

Models of recognition algorithms based on linear threshold functions

Sh Fazilov¹, R Khamdamov¹, G Mirzaeva¹, D Gulyamova¹, N Mirzaev^{1,2}

¹Tashkent university of information technologies named after Muhammad al-Khwarizmi, 108, Amir Temur street, Tashkent, 100200, Republic of Uzbekistan

²E-mail: nomazmirza@rambler.ru

Abstract. The problem of constructing a model of recognition algorithms in the conditions of high dimensionality of feature space is considered. To solve this problem, a new approach suggests that it takes into account the interconnectedness of the given features. The main concern of this approach is to isolate representative traits and form preferred combinations of traits. A distinctive feature of the proposed model of algorithms is to determine a suitable set of elementary threshold rules within the framework of the preferred combinations of representative features and build a proximity function based on this set of feature combinations. Scientifically, the results of this work in the aggregate represent a new solution to a scientific problem related to improving the reliability of recognition algorithms based on radial functions. The practical significance of the results lies in the fact that the developed algorithms can be applied in medical and technical diagnostics, geological forecasting, biometric identification. To test the performance of the proposed model, experimental studies were performed.

1. Introduction

In recent years, the circle of specialists paying attention to the problem of pattern recognition has been constantly expanding, and the number of scientific publications on this subject has continuously increased. This is due to the fact that pattern recognition is increasingly used in science, technology, manufacturing and everyday life.

It is known [1-3] that the development of pattern recognition theory is divided into two stages. For a long time, the vast majority of applications of pattern recognition theory were associated with poorly formalized areas - medicine, geology, sociology, chemistry, etc. Therefore, at the first stage of development of recognition, many algorithms appeared that were in the nature of projects of various technical devices or algorithms for solving specific applied problems. The value of the developed pattern recognition methods is determined primarily by the achieved experimental results [2]. The second stage of development is characterized by the transition from individual algorithms to the construction of models – a family of algorithms for a unified description of methods for solving classification problems [1, 4]. The need to synthesize models of pattern recognition algorithms was determined by the need to fix the class of algorithms when choosing the optimal or at least acceptable procedure for solving a specific problem. At this stage of development, it is shown [1] that an arbitrary



recognition algorithm can be represented as a sequential execution of the operators \mathfrak{B} (recognizing operator) and \mathfrak{C} (decision rule):

$$= \mathfrak{B} \circ \mathfrak{C}. \quad (1)$$

From (1) it follows that each recognition algorithm \mathfrak{A} can be divided into two successive stages. At the first stage, the recognition operator \mathfrak{B} converts the admissible object S_u into a numerical estimate represented by the vector \tilde{b}_u :

$$\mathfrak{B}(S_u) = \tilde{b}_u, \quad (2)$$

where $\tilde{b}_u = (b_{u1}, \dots, b_{uv}, \dots, b_{u\ell})$.

At the second stage, according to the numerical estimate b_{uv} the decision rule \mathfrak{C} determines whether the object S_u to the classes $\mathcal{C}_1, \dots, \mathcal{C}_j, \dots, \mathcal{C}_l$:

$$\mathfrak{C}(b_{uv}) = \begin{cases} 0, & \text{if } b_{uv} < c_1; \\ \Delta, & \text{if } c_1 \leq b_{uv} \leq c_2; \\ 1, & \text{if } b_{uv} > c_2, \end{cases} \quad (3)$$

where c_1, c_2 – decision rule parameters. In this case, the estimate b_{uv} is calculated using the operator (2).

To date, several types of models have been developed and deeply studied in detail, which make it possible to distinguish the following well-known pattern recognition algorithms: 1) models based on the use of the separation principle [1-7]; 2) models based on the method of potential functions [4-6]; 3) statistical models [4, 5-8]; 4) models built on the basis of mathematical logic [1, 9, 10]; 5) models based on the calculation of estimates [1, 4, 11]. However, the analysis of these models shows that at present, models of recognition algorithms are mainly being developed, mainly focused on solving problems, where objects are described in the space of independent features (or the relationship between features is rather weak).

In practice, often the applied tasks of pattern recognition are given in the space of features of a large dimension. When solving such problems, the assumption of independence of signs is often not fulfilled [3, 12]. Consequently, the question of creating recognition algorithms that can be utilized to solve applied recognition problems for large dimensions of an attribute space and the presence of interconnectedness of properties remains insufficiently resolved.

The purpose of this work is to develop a model of modified recognition algorithms in conditions of high dimensionality of feature space. The model of recognition algorithms based on radial functions is considered as an initial model.

2. Statement of the problem

Given a set of admissible objects \mathbb{S} . Let the set \mathbb{S} consist of l subsets (classes) $\mathcal{C}_1, \dots, \mathcal{C}_j, \dots, \mathcal{C}_l$:

$$\mathbb{S} = \bigcup_{j=1}^l \mathcal{C}_j, \mathcal{C}_i \cap \mathcal{C}_j = \emptyset, i \neq j, i, j \in \{1, 2, \dots, l\}. \quad (4)$$

In this case, it is assumed that the partition (4) is not completely defined. However, there is some a priori (initial) information I_0 about classes $\mathcal{C}_1, \dots, \mathcal{C}_j, \dots, \mathcal{C}_l$. Usually I_0 is defined as classified objects. To determine I_0 from the set \mathbb{S} we select m objects: $\tilde{S}^m = \{S_1, \dots, S_u, \dots, S_m\}$. We introduce the following notation: $\tilde{\mathcal{C}}_j = \tilde{S}^m \cap \mathcal{C}_j$, $\tilde{\mathcal{D}}_j = \tilde{S}^m \setminus \tilde{\mathcal{C}}_j$. Then the initial information I_0 can be represented as a set of pairs consisting of S_u and $\tilde{\alpha}(S_u)$:

$$I_0 = \{ \langle S_1, \tilde{\alpha}(S_1) \rangle, \dots, \langle S_u, \tilde{\alpha}(S_u) \rangle, \dots, \langle S_m, \tilde{\alpha}(S_m) \rangle \}. \quad (5)$$

Here $\tilde{\alpha}(S_u)$ – is the information vector of the object S_u ($S_u \in \mathbb{S}$): $\tilde{\alpha}(S_u) = (\alpha_{u1}, \dots, \alpha_{uj}, \dots, \alpha_{u\ell})$, where α_{uj} – is the value of the predicate that has the following form:

$$P_j(S_u) = \begin{cases} 1, & \text{if } S_u \in \tilde{\mathcal{C}}_j; \\ 0, & \text{if } S_u \notin \tilde{\mathcal{C}}_j. \end{cases}$$

Let some sample be given consisting of q objects: $\tilde{S}^q = \{S'_1, \dots, S'_u, \dots, S'_q\}$ ($\tilde{S}^q \subset \mathbb{S}$). Each object S'_u from the sample \tilde{S}^q in the feature space X corresponds to a certain feature vector:

$$S'_u = (a'_{u1}, \dots, a'_{ui}, \dots, a'_{un}).$$

At the same time, the dimension n of the space of initial features is rather large. Under these conditions, most signs are interrelated, which makes it difficult to use many recognition algorithms [3]. The task is to construct such a recognizing operator \mathfrak{B} (see formula 2), which using the decision rule \mathfrak{C} (see formula 3) allows us to calculate the values of the predicate $P_j(S'_u)$ ($P_j(S'_u) = "S'_u \in \mathcal{C}_j", u = \overline{1, q}$) according to the initial information (5):

$$\mathfrak{B}(\tilde{S}^q) = \|\mathfrak{b}_{uv}\|_{q \times l}, \mathfrak{C}(= \|\mathfrak{b}_{uv}\|_{q \times l}) = \|\beta_{uv}\|_{q \times l}, \beta_{uv} \in \{0, 1, \Delta\}.$$

Here β_{ij} interpreted as follows [1]. In two cases, it is considered that using the algorithm \mathfrak{A} the value of the characteristic function on the admissible object S_i : 1) if $\beta_{ij} = 1$, then the object S_i belongs to the class \mathcal{C}_j ; 2) if $\beta_{ij} = 0$, then the object S_i is not in the class \mathcal{C}_j . In the case of $\beta_{ij} = \Delta$ it is considered that algorithm could not determine the value of the predicate $P(S'_i)$.

3. Proposed approach

To solve the formulated problem, an approach suggests that is a logical continuation of the work of Academician Yu.I. Zhuravleva and his students. Based on this approach, a model of recognition algorithms has been modified, based on identifying independent subsets of interrelated features and highlighting the preferred model for making intermediate decisions. As the initial model, recognizing operators based on radial functions are taken into consideration [12-14].

1. The selection of subsets of tightly coupled features. At this stage, a system of “independent” subsets of features is determined, the composition of which will depend on the parameter n' [12, 15]. By specifying various integer values for this parameter, we obtain various recognition algorithms.

2. Formation of a set of representative features. As a result of this stage, we obtain a reduced feature space, the dimension of which is much less than the original one. Further, the generated feature space is denoted by X' ($X' = (x'_1, \dots, x'_{n'})$) [15-17].

3. Definition of models of elementary transformations. At this stage, the construction of models of elementary transformations is carried out [1]. Suppose that on objects belonging to the class \mathcal{C}_j , several simple models are given in the space of representative features:

$$y_1 = f(x'_1, x'_2, x'_3), y_2 = f(x'_1, x'_2, x'_4), \dots, y_u = f(x'_{u_1}, x'_{u_2}, x'_{u_3}), \dots, y_n = f(x'_{n'-2}, x'_{n'-1}, x'_{n'}), \quad (6)$$

$$u_1, u_2, u_3 \in \{1, 2, \dots, n'\}, u_1 \neq u_2 \neq u_3, u \in \{1, 2, \dots, q\}, n = C_{n'}^3,$$

where f – a function from a given set of models of elementary transformations of \mathbb{F} .

As a given set of models of elementary transformations \mathbb{F} we consider linear models. Then the model of elementary transformations (6) is given in the form

$$y_u = c_{u_0} + c_{u_1}x'_{u_1} + c_{u_2}x'_{u_2} + c_{u_3}x'_{u_3},$$

where $c_{u_0}, c_{u_1}, c_{u_2}, c_{u_3}$ – are the parameters determined on the basis of the criterion of least squares.

At this stage, the parameters $\tilde{\zeta} = \{\bar{c}_1, \bar{c}_2, \dots, \bar{c}_u, \dots, \bar{c}_n\}$ for the class K_j ($j = \overline{1, l}$), where $\bar{c}_u = (c_{u_0}, c_{u_1}, c_{u_2}, c_{u_3})$.

4. Selection of subsets of strongly related models of elementary transformations. At this stage, a system of “independent” subsets of models of elementary transformations is determined. As a result of the implementation of this stage, the system of n' “independent” subsets of the strongly related attributes \mathcal{P}_q ($q = \overline{1, n'}$) is distinguished.

Depending on the method of specifying the measure of proximity between the subsets of tightly coupled models of elementary transformations (\mathcal{P}_p and \mathcal{P}_q) and the quality functional of the classification analysis, you can get various procedures for selecting independent sets of tightly coupled models of elementary transformations.

5. *Formation of a set of representative models of elementary transformations.* This stage, a set of representative models of elementary transformations is formed, each of which is a typical representative of a selected subset of tightly coupled models of elementary transformations. This procedure is similar to the procedure given in [12].

6. *Definition of elementary threshold decision rules.* We formulate elementary threshold decision rules. They characterize elementary decisive (discriminant) functions within the framework of the transformation model under consideration. Let an arbitrary admissible object ($S \in \mathcal{C}_j$), given in the space of representative features: $S = (a_1, \dots, a_i, \dots, a_{n'})$. To Then, on the basis of the elementary transformation model $y_u = f(x'_{u_1}, x'_{u_2}, x'_{u_3})$ can determine the elementary threshold decision rules \mathcal{G}_u ($u = 1, 2, \dots, n$) as:

$$\mathcal{G}_u(K_j, S) = \begin{cases} 0, & \text{if } f(a'_{u_1}, a'_{u_2}, a'_{u_3}) > \delta_u; \\ 1, & \text{else,} \end{cases} \quad (7)$$

where $a'_{u_1}, a'_{u_2}, a'_{u_3}$ – is the value of u_1, u_2 and u_3 representative feature of the object S ; δ_u – is a given threshold ($u = 1, \dots, C_n^3$).

7. *Selection of preferred models of elementary threshold decision rules.* As a result of this stage, the preferred models of elementary threshold decision rules are determined. The search for the preferred model of elementary threshold decision rules in the subset \mathfrak{P}_j is based on the assessment of the dominance of the considered models for objects that belong to the set \tilde{S}^m [14, 15]:

$$\mathfrak{D}_u = \sum_{j=1}^l \left(\left(|\tilde{K}_j| \sum_{S \in \tilde{\mathcal{D}}_j} \mathcal{G}_u(K_j, S) \right) / \left(|\tilde{\mathcal{D}}_j| \sum_{S \in \tilde{\mathcal{C}}_j} \mathcal{G}_u(K_j, S) \right) \right).$$

The larger the value of \mathfrak{D}_u , the more preferred is the u -th model of the elementary threshold decision rules. If several models get the same preference, then any one of them will be selected.

8. *Definition of the function of the distinction between class and object.* At this stage, the function sets the difference between the class and the object [12], calculated using the threshold function (7):

$$d(K_j, S) = \sum_{u=1}^{n'} \gamma_u \mathfrak{d}_i(\mathcal{C}_j, S),$$

where γ_u – is a parameter of the algorithm ($u = 1, \dots, n'$).

10. *Evaluation for a class based on the aggregate of preferred elementary threshold rules.* The estimation of the belonging of the object S to the class K_j ($j = \overline{1, l}$) is calculated as follows:

$$B(S) = (\mathfrak{G}_1(S), \dots, \mathfrak{G}_j(S), \dots, \mathfrak{G}_l(S)),$$

$$\mathfrak{G}_j(S) = 1 / (1 + \tau d(K_j, S)),$$

where τ – is an algorithm parameter.

9. *The decisive rule.* The decision rule is given in the form (3).

Thus, we have defined a class of modified recognition algorithms based on radial functions. Any algorithm \mathfrak{A} from this model is completely determined by specifying a set of parameters π . The set of all recognition algorithms from the proposed model is denoted by $\mathfrak{A}(\pi, S)$. The search for the best algorithm is carried out in the parameter space π [11].

4. Experiment and Results

An experimental study of the performance of the proposed model of recognition algorithms is carried out on the example of solving a model problem.

The source data of recognizable objects for a model task is generated in the space of dependent attributes. The number of classes in this experiment is two. The size of the initial sample is 400 implementations (200 implementations for objects of each class). The number of features in the model example is 400. The number of subsets of tightly coupled features is 6.

The following models of recognition algorithms were chosen: the classical model of recognition algorithms of the type of potential functions (\mathfrak{A}_1), the model of recognition algorithms based on the calculation of estimates (\mathfrak{A}_2), and the model (\mathfrak{A}_3), proposed in this work. Comparative analysis of the listed models of recognition algorithms for solving the considered problem was carried out according to the following criteria: 1) recognition accuracy of test sample objects; 2) time spent on training; 3) the time spent on recognition of objects of the control sample.

To calculate the specified criteria when solving the considered problem, in order to exclude successful (or unsuccessful) splitting of the initial sample into two equal parts, the sliding control method [18] is used.

The recognition accuracy in the learning process (see formula 3) for \mathfrak{A}_1 is equal to 94.8%, for \mathfrak{A}_2 – 97.6%, and finally, for \mathfrak{A}_3 – 99.1%. The results of solving the problem under consideration with the use of \mathfrak{A}_1 , \mathfrak{A}_2 and \mathfrak{A}_3 in the process of control are given in Table 1.

Table

The results of solving the problem using various recognition algorithms

Table 1. The results of user authentication by keyboard handwriting.

Recognizing operators	Time (insec.)		Accuracy of recognition (in percentage form)
	Learning	Recognition	
1	4.141	0,014	80,5
2	6,127	0,012	82,7
3	9,041	0,002	93,4

Comparison of these results shows (see table) that the proposed model of recognition algorithms \mathfrak{A}_3 has improved the accuracy of recognition of objects described in the space of interrelated features (more than 10% higher than with \mathfrak{A}_1 and \mathfrak{A}_2). This is because the models \mathfrak{A}_1 and \mathfrak{A}_2 do not use the preferred elementary threshold rules. However, for model \mathfrak{A}_3 there is a certain increase in the learning time due to the implementation of an additional procedure for the formation of independent subsets of interrelated features and preferred threshold rules.

Conclusion

A new approach has been proposed for building a model of recognition algorithms, and a modified model of recognition algorithms based on the evaluation of estimates has been constructed on the basis of this approach. The proposed model of algorithms is based on the calculation of intermediate estimates in the subspaces of attributes of various sizes. This model is focused on solving the problem of recognizing objects defined in the space of features of a large dimension. A distinctive feature of the proposed model of recognition algorithms is the use of the aggregation method in the construction of linear decision functions.

The results of solving a model problem showed that the proposed model of recognition algorithms improves accuracy and significantly reduces the number of computational operations by recognizing an unknown object specified in the space of interrelated signs. It can be used in the compilation of various software systems focused on solving applied recognition problems.

References

- [1] Zhuravlev Yu I 1998 *Selected scientific works* (Moscow: Master)
- [2] Kovalevsky V A 1980 *Image Pattern Recognition* (New York: Springer)
- [3] Fazilov Sh Kh, Mirzaev O N and Radjabov S S 2015 The current state of pattern recognition problems *Problems of computational and applied mathematics* **2** 99-112
- [4] Zhuravlev Yu I, Ryazanov V V and Senko O V 2006 *Recognition. Mathematical methods. Software system. Practical applications* (Moscow: Fazis)
- [5] Duda R O and Hart P E, and Stork D G 2001 *Pattern Classification* (New York: John Wiley)
- [6] Tou J and Gonzalez R 1978 *Pattern recognition principles* (Moscow: Mir)
- [7] Theodoridis S and Koutroumbas K 2009 *Pattern Recognition: Theory and Applications* (New York: Academic Press)
- [8] Fomin Ya A 2012 *Pattern recognition: theory and application* (Moscow: Fazis)
- [9] Lbov G S and Berikov V B 2005 *Stability of decision functions in problems of pattern recognition and analysis of heterogeneous information* (Novosibirsk: Publishing house of the Institute of Mathematics)
- [10] Kudryavtsev V B, Andreev A E and Gasanov E E 2007 *Theory of test recognition* (Moscow: Fizmatlit)
- [11] Zhuravlev Yu I, Kamilov M M and Tulyaganov Sh E 1974 *Algorithms for calculating estimates and their application* (Tashkent: Fan)
- [12] Mirzaev N M and Saliev E A 2018 Recognition Algorithms Based on Radial Functions *Proceedings of the 3rd Russian-Pacific Conference on Computer Technology and Applications* (Vladivostok: FEFU) doi: 10.1109/RPC.2018.8482213
- [13] Buhmann M D 2009 *Radial Basis Functions: Theory and Implementations* (Cambridge: Cambridge University Press)
- [14] Russell J and Cohn R 2012. *Radial Basis Function* (New York: Book on Demand)
- [15] Fazilov Sh Kh, Mirzaev N M, and Mirzaeva G R 2019 Modified recognition algorithms based on the construction of models of elementary transformations *Procedia Computer Science* **150** 671-678 doi: 10.1016/j.procs.2019.02.037
- [16] Mirzaev N M 2010 About one model of image recognition *Proceedings of the First Russia and Pacific Conference on Computer Technology and Applications* 394-398
- [17] Fazilov Sh Kh, Mirzaev N M, Radjabov S S and Mirzaev O N 2018 Determining of Parameters in the Construction of Recognition Operators in Conditions of Features Correlations *Proceedings of the 7th International Conference on Optimization Problems and Their Applications* 118-133 url: <http://ceur-ws.org/Vol-2098/paper10.pdf>
- [18] Braga-Neto U M and Dougherty E R 2016 *Error Estimation for Pattern Recognition* (New York: Springer)