Hybrid system for recognition of handwritten symbols on the base of structural methods and neural networks

R.Kh.Sadykhov, O.G.Malenko, A.N.Klimovich, M.L.Selinger

Belarusian State University of Informatics and Radioelectronics, Computer System Department. Institute of Engineering Cybernetics of the National Academy of Sciences of Belarus.
6, Brovka str., 220600, Minsk, Belarus
E-mail: rsadykhov@gw.bsuir.unibel.by

Abstract

The problem of handwritten symbol recognition has been investigated. An algorithm of approximation and breaks elimination has been developed. This approach allows to simplify the description of symbols and remove available errors. A method of structural recognition based on the description of a structure for handwritten symbols with the help of primitive sequences, robust to geometrical distortions has been used. A structural classifier has been used in a combination with the specified method of the description. The technique for the analysis and optimization of a choice of feature extraction algorithm and its parameters based on an estimation of clasterization quality of training sample with the help of self-organizing neural networks has been implemented. Computation algorithm of Legendre moments is presented. A new method for training RBF-neural network is represented. Classification results for binary images (handwritten Arabic numerals) are presented. On the base of classification results the recommendations for choice of maximal order Legendre moments and various classifiers are given.

1. Introduction

The system consists from two parts: structural recognition subsystem, and neural recognition subsystem [6, 7]. Methods of structural recognition have more potential when it is necessary to achieve high recognition quality at significant modifications and distortions in recognized objects in comparison with an ideal one. We suggest the structure of recognition system on the basis of comparison with the standard. The structural methods, permitting to select and recognize structure, are characterized by a stability to the distortions peculiar to handwritten symbols: variations of a size, proportions, an angle of declination, thickness of a line. The possibility of training allows to adjust a system to various types of symbols (numerals, letters of various languages, other symbols).

The structural subsystem operates according to the following stages:

- filtration and thinning;
- approximation of the symbol on the base of graph representation;
- elimination of the break-downs;
- exposition of a symbol structure via primitive sequences;
- recognition of an image by the search close ideal description via the base of standards.

Neural subsystem performs feature extraction task from the image on the basis of Legendre moments and classifies them with the help of RBF network. So, real output of the system is based on the best result of two subsystems. Both subsystems has been trained on the same data set of grayscale images.

2. Break-down elimination

The basic direction of the feature extraction (compression) in a structural recognition of handwritten numerals is selection of symbol structure with the help of thinning [1], graph extraction and approximation. The serious problem is the break-downs that takes place because of the errors of spelling and scanning, and causes the large modifications in description of a symbol structure. The break-downs of handwritten symbols have been investigated, and the fast heuristic algorithm of their elimination has been developed. The break-downs can be divided into three types. Figure 1 shows them.

It should be noted that the characteristic case is the "point-line" break-down. It is appropriate to such situation when square of distance between vertex of graph and the nearest point of some edge, not incident to vertex and not connected with it, is less than threshold *P*. The threshold is usually defined by percentages from geometrical size of a symbol that makes algorithm insensitive to symbol magnitude.

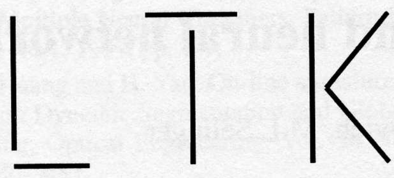

Fig. 1. Classification of break-downs

Let the edge A_1A_2 (figure 2) be given and there is a point B, not lying on it. It is required to find coordinates (x, y) of point C, belonging to the edge and the closest to B.

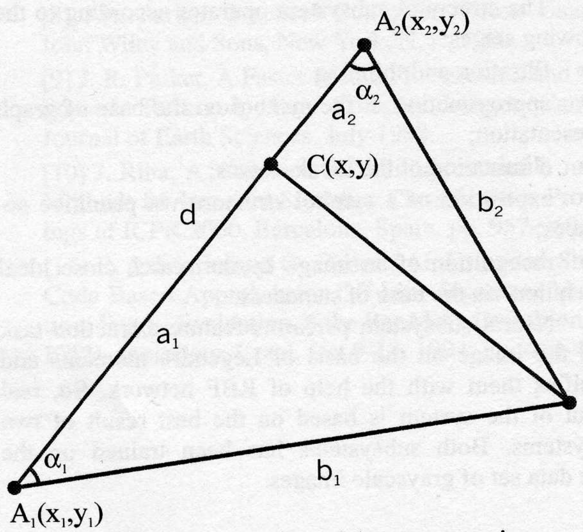

Fig. 2. Determination of the nearest point

We have obtained the following relation:

$$C(x,y)=\begin{cases} A_1(x_1,y_1), & if \quad d^2-b_2^2+b_1^2\le 0 \\ A_2(x_2,y_2), & if \quad d^2-b_1^2+b_2^2\le 0 \\ \left(x=\dfrac{(z-1)x_2+x_1}{z}, y=\dfrac{(z-1)y_2+y_1}{z}\right), & if \quad \begin{matrix} d^2-b_2^2+b_1^2>0 \\ d^2-b_1^2+b_2^2>0 \end{matrix} \end{cases},$$

where $z=\dfrac{2d^2}{d^2-b_1^2+b_2^2}$.

Thus, the nearest point is defined and if the distance square BC is less than threshold P, then we have break-down, which is easily eliminated by introduction of edge BC. It is necessary to note, that as a result of break-down elimination in the graph the edges that did not expose for approximation can appear. That is the approximation algorithm and break-downs elimination are preferable for applying in a complex.

3. Primitive sequences

The method for simplified description and classification, based on primitive sequences [2-4] (figure 3), representing a curve of the defined direction has been developed. For each primitive sequence it is possible to determine the numerical value based on magnitude of its turn. Thus turn is estimated on a difference of bisector angles for each angle of a curve. Such approach is more robust to geometrical distortions, than a simple difference of primitive angles.

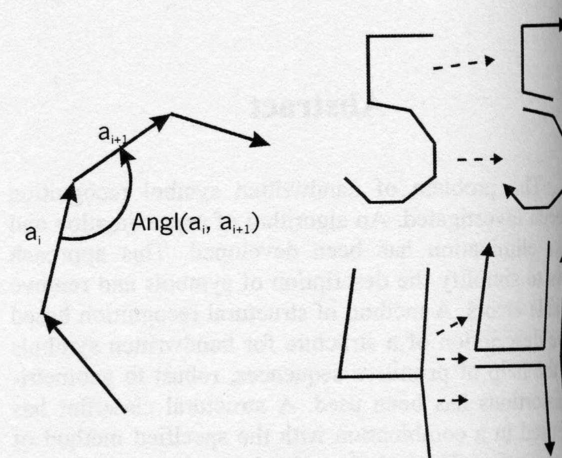

Fig. 3. Primitive sequences

For simplification of calculations the angle of bisector is coded as j according to its angle range, when $j=\overline{0,3}$. Then, for a primitive sequence:

$$a_1 \xrightarrow{j_1} a_2 \xrightarrow{j_2} a_3 \xrightarrow{j_3} \ldots \xrightarrow{j_{n-2}} a_{n-1} \xrightarrow{j_{n-1}} a_n,$$

where j_i is value of the appropriate bisector, and characteristic of the sequence, defining turn, is determined on the base of relation:

$$P = 2 + \sum_{i=1}^{n-2}\{(j_{i+1}-j_i)(mod\,4)\}$$

The relation allows to estimate a sequence correctly even in case when an angle difference between first and last primitive is more than 2π. If the sequence consists only of two primitives, then $P=2$; if from only one we count, that $P=1$. Number P is the basic value of a sequence which is robust for both turn and declination. So, in some cases for a successful recognition the information on declination is very important. Therefore an index of declination value of the first bisector j_1 must be included to describe the sequence. Thus, the pair (P, j_1) may be accepted as the description of one primitive sequence.

4. Training

The description based on these values is used at the stage of recognition. The ordered list of description of all primitive sequences for handwritten symbols has been applied. Basing on such approach both recognized pattern and each class in the base of standards are coded. The recognition is reduced to searching the closest class. A recognition sequence is the follows:

- in the base of standards of the classifier similar standards are searched with the help of fast function of distance.
- more detailed comparison is made to find standards, with the use of the information about mutual positioning of primitive sequences and methods of graph comparison.
- if the recognized image does not correspond to any standard it is rejected as erroneous (operating mode) or its compressed description is added into the base of standards (training mode).

Let us consider an example of distance function. Let $(P^{r1}, j_1^{r1}), (P^{r2}, j_1^{r2}), ..., (P^{rS}, j_1^{rS})$ be the list of descriptions of primitive sequences for the recognized class, and $(P^{i1}, j_1^{i1}), (P^{i2}, j_1^{i2}), ..., (P^{iM}, j_1^{iM})$ be the list appropriate to a class standard i. At each comparison there is such sequence $(P^{r1}, j_1^{r1}), (P^{r2}, j_1^{r2}), ..., (P^{rk}, j_1^{rk})$, that the value:

$$R_i = \sum_{k=1}^{S} [A(P^{ik} - P^{rk}) + B\{(j_1^{ik} - j_1^{rk})(mod\, 4)\}],$$

should be minimum for a class i, where A and B are adjusted factors. If the given function of distance is less than threshold C then detailed comparison is implemented.

Values A, B allow to set a degree of influence of each performance on the result of the recognition and determine influence P and j_1, and value C influences on a performance of a system functioning and quality of recognition.

Training of the system is made by adding into a database of standards a description of those symbols which recognition according to current base was not successful. The quality of a recognition system strongly depends on training sample choice.

According to the sequentially adding standards there is a redundancy of that the added standard may cover with the function of distance a part of already existing standards. We have used a method of minimization of training sample for the structural classifier on the basis of comparison with the standard with the help of distance function. This method is specially oriented on work in the structural classifier with distance function where classical methods of cluster analysis are inapplicable. The basis of this method is the concept of a covering by one ideal description another which can be rejected without decreasing of a recognition quality if the given descriptions are corresponding to one class. Application of a method has allowed to reduce volume of the base of ideal descriptions (standards) to 25 %, without essential loss of quality (Table 1).

Table 1. Structural subsystem result comparison

Pattern type	Number of classes	Base number of standards	after minimization	Test sample size	Correct	Time, mc
Stylized symbol	36	360	360	3500	3320 (94,8%)	19
Stylized symbol	36	360	275 (76,4%)	3500	3309 (94,5%)	18
Any symbol	36	492	492	5000	4389 (87,8%)	23
Any symbol	36	492	370 (75,2%)	5000	4361 (87,2%)	20
Digit	10	64	64	700	691 (98,7%)	10
Digit	10	64	52 (81,2%)	700	690 (98,6%)	9

5. Legendre moments

In feature extraction task considerable attention for methods that use moment functions is given. Moment invariant properties are investigated in [12,14]. There are invariant on shifts, scaling and rotating of source object. During research time various types of moment functions were introduced, and fast computation algorithms for different types of moments were created. This part describes Legendre moments using for handwritten character informative feature extraction.

Two-dimensional Legendre moments for image intensive function $f(x,y)$ are defined as [6, 8]:

$$L_{kl} = \frac{(2k+1)(2l+1)}{4} \int_{-1}^{1}\int_{-1}^{1} P_k(x)P_l(y)f(x,y)dxdy,$$

where $f(x,y)$ – picture element with coordinates (x,y);

$P_0(x)=1;\ P_1(x)=x;\ P_k(x)=[(2k-1)xP_{k-1}(x) - (k-1)P_{k2}(x)]/k$ - Legendre polynomial by power k, $k>1$;

$l\geq 0$ и $k\geq 0$ defines the order of moments.

Since definition area of Legendre polynomials is $-1\leq x\leq 1$, then definition area of two-dimensional Legendre moments is unit square, so a rectangle image of $N\times M$ pixels with intensity function $f(i,j)$, $1\leq i\leq N$, $1\leq j\leq M$ will have to be scaled the region $1\leq x,y\leq 1$, and image center of gravity must be located in the coordinate system origin. For this:

- source image center of gravity (i_c, j_c) is computed;

- distance D from the center of gravity to the farthest from it point of image is determined according to equation

$$\forall(i,j): D = \max\{|i_c - i|, |j_c - j|\};$$

- scaling of image is performed according to

$$(x,y) = \left(\frac{j - j_c}{D}, \frac{i_c - i}{D}\right).$$

Legendre moments to maximum order *MAX_ORDER* can be computed by pseudo-code:

```
for k:=0 to MAX_ORDER
  for l:=0 to k
    L(k-l,l):=0
    for i:=1 to N
      for j:=1 to M
        x:=(j - jc)/D
        y:=(ic – i)/D
        L(k-l,l) := L(k-l,l) + Pk-l(x)*Pl(y)*f(i,j)
      end
    end
    L(k-l,l) := L(k-l,l)*(2k-2l+1)*(2l+1)/(N-1)/(M-1)
  end
end
```

Legendre polynomials $P_k(x)$ forms full orthogonal basis inside unit circle, so source image may be reconstructed from the finite number of Legendre moments as follows:

$$f(x,y) \cong \sum_k \sum_l L_{kl} P_k(x) P_l(y).$$

Figure 4 shows source halftone image and reconstructed images using various number of Legendre moments.

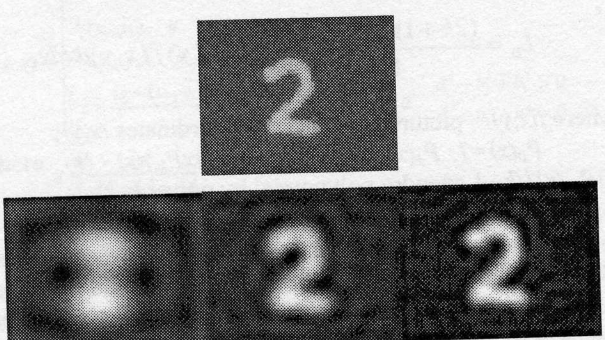

Fig.4. Source and reconstructed images using Legendre moments with maximum order 20, 40 and 60

6. Classifier on the basis of the RBF-neural network

In the tasks of classification the large attention is given to construction of classifiers on the basis of neural networks [8, 11, 15]. Radial basis function (RBF) neural network is two-layer neural network offered by Moody and Darken in 1989 [13] (figure 5).

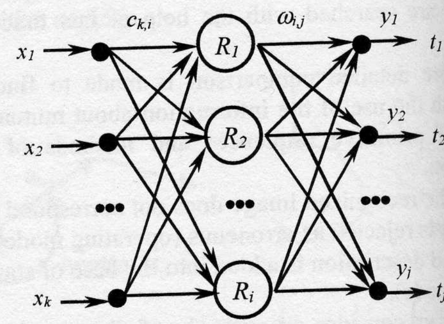

Fig. 5. The RBF- neural network architecture

The RBF-networks represent multilayer neural networks with RBF-neuron layer, which activation function correspond to the radial basic function [9]

$$R(D_i) = \exp\left(-\frac{D_i^2}{2\cdot\sigma_i^2}\right),$$

where $D_i = |\vec{x} - \vec{c}_i|$ is distance between an entry pattern $\vec{x}$ and i–th cluster of the radial basic function $\vec{c}_i$, σ - width of a cluster.

The RBF-neuron weight coefficients are associated with cluster of the radial basis function. Thus, the output data of RBF-layer represents a vector of closeness measures of entry pattern to all RBF-clusters.

The subsequent layers of such networks usually evaluate a linear combination of these functions.

$$y_j = \sum_i w_{ij} R_i$$

Key aspect for the RBF-network training is the possibility of layer-by-layer training, that results in two-phase training algorithm: a RBF-layer training and perceptron layer training.

RBF-layer training.

The RBF-layer training consists of two tasks: definition of necessary amount of RBF-neurons, and their weight coefficients setting. The training is produced by the following rules:

If in RBF-layer there are no such neurons that $D_i < \sigma$ or amount of neurons is equal to 0, then it is necessary to add a new neuron initializing its weights by training pattern vector value.

Else modification i-th neuron weights for $\min_{\forall i}(D_i)$ is performed in accordance with

$$\vec{c}_i(t+1) = \vec{c}_i(t) + \frac{1}{t+1}\cdot[\vec{x}(t+1) - \vec{c}_i(t)],$$

where $\vec{x}, \vec{c}_i$ are training vector and cluster vector accordingly, t – number of the training vectors came into i-th cluster.

Perceptron layer training

The perceptron layer training is made by a gradient descent method with the purpose of the error function minimization in weight coefficient space w_{ij}:

$$Err = \frac{1}{2}\cdot\sum_n\sum_j\left(\sum_i w_{i,j}\cdot R_i^n - t_j^n\right)^2 \to 0,$$

where n=*[1..N]* is amount of learning images, t_j^n - target value of an *j-th* output for a training pattern *n*.

A part of binary image database used in experiments as training sample is presented in figure 6.

For classification the RBF neural network classifier and minimal distance classifier [10] have been used. The database for experiments has been consisted of 4000 images on 400 each class of numerals. Classification results for different maximum order of Legendre moments are presented in Table 2.

Table 2. Classification results comparison

MAX_ORDER	Recognition percent, %	
	Minimal distance classifier	Neural network
5	76,4	92
10	90,6	98,2
15	91,6	98,7
20	88,4	98,7
25	85,2	98,5

The competitive neural network and Kohonen network have been applied to optimize the parameters of recognition system [5]. The base rule of training such neural networks is selforganizing by increase weights of a neuron-winner. The neuron-winner is determined with the help of the minimum:

$$d_i = |x - w^i| f_i,$$

where $|x-w^i|$ - Euclidean distance between input and weight vector, and f_i – frequency of neuron wins.

Clusterization results of training sample by competitive neural network are compared with training sample splitting into classes. Comparisons are made at various neural network configurations (number of inputs and outputs). On their base next choices may be made:

1. The choice of the best algorithm of extraction of informative features from the several perspective ones.
2. The estimation of a possibility of informative features number reduction or class quantity increase.
3. Check and filtration of training sample, removal of defective patterns.

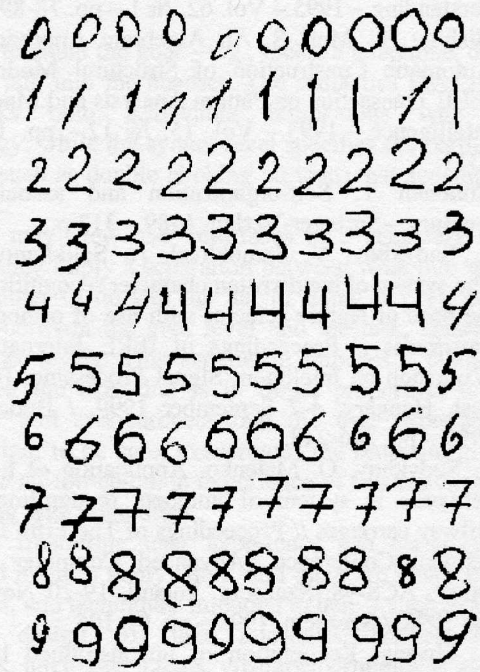

Fig 6. Part of training sample

Conclusion

The hybrid system for recognition of handwritten characters with the use of structural methods and neural networks has been developed. The application of Legendre polynomials and structural technologies to extract the informative features is shown. Furthermore these methods allow both to reduce space of features and increase performance of calculations. The outcomes of experiments on database consisting of 4000 images have ensured the high average index (98%) on recognition accuracy.

References

[1] Choy Steven S. O., Choy Clifford Sze-Tsan, Siu Wan-Chi. New single-pass algorithm for parallel thinning // Computer Vision and Image Understanding.– 1995.– Vol. 62, №. 1.– pp. 69-77.

[2] Nishida H. Curve Description Based on Directional Features and Quasi-Convexity/Concavity // Pattern Recognition.– 1995.– Vol. 28, № 7.– pp. 1045-1051.

[3] Nishida H. Structural Feature Extraction Using Multiple Bases // Computer Vision and Image Understanding.– 1995.– Vol. 62, № 1.– pp. 78-89.

[4] Nishida H., Mori S. An Algebraic Approach to Automatic Construction of Structural Models // IEEE Transaction on Pattern Analysis and Machine Intelligence.– 1993.– Vol. 15, № 12.– pp. 1298-1311.

[5] Kohonen T. Self-organization and associative memory. –Springer-Verlag, 1989.– 312 p.

[6] R. Sadykhov, V. Samokhval, A. Selikhanovich. The system of handwritten character recognition on the base of feature selector with use of orthogonal transforms // Proceedings of IEEE International Workshop on Intelligent Signal Processing, Budapest, Hungary, 4-7 September 1999. / Budapest, 1999.– pp. 246-248.

[7] R. Sadykhov, O. Malenko. Application of neural networks in system of numbers recognition for railway carriages // Proceedings of The Fifth International Conference “Advanced Computer Systems”, ACS-98, Szczecin, Poland, 19-20 November 1998. / Szczecin, 1998.– pp. 99-105.

[8] Y. Hwang. Recognition of unconstrained handwritten numerals by a radial basis function neural network classifier // Pattern recognition letters. – 1997. – Vol. 18, № 7. – pp. 657-664.

[9] Christopher M. Bishop., “Neural networks for pattern recognition”, *Oxford: Clarendon press*, 1997, pp.165-191.

[10] Richard O. Duda, Peter E. Hart, *Pattern classification and scene analysis*, Mir, Moscow, 1976.

[11] Foor W., “Adaptive optical radial basis function neural network for handwritten digit recognition”, *Applied optics*, 1995, v.34, pp.7545-7555.

[12] S.Liao. On image analysis by moments, IEEE Transactions on Pattern Analysis and Machine Intelligence, v.PAMI-18, 1996, №3, pp.254-266

[13] Moody, J. and C.J. Darken, “Fast learning in networks of locally tuned processing units”, *Neural Computations*, 1989, vol. 1, No. 2, pp. 281-294.

[14] Mukundan R., Ramakrishnan K. R., "Fast Computation of Legendre and Zernike Moments", *Pattern Recognition*, 1995, vol. 28, No. 9, pp. 1433-1442.

[15] Powell, M.J.D., “Radial basis function for multivariable interpolation: a review”, *In J.C. Mason and M.G. Cox (Eds), Algorithms for Approximation,* 1987, pp. 143-167.