УДК 004.93

PATTERN RECOGNITION IMPROVEMENT BY QUASI-TRAINING SAMPLES*

УСОВЕРШЕНСТВОВАННОЕ РАСПОЗНАВАНИЕ ОБРАЗОВ НА ОСНОВЕ КВАЗИЭТАЛОННЫХ ОБРАЗОВ*

A.S. GOLUBEV, M.YU. ZVYAGIN, L.A. SYEMIN, A.V. LOKHANOV
А.С. ГОЛУБЕВ, М.Ю. ЗВЯГИН, Л.А. СЕМИН, А.В. ЛОХАНОВ

(Vladimir State University named after Alexander and Nikolay Stoletovs)
(Владимирский государственный университет
имени Александра Григорьевича и Николая Григорьевича Столетовых)
E-mail: oid@vlsu.ru; andrey.golubev@vlsu.ru; muz1953@yandex.ru; semin.lev@gmail.com

The work addresses the problem of pattern recognition for the case of one-training sample per class. An original approach based on auxiliary set of "quasi"-training samples is discussed. We provide a formal mathematical definition along with practical results from experimental recognition system.

The research has been done in the framework of the state job VlSU 2014 in the field of scientific work.

* Научная статья подготовлена в рамках государственного задания ВлГУ 2014 г. на выполнение государственных работ в сфере научной деятельности.

В работе рассматривается задача распознавания образов в условиях единичного эталона и метод ее решения на основе вспомогательного множества «квазиэталонных» образов. Приводятся математическое описание и практические результаты, полученные с помощью экспериментальной системы распознавания.

Keywords: pattern recognition, single training sample, quasi sample, estimates calculation algorithm.

Ключевые слова: распознавание образов, единичный эталон, квазиэталон, алгоритм вычисления оценок.

The problem of pattern recognition is extremely useful in a wide area of applications, for example industrial automation. One of yet unresolved issues is recognition under a small number of training samples per class, with a single sample per class as an extreme case [1, p.161]. When the problem is formulated like that, it is impossible to use statistic models, methods of cluster analysis, leaning-based methods, i.e. the most effective instruments of the pattern recognition theory. In our research [3…6], which directly concerns the condition of a single training sample per class we decided to concentrate our efforts on the idea of Estimates Computation Algorithms, also known as ECA [2, p.67].

Object classification during recognition by a single training sample can be achieved within the limits of two approaches. "Classical" approach implies, at first, making the function $F: X \to D$, which maps the global set of object samples (denoted as X) to the metric set of features (denoted as D), and then introduction of "threshold" value ε. Let $Y \subset X$ be a set of training samples. By definition, any training sample corresponds to unique class of objects (single sample per class). Hence, each known class can be represented by its training sample $y \in Y$. The test sample being recognized $x \in X$ is classified (recognized) as sample from class $y' \in Y$ if $F(y')$ is the nearest element to F(x) in terms of set D metrics. Meanwhile the distance between $F(y')$ and $F(x)$ must be less than some fixed threshold ε, otherwise the test is considered to be not recognized. Thus the necessity of introducing a threshold is conditioned by the demand for revealing the objects which do not have training samples (this demand differs the problem of recognition from the problem of classification). The main difficulty in the classical approach is to choose the value of ε, as this value can be specified only empirically. Besides, it is necessary to mention that this approach envisages strict requirements to function F in the sense that the samples of the same object must be matched to the "nearest" elements of set, i.e. it is necessary to follow the hypothesis of compactness of "feature" subset mapped from "sample" subset of each class.

The main idea of another approach is in applying the principle of a multiple-factor decision making that characterizes ECA method. This approach has been demonstrated in [3] when used for the problem of the pattern recognition by a single training sample. The essence of this approach is that it suggests to use the ensemble of *n* functions $\{f_i : X \to D_i\},\ i = 1,\ldots,n$ instead of one global function F. Thus, each function f_i specifies partial order relation on the set X, and for any test $x \in X$ it is possible to fetch n sequences $\{y_1^i,\ldots,y_m^i\},\ y_m^i \in Y,\ i = \overline{1,n}$; each of them being composed of *m* classes which are the nearest (in the appropriate metrics) to x. The decision about the membership of x in this or that class is made after analysis of fetched sequences, e.g. by a simple election. The value of m is not fixed. It is determined by analyzing the decision rule and can vary in the range of 1 to $|Y|$. The success of this method is based on the assumption about the independence of the result calculation of each ensemble function $\{f_i\}$.

The advantage of this approach is that each separate function of the ensemble can be a relatively weak classifier, i.e. it does not need to provide a complete division of classes in the appropriate feature space being mapped D_i. Besides, the decision rule analyzing the sequences $\left\{\{y_j\}^i\right\}$, does not necessarily need an artificially specified threshold. For example, the rule can make a decision about negative classification if two or more classes in the sequences have got the best estimation. At the same time if the decision rule is in some way based on the application of a threshold, its significance depends on the external factors to a lesser degree as compared to the classical approach. In practice it means a simpler adaptation of the recognition system to some concrete problem and conditions of application.

The problem of the described approach to classification occurs in the case when the number of training samples is comparable with n (i.e. one must not make the size of the ensemble too large, and, which is more important, one must not identify small sets of objects as, for example, when recognizing alphabet symbols.) Under these conditions elements in $\left\{\{y_j\}^i\right\}$ are most likely to be matched even if the necessary training sample is absent in Y. In the latter case the decision rule will practically always give a wrong result. This statement is a good demonstration of the extreme case when one class is identified, i.e. set Y consists of a single element y'. Since all sequences $\{y_j\}^i$ will be composed of repetitions y' for any test, the method will always lead to the same result being positive regardless whether the test is related to the object.

In the canonical formulation of ECA method this problem is excluded due to the fact that each of the functions $\{f_i\}$ can definitely determine if the test refers to one of the classes being recognized. But for the condition of a single training sample per each class it means the introduction of threshold values ε_i for all ensemble functions. Obviously it nullifies all the advantages of the multifactor approach over the "classical" one.

Practice of applying rerecognition systems leads us to the conclusion that having the above mentioned scheme of comparison it is reasonable to use special objects (named as *quasi training samples* or simply *quasi-samples*) the meaning and role of which will be explained further. The nature of such objects (as will be shown in the next chapter) can be various and, to a great extent, their good choice predetermines successful recognition.

So, to be more specific, let us introduce the subset of quasi-samples W on set X. By its construction the subset W must have the following properties. Firstly, the set of quasi-samples classes $C(W)$ and the set of training samples classes $C(Y)$ must not intersect at all: $C(W)\cap C(Y)=\varnothing$. Secondly, $|C(W)|>>|C(Y)|$. In other words W is a (quite big) set of samples which do not belong to objects under current recognition.

Let's postpone the question about how it is possible to obtain set W in practical tasks and see how to solve the above problem of building the decision rule in ECA by means of the new concept**.** It should be mentioned that several variants of solution can be suggested. We are going to demonstrate some of them in the order of increasing their complexity.

Variant 1. For test $x \in X$ only those elements are included into each sequence $\{y_j\}^i$ generated by function f_i which are nearer in the metrics of set D_i than any other element of set W. Thus, quasi-samples implicitly specify threshold values ε_i for every x in a natural way:

$$\varepsilon_i = \min_{w \in W} d_i(x, w),$$

where $d_i(x,w)$ – distance between $f_i(x)$ and $f_i(w)$.

The main idea of this construction is that threshold values are not fixed artificially but get an intuitively understandable justification. It is of no small importance that due to the modification of the sets of quasi- samples W there appears a possibility to adapt threshold values to the concrete environment**.**

<u>Variant 1.1.</u> The previous method can be made more flexible if take a mean distance instead of a minimum one and add a control coefficient α:

$$\varepsilon_i = \alpha \cdot \underset{w \in W}{\operatorname{avg}} d_i (x, w), \ \alpha \in R^+.$$

<u>Variant 2</u>. For every sample $x \in X$ (including the training samples) and every function f_i we'll build the sequence W_i^x composed of elements of set W arranged in the ascending order of their distances to x. As a result every sample will be mapped to the list $\{W_1^x, \ldots, W_n^x\}$ of such sequences. We call this list quasi-coordinates of x. We are going to determine the membership of test x in the class y by comparing their quasi-coordinates. A great variety of variants is also possible here. For example, we leave one element (the first one) in each i-th quasi-coordinate. We receive two vectors: $\{w_1^x, \ldots, w_n^x\}$ and $\{w_1^y, \ldots, w_n^y\}$ for x and y respectively. Subsequently, a binary vector $b^{xy} = \{b_i^{xy}\}$ is composed according to the following rule: b_i^{xy} is equal to 1 if $w_i^x = w_i^y$ and is equal to 0 otherwise. Estimation of y for test x is considered to be the sum of b^{xy} vector elements. The method can be generalized by selecting from every i-th quasi-coordinate not one element but m elements ($1 \le m \le |W|$), and by calculating the estimation as a sum of elements of binary matrix $\{b^{xy}\}_{m \times n}$.

Our research shows that such estimations enable to successfully solve the problem of recognition on a small set Y. To illustrate this statement we are giving the results of a simple experiment made by means of the face image recognition algorithm [7]. During the experiment we carried out the identification of 30 people ($|Y| = 30$) with the use of variable set of quasi-samples ($|W| = 100, 200, \cdots, 600$). Growth of type 1 error turned out to be relatively low (from 9,20% to 13,67%, i.e. less than 5%) though the number of classes increased from 30 to 630. At the same time there was the reduction of type 2 error by more than a factor of ten (from 37% to less than 3%).

Now we are coming back to the question about the way of receiving set W in practical tasks. It has been said already that the set of quasi-samples consists of the classes which are known to be unreachable during ordinary recognition process. It is obvious that many problems of recognition imply "the natural interpretation" of such set. A popular problem of recognizing a person by the image of his face envisages, as a rule, the recognition of some fixed target group (wanted, employers of some particular institution) made up of the unfixed set of people being under control. In most cases it is easy to form an extra base of faces which must be safe from intersecting with the target base. The sources of such pictures can be open data bases in the Internet, face photographs of some other target group, pictures taken in a hive separated geographically from the zone of control, etc.

At the same time there are problems where "natural" interpretation of a quasi-etalon set is either not obvious or difficult for technical realization. Such situations are of great interest as they imply a synthetic generation of quasi-samples. We are considering the following variants of such generation.

1) The synthesis of samples on the basis of a summarized parametric model of the objects being recognized. The method can be used only for some tasks when it is possible to describe all classes by some mathematical model with a limited number of parameters. Varying the values of these parameters we can generate the necessary number of objects and then use their images as quasi-samples. It is obviously difficult to create the exact model but in a number of areas of machine vision, e.g. in technological processes of the quality control, this problem can be solved.

2) Random generation. In this case not the elements $w \in W$ but their mapped fea-

tures $f_i(w)$ in each of the spaces D_i are synthesized. So far as each of these features is, as a rule, a real vector, its components can be specified in a random manner by means of this or that distribution. This method is limited by the demand for a small dimension D_i. But later this enables to use optimization methods for a better division of the classes being recognized.

3) The synthesis of set W (or its mappings $f_i(W)$) on the basis of the set of available training samples. Actually, this method is the application of some parametric operator $g(\lambda, y)$ to the training samples from Y. Variation of parameter λ enables to synthesize several quasi-samples from single training sample. For example, if Y is a set of digital images, it is convenient to consider some sequence of the image distortions (such as linear or centered tensions, distancing or replacing the parts, noise introduction, etc.) as operator g. A similar operator can be applied not to the images $\{y \in Y\}$ themselves, but to their mappings $f_i(y)$.

Thus, the sets of quasi-samples can be realized practically in any task. Conceptually this approach is close to the ideas of training set expansion by synthesizing extra images of identified classes; this being the attempt to remove the restrictions of single training sample per class. In the case with quasi-samples this restriction holds valid and due to the natural realization the synthesis is not necessary at all. As it was shown, such set enables to cardinally solve the problem of type 2 error and allow us to remain under the concept of the Estimates Computation Algorithms without using threshold values.

From our point of view the idea of applying the ensemble of “weak classifiers” for solving the problem of recognition is very promising not only under the conditions of a single training sample. It is no mere chance that this principle is the basis of many algorithms which are successfully used for solving many adjacent problems (e.g. detection of the object on the image). We are planning to continue our research in some areas having the following purposes: to improve the procedure of accounting quasi-sample images while calculating the estimations; to involve as much data available for identification as possible; to extend the number of base algorithms for calculating estimations.

REFERENCES

1. *Anisimov B.V., Kurganov V.D., Zlobin V.K.* Raspoznavanie i cifrovaja obrabotka izobrazhenij: Ucheb. posobie dlja studentov vuzov. – M.: Vyssh. shk., 1983.

2. *Zhuravlev Ju.I., Rjazanov V.V., Sen'ko O.V.* Raspoznavanie. Matematicheskie metody. Programmnaja sistema. Prakticheskie primenenija. – M.: Fazis, 2005.

3. *Rozhkov M.M., Prokoshev V.G., Shamin P.Ju.* Problema avtomaticheskogo raspoznavanija lic s odnim jetalonnym izobrazheniem // Nauchno-tehnicheskie vedomosti Sankt-Peterburgskogo gosudarstvennogo politehnicheskogo universiteta. Serija “Informatika. Telekommunikacii. Upravlenie”. – SPb, 2010, №5.

4. *Golubev A.S., Zvjagin M.Ju., Kvasov D.S., Kokorin I.G., Zinov'ev I.I., Shamin P.Ju.* Apparatno-programmnyj kompleks avtomaticheskoj registracii i biometricheskoj identifikacii ljudej // V kn.: Materialy XVII Vserossijskoj nauchno-metodicheskoj konferencii “Telematika 2010”. – SPb., 2010.

5. *Rozhkov M.M., V.G., Rozhkov M.M., Shamin P.Ju., Golubev A.S.* Postroenie podprostranstv atributov na baze odnogo jetalona dlja obespechenija ustojchivosti raboty v perspektivnyh sistemah avtomaticheskogo raspoznavanija lic // Prikladnaja informatika. – M., 2011, №2 (32).

6. *Rozhkov M.M.* Ispol'zovanie teksturnyh kart Lavsa i diskretnogo kosinusnogo preobrazovanija v zadache raspoznavanija lic // Prikladnaja informatika. – M., 2011, №1 (31).

7. *Golubev A.S., Zvjagin M.Ju., Prokoshev V.G., Rozhkov M.M.* Osobennosti raspoznavanija metodom blizhajshego jelementa v algoritmah vychislenija ocenok // Prikladnaja informatika. – M., 2013, №1.

Рекомендована кафедрой физики и прикладной математики. Поступила 09.07.15.