mu_1(S),...,\mu_2(S),...,\mu_l(S)),$$

$$\mu_i(S) = \sum_{u=1}^{n'} \gamma_u B_u(K_j, S),$$

where γ_u is parameter of the algorithm ($u = \overline{1,n'}$), n' - set of strongly correlated features.

8. Decision rule. Decision rule $C(c_1, c_2)$ is given by the same way as that of the first type recognition algorithms (see the ninth stage of the first type recognition algorithms).

C. *The third type recognition algorithms*

This type recognition algorithm is based on the principle of integration of the first and the second type recognition algorithms. Forming of generalized algorithms is curried out on the base of algebraic method of building of recognition algorithms [2].

Let random type (the first or the second) recognition operators $B_u(J_0, \widetilde{S}^q) = \left\|a_{ij}^u\right\|_{qxl}$ ($u \in [1,2]$) are given. For these operators we introduce the operations of addition and multiplication by a scalar:

$$B_1(J_0, \widetilde{S}^q) + B_2(J_0, \widetilde{S}^q) = \left\|a_{ij}^1 + a_{ij}^2\right\|_{q \times l},$$

$$gB_u(Z) = \left\|g \times a_{ij}^u\right\|_{q \times l},$$

where g is real constant.

These operations make it possible to build such recognition operator that allows to more accurately solving the recognition problem when the number of features is sufficiently large.

Formation of recognition operators within the generalized pattern recognition algorithms are carried out in various ways, such as the sum of two operators:

$$B(S) = g_1 B_1(S) + g_2 B_2(S),$$

where B_1 and B_2 are the first and the second type recognition operators respectively; g_1, g_2 are parameters of the generalized algorithm ($0 \le g_1, g_2 \le 1$).

Decision rule of the generalized algorithm $C(c_1, c_2)$ is given by the same way as that of the first and the second types recognition algorithms (see the eighth stage of the first type recognition algorithms).

We defined class of the generalized recognition algorithms based on the calculation of estimates. Any algorithm A from this model is fully defined by the set of parameters $\pi = (n', \{\lambda_u\}, \{\bar{c}\}, \Delta_i, \{\gamma_u\}, \{b_j, \widetilde{b}_i\}, \{\tau\}, g_1, g_2, c_1, c_2)$. We define the collection of all recognition algorithms form the proposed model by $A(\pi, S)$. The search for the best algorithm is carried out in the parameters' space π, and their effectiveness (as well as all the heuristic algorithms) is determined by the results of their application in solving a number of practical problems.

IV. Experimental Results

For practical use, and verify the performance of considered algorithms the recognition software have been developed. The efficiency of the developed software is tested on solving of practical problems.

There is no doubt that the proposed algorithms perform better than classical recognition algorithms in condition of features' correlations. To show this, we consider the problem of person identification on the base of face images as a practical example for checking the efficiency of the developed recognition algorithms. Choosing of this problem can be justified for the following reasons. Typically, the human face is symmetrical, and therefore the characteristic features of the face (i.e. the distance between the points, selected as features) have some correlations.

The initial set of features in solving of this problem is the distance between anthropometric points in the image [10, 12].

The number of classes is $l = 4$, sample size – 400 objects, the numbers of objects in the classes are different (between 80 to 120 objects).

Conducted experimental studies in solving the problem of person recognition on the base of face images showed a higher effectiveness of the proposed algorithms in comparison with the classical estimates calculation algorithms (recognition accuracy is 9.7% higher).

V. CONCLUSIONS

The generalized recognition algorithms based on the estimates calculation were developed on the base of conducted experiments. The developed recognition algorithms can be used to create computer systems aimed at solving applied problems of computer vision, medical and technical diagnosis, biometric identification.

REFERENCES

[1] M.A. Izerman, E.M. Braverman, L.I. Rozonoer. Potential functions' method in the theory of machine learning. "Nauka" Publishing House, Moscow (1970), 348 p. (in Russian).

[2] Yu.I. Zhuravlev. Selected papers. "Magistr" Publishing House, Moscow (1998), 420 p. (in Russian).

[3] Yu.I. Zhuravlev, M.M. Kamilov, Sh.E. Tulaganov. Estimates calculation algorithms and their application. "Fan" Publishing House, Tashkent (1974), 119 p. (in Russian).

[4] Yu.I. Zhuravlev, V.V. Ryazanov, O.V. Senko. Recognition. Mathematical methods. Software system. Practical application. "Fasiz" Publishing House, Moscow (2006), 159 p. (in Russian).

[5] M.M. Kamilov, Sh.Kh. Fazilov, N.M. Mirzaev. Recognition algorithms on the base of features' correlations' estimate // In Proc. of "Mathematical methods of pattern recognition", Moscow (2007), pp.140-143 (in Russian).

[6] G.S. Lbov, N.G. Starseva. Logical decision functions and the problems of statistical stability of solutions. IM SB RAS Publishing House, Novosibirsk (1999), 211 p. (in Russian).

[7] Duda R.O., Hart P.E., Stork D.G. Pattern Classification, Second Edition. John Wiley, Inc., New York (2001), 680 p.

[8] Vapnik V. Statistical Learning Theory. John Wiley Sons, Inc., New York (1998), 732 p.

[9] M.M. Kamilov, N.M. Mirzaev, S.S. Radjabov. Urgent problems of building of models of recognition algorithms // Chemical technology. Control and management, Tashkent (2009), pp.67-72 (in Russian).

[10] M.M. Kamilov, Sh.Kh. Fazilov, N.M. Mirzaev, S.S. Radjabov. Person recogntion algorithms on the base on face image's geometric features' correlations' estimate // Problems of informatics and energetics, Tashkent (2007), №5-6, pp. 3-11 (in Russian).

[11] Sh.Kh. Fazilov, N.M. Mirzaev, O.N. Mirzaev. About one model of modified potential functions type recognition algorithms // In Proc. of "Mathematical methods of pattern recognition", Moscow (2009), pp. 200-203 (in Russian).

[12] D.I. Samal. Building of person identification systems on the base of anthropometric points of a face // In Proc. of "Digital image processing", Minsk (1998), pp. 72-78 (in Russian).