A METHOD FOR HANDWRITTEN CHARACTERS RECOGNITION BASED ON A VECTOR FIELD

Tetsuya Izumi, Tetsuo Hattori, Hiroyuki Kitajima, Toshinori Yamasaki
Graduate School of Engineering, Kagawa University, 2217-20 Hayashi-Cho, Takamatsu City, 761-0396 Japan

Keywords: Handwritten Characters Recognition, Feature Extraction, Vector Field and Fourier Transform

Abstract: In order to obtain a low computational cost method for automatic handwritten characters recognition, this paper proposes a combined system of two rough classification methods based on features of a vector field: one is autocorrelation matrix method, and another is a low frequency Fourier expansion method. In each method, the representation is expressed as vectors, and the similarity is defined as a weighted sum of the squared values of the inner product between input pattern and the reference patterns that are normalized eigenvectors of KL (Karhunen-Loeve) expansion. This paper also describes a way of deciding the weight coefficients based on linear regression, and shows the effectiveness of the proposed method by illustrating some experimentation results for 3036 categories of handwritten Japanese characters.

1 INTRODUCTION

Since there are very many kinds of categories (or pattern classes) in Japanese characters (Hiragana and Chinese characters) and so there are many similar patterns in those characters, it needs much computational cost, i.e., computing time and memory storage, to automatically recognize those handwritten character patterns at high correct recognition rate. For this problem, many researches have been done in recent years (see References).

However, we consider that they still require considerably high computation cost for the automated recognition of all Japanese handwritten characters. Therefore, in order to obtain a low cost recognition system with high accuracy, we think we still have to pursue simple and efficient rough classification method based on more effective feature extraction and similarity measure.

In this paper, we propose a recognition method using a vector field, aiming to effectively obtain the feature information on directions of character lines and their juxtaposition situation and so on.

Based on the feature point vector field, we present two rough classification methods and the combined one. The two rough classification methods depend on different representations for the distribution of feature point's vectors: one is an autocorrelation matrix and another Fourier expansion on low frequency domain that can be interpreted as a complex-valued function. In each of the methods, the representation is expressed as high dimensional vector, and the similarity is defined as a weighted sum of the squared values of the inner product between input pattern and the reference patterns that are eigenvectors of KL (Karhunen-Loeve) expansion.

This paper shows the effectiveness of the proposed combined method by giving the experimental results that the correct recognition rate of 92.2% for unknown pattern is obtained in 3036 categories of handwritten Japanese characters.

2 FEATURE EXTRACTION

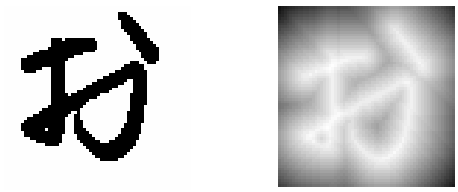
2.1 Vector Field

After a distance transformation is done for the binarized input pattern (Figure 1), and two-dimensional vector field is constructed by (1), where each vector corresponds to the gradient of the distance distribution at each point *P*, as shown in Figure 2. Let *T*(*P*) and *V*(*P*) be the value of distance transformation and two-dimensional vector at the point *P*, respectively. The *V*(*P*) is defined as follows.

Izumi T., Hattori T., Kitajima H. and Yamasaki T. (2004).
A METHOD FOR HANDWRITTEN CHARACTERS RECOGNITION BASED ON A VECTOR FIELD.
In *Proceedings of the First International Conference on Informatics in Control, Automation and Robotics*, pages 417-420
DOI: 10.5220/0001149004170420

$$V(P) = \sum_{i=1}^{8} \{T(P) - T(Q_i)\} \cdot e_i \qquad (1)$$

where, Q_i ($1 \leqq i \leqq 8$) shows each point of the eight neighborhood of point P, and e_i shows a unit length vector in the direction from the point P to Q_i.



(a) (b)

Figure 1:(a) Binarized pattern. (b) Distance transformation. (64x64 pixels.)

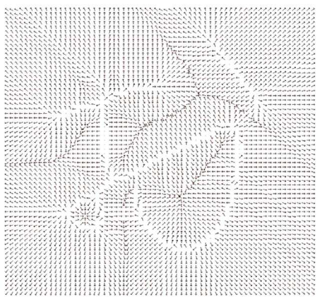


Figure 2: Vector Field.

2.2 Normalization and Divergence

The length of each vector on the field is normalized to be one or zero by a threshold. By divergence operation on the field, source points and sink points can be extracted as feature points. Those are called "*flow-out point*" and "*flow-in point*", respectively. Then at the same time, feature point vectors are obtained (Figure 3), which are vectors on the source points and sink points, what we call "*flow-out point vectors*" and "*flow-in point vectors*", respectively in the same manner to the above naming.

Figure 3: Feature point's vector field.

As a characteristic property of the feature point's vector field, flow-out and flow-in point vectors are located on the character lines (or strokes) and the background, respectively. They show not only the directional information on the strokes but also the juxtaposition situation of those strokes.

3 FEATURE REPRESENTATION AND SIMILARITY

3.1 Outline

After the construction of the above feature point's vector field, a combined method of two rough classifications is performed. The two classifications are based on different expressions of feature point's vector field: one is Autocorrelation Matrix Representation Method (AMRM) and another Fourier expansion in low frequency domain method (shortly we call Low Frequency Domain Method (LFDM)).

3.2 Autocorrelation Matrix Representation Method (AMRM)

The neighborhood vector pattern X of a feature point vector, i.e., 2-dimensional vectors on 3x3 points centering the feature point, can be represented as a 9-dimensional complex vector in which each complex-valued component means 2-dimensional vector. So, the X can also be regarded as an 18-dimensional real vector. In order to express the neighborhood pattern X effectively, we use an orthonormal (orthogonal and normalized) system that can be made from a set of nine typical neighborhood patterns by well-known Gram-Schmidt's Orthogonalization Method (GSOM). Actually, an orthonormal basis $\{\mu_i\}$ (i=1,…, 6) is obtained by the GSOM.

Then, we can represent the neighborhood vector pattern X of a feature point P as the following 8-dimensional real vector $\chi(P)$, using the coordinate (i, j) of the point P and a set of the real-valued inner products between the neighborhood pattern and each element of the above orthonormal system, i.e., $\{ \langle X \mid \mu_i \rangle \}$.

$$\chi(P) = \left(i, j, \langle X \mid \mu_1 \rangle, \cdots, \langle X \mid \mu_6 \rangle\right)^T \qquad (2)$$

So, a set of $\{ \chi(P) \}$ is extracted from the feature point vector field. Then, we express the distribution of the set of $\{ \chi(P) \}$ in the 8-dimensional real vector space by an autocorrelation matrix in 8x8 size.

Because the matrix is symmetric, it can be corresponded to a 36-dimensional real vector.

3.3 Low Frequency Domain Method (LFDM)

In this method, as a representation of input pattern, the Fourier expansion on the low frequency domain is used after the Fourier transform over the feature point vector field. The Fourier transform is described as follows. Let x and ω be two 2-dimensional real positional vectors on a real plane and the frequency domain, respectively. The Fourier transform $F(\omega)$ of the input pattern or complex-valued function $f(x)$ is defined in the following (3).

$$F(\omega) = \int_{R^2} f(x) e^{-j\langle \omega | x \rangle} \, d\mu(x) \qquad (3)$$

where j and $d\mu(x)$ mean an imaginary number unit and an area element, respectively.

For example of the Fourier transform, a character pattern and its amplitude spectrum image on the frequency domain are shown in Figure 4. In this figure, we can see that much information of the input pattern is in the low frequency domain (near the center of the image).

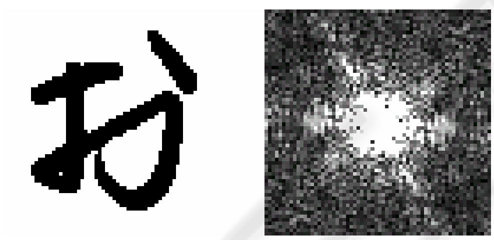


Figure 4: Input pattern (left, the same pattern in Figure 1(a)) and its amplitude Fourier spectrum image (right) where the original point is the center of the image.

Actually, as a feature representation of input pattern, we use the information on 10×10 points around the original point in the frequency domain of the Fourier transform from the feature point vector field. Therefore, an input pattern is corresponded to a 100-dimensional complex vector.

3.4 Reference Pattern and Similarity

As aforementioned, an input pattern is represented as a correspondent feature vector in each of the two classification methods. Then, an orthonormal basis (or orthonormal set) is made from eigenvectors of KL (Karhunen-Loeve) expansion for learning samples of each character pattern class (or category). The elements of the orthonormal basis are used as reference patterns for the category.

The similarity between input pattern and each category is defined as a weighted sum of the squared values of inner product between the feature vector and the reference patterns belonging to the category, as in the following (4).

Let f and g be an input pattern (or feature vector) and a category, respectively. Let $\{g^k{}_i\}(i=1,\ldots,n)(k=1,\ldots,m)$ be a set of reference patterns of the category g^k, and let $sim(f, g^k)$ be the similarity between f and g^k , the definition is given by (4).

$$sim(f, g^k) = \sum_{i=1}^{n} \frac{W_i \times \left| \langle f \mid g_i^k \rangle \right|^2}{\|f\|^2} \qquad (4)$$

where $\|f\| = \sqrt{\langle f \mid f \rangle}$, W_i ($W_i > 0$, i=1,…,n) shows one of weight coefficients .

4 DECISION OF WEIGHT COEFFICIENT

After the similarity computation, input pattern is classified into a category that gives the highest similarity in the above computation. Therefore, the weight coefficient is very influential in the similarity evaluation. In many cases, a set of the coefficients is defined by the eigenvalues of the KL expansion as in the following (5), what we simply call *Eigenvalue Similarity*.

$$W_i = \frac{\lambda_i}{\lambda_1} \qquad (5)$$

where λ_i ($\lambda_i > 0$, i=1,…, n) shows the i-th largest eigenvalue in the KL expansion.

However, from our experiences in this kind of character recognition, the largest eigenvalue is often much greater than the other eigenvalues, and so the similarity is decided by the first term of the inner product between the input and the first reference pattern. As a result, the recognition rate is sometimes worse than the case when W_i =1 for all i.

In order to decide the suitable weight coefficients for good recognition rate, we have adopted an iteration method based on a linear regression model, starting the initial condition that W_i =1 for all i. And, substituting the product of old and new coefficient into W_i , (i.e., $W_i^* W \rightarrow W_i$), the updated coefficients are obtained. Thus, we can iteratively search the suitable coefficients. The iteration terminates when no improvement of the recognition rate can be seen.

5 EXPERIMENTATION

The aforementioned two classification methods are combined by using a synthesized similarity as defined in (6). Let x and y be the similarity value between the input pattern and each category in the AMRM and LFDM, respectively. The following sum of squared similarity (like Euclid norm) is used.

$$\text{S\_Similarity} = x^2 + y^2 \tag{6}$$

Thus we have experimented the above three kinds of methods for 3036 categories of Japanese handwritten characters (total number of character patterns: 3036 x 20 patterns per category = 60,720) in ETL9B (Electro Technical Laboratory in Japan) database. The data used for experimentation includes not only Chinese characters but also Japanese Hiraganas.

In the experimentation, 10 samples (or character pattern) per category were used for learning, i.e., decision of reference patterns and the weight coefficients. Therefore, they are what we call *learning patterns*. Actually, we have decided that the number of the reference patterns (or eigenvectors of KL expansion) per category is eight, because the number has been the most effective for recognition of the learning patterns used in experimentation. The rest 10 patterns are experimented as *unknown pattern*.

The specification of the computer, OS, etc. that we used in this experimentation is as follows.

OS：Microsoft Windows XP Professional.
CPU：Intel PentiumⅣ(2.4GHz).
Main memory：1024Mbytes.
Programming language: Borland C++5.02J.

The experimental results are shown in Table 1 through 3. In order to compare the effects of three kinds of weight coefficient, i.e., no weight (W_i =1 for all i), eigenvalue, and weight coefficient decided by the linear regression model, the results in the three cases are also shown in the tables.

Table 1: Recognition rate by the AMRM.

Input Pattern \ Weight Coefficient	No Weight	Eigenvalue	Linear Regression Model (LRM)
Learning Pattern	94.01%	68.13%	94.23%
Unknown Pattern	60.72%	60.10%	71.91%

Execution time in LRM: 48 msec/pattern.
Required storage: 6.6 Mbytes.

Table 2: Recognition rate by the LFDM.

Input Pattern \ Weight Coefficient	No Weight	Eigenvalue	Linear Regression Model (LRM)
Learning Pattern	99.92%	91.83%	99.61%
Unknown Pattern	85.79%	81.30%	87.24%

Execution time in LRM: 66 msec/pattern.
Required storage: 18.5 Mbytes.

Table 3.: Recognition rate by the combined method.

Input Pattern \ Weight Coefficient	No Weight	Eigenvalue	Linear Regression Model (LRM)
Learning Pattern	99.93%	95.62%	99.81%
Unknown Pattern	90.13%	89.45%	92.20%

Execution time in LRM: 68 msec/pattern.
Required storage: approximately 25 Mbytes.

6 CONCLUSION

We have presented two classification methods and a combined one for handwritten characters recognition using features of the vector field. We have also presented a set of weight coefficients in the similarity, using the linear regression model (LRM). Moreover, we have revealed the experimental results. From the results, we can see that it is very effective to use the feature of the vector field and the decision of weight coefficients based on LRM. Therefore, we consider that the feature point's vector field method is promising and worthwhile refining in order to find more effective and low computational cost (in the sense of time and storage) method.

REFERENCES

Masato S. et al., 2001. A Discriminant Method of Similar Characters with Quadratic Compound Function, IEICE Transactions, Vol.J84-D2, No.8, pp.1557-1565, Aug. 2001 (in Japanese).

Takashi N. et al., 2000. Accuracy Improvement by Compound Discriminant Functions for Resembling Character Recognition, IEICE Transactions, Vol.J83-D2, No.2, pp.623-633, Feb. 2000 (in Japanese).

Kazuhiro S. et al., 2001. Accuracy Improvement by Gradient Feature and Variance Absorbing Covariance Matrix in Handwritten Chinese Character Recognition, IEICE Transactions, Vol.J84-D2, No.11, pp.2387-2397, Nov. 2001 (in Japanese).