Using Eigen-Deformations in Handwritten Character Recognition

Seiichi Uchida, Mohammad Asad Ronee, Hiroaki Sakoe
Department of Intelligent Systems, Kyushu University, Fukuoka, 812-8581, Japan
uchida@is.kyushu-u.ac.jp

Abstract

Deformations in handwritten characters have class-dependent tendencies. For example, characters of class "A" are often deformed by global slant transformation and never deformed to be similar to "R". In this paper, the extraction and the utilization of such tendencies called eigen-deformations are investigated for better performance of elastic matching based recognition systems. The eigen-deformations are extracted by the principal component analysis of actual deformations automatically collected by elastic matching. From experimental results it was shown that the extracted eigen-deformations represent typical deformations of each class. It was also shown that the recognition performance can be improved significantly by using the eigen-deformations in detecting overfitting, which often results in misrecognition.

1. Introduction

Elastic image matching is one of the promising techniques for handwritten character recognition. In elastic matching based recognition, each reference image pattern is nonlinearly warped so that its shape becomes as close to the shape of an input image pattern as possible, and then a (dis)similarity between those patterns is evaluated. Although the elastic matching based recognition has shown good performance, it often suffers from misrecognitions due to *overfitting*, where the similarity to a wrong reference pattern is overestimated by an undesirable matching result. For example, an input pattern of "R" may be misrecognized as "A" because they have the same topological structure and thus their similarity tends to be overestimated by elastic matching.

The purpose of this paper is the reduction of the misrecognitions due to overfitting by using *eigen-deformations*, which are intraclass deformation tendencies, such that "A" is often deformed by global slant transformation and is never deformed to be similar to "R". For this purpose, the following two problems are addressed: (1) the extraction of eigen-deformations, and (2) the development of an improved recognition technique with eigen-deformations.

The key idea for the first problem is the principal component analysis (PCA) of actual handwritten deformations automatically collected by elastic matching. Consequently, eigen-deformations are provided as the principal axes (i.e., the directions with larger variances) of the distribution of the collected deformations.

The key idea for the second problem is the incorporation of *a posteriori* evaluation based on eigen-deformations into conventional elastic matching based recognition. Specifically, if the result of the elastic matching between an input and a reference pattern does not agree with the eigen-deformations of the reference, that result is considered to be overfitting. **Figure 1** illustrates the present technique based on those ideas.

In conventional elastic matching based recognition [1]-[7], universal deformation characteristics such as monotonicity, continuity, and smoothness, are often considered, but class-dependent deformation characteristics, i.e., eigen-deformations are not. HMM-based techniques [8], [9] are promising but their Markovian property limits manageable deformations to local ones.

The present technique for the extraction of eigen-deformations can be considered as an extension of Point Distribution Model (PDM) [10], [11], where the deformations of a contour are manually collected as the displacement of landmarks on the contour and then subjected to PCA. Compared to PDM, the present technique has the following features: (1) deformations are automatically collected by elastic matching, (2) deformations are analyzed categorywise, (3) two-dimensional, or planar deformations of character images are considered, (4) the relation between eigen-deformations and elastic matching is of interest.

2. Extraction of eigen-deformations

2.1. Collection of displacement fields

In general, elastic matching can be defined as the technique to obtain the optimal pixel-to-pixel correspondence between a pair of $I \times I$ images, $\boldsymbol{A} = \{\boldsymbol{a}(i,j)\}$ and $\boldsymbol{B} = \{\boldsymbol{b}(x,y)\}$ (**Fig. 2**). Let $(x_{i,j}, y_{i,j})$ denote the pixel on $\boldsymbol{B}$ corresponding to pixel (i,j) of $\boldsymbol{A}$. The correspondence, i.e., elastic matching $\{(x_{i,j}, y_{i,j}) \mid i,j = 1,\ldots,I\}$

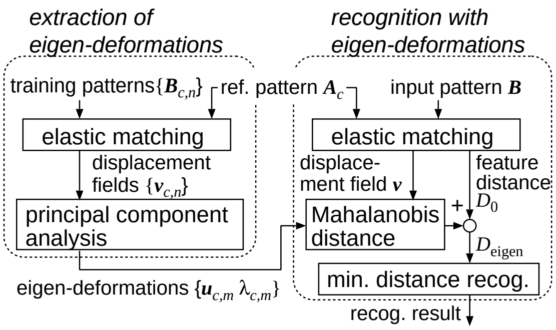

Figure 1. Diagram of the present technique.

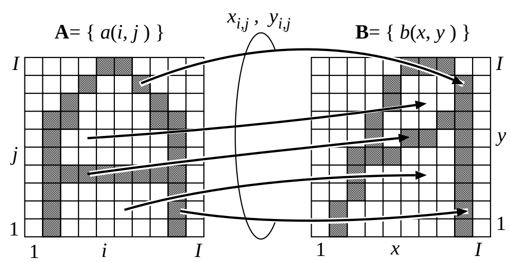

Figure 2. Elastic matching of a pair of images.

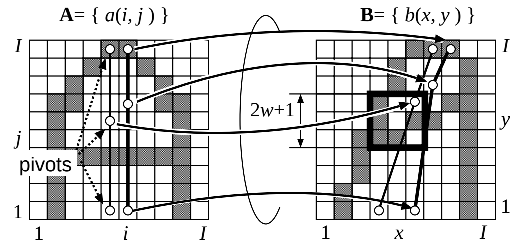

Figure 3. Piecewise linear two-dimensional warping [7], [16].

is obtained by minimizing an objective function such as

$$J = \min_{\{(x_{i,j}, y_{i,j})\}} \sum_{i=1}^{I} \sum_{j=1}^{I} \delta(\boldsymbol{a}(i,j), \boldsymbol{b}(x_{i,j}, y_{i,j})) \quad (1)$$

under some constraints on $\{(x_{i,j}, y_{i,j})\}$, where $\delta(\cdot,\cdot)$ is a function for the distance between two pixel feature vectors.

Let $\boldsymbol{A}_c$ denote the reference pattern of class $c \in \{1,\ldots,C\}$ and $\{\boldsymbol{B}_{c,n} \mid n = 1,\ldots,N\}$ denote the set of training patterns of the same class c. The deformation in $\boldsymbol{B}_{c,n}$ can be represented as $\boldsymbol{v}_{c,n} = ((1 - x_{1,1}, 1 - y_{1,1}), \ldots, (i - x_{i,j}, j - y_{i,j}), \ldots, (I - x_{I,I}, I - y_{I,I}))$ where $\{(x_{i,j}, y_{i,j})\}$ is automatically provided by the elastic matching between $\boldsymbol{A}_c$ and $\boldsymbol{B}_{c,n}$. Hereafter, this vector $\boldsymbol{v}_{c,n}$ is called *displacement field* and its dimension is denoted as M (here, $M = 2I^2$). Thus, using all $\boldsymbol{B}_{c,n}$, N displacement fields are collected as $\boldsymbol{V}_c = \{\boldsymbol{v}_{c,1}, \ldots, \boldsymbol{v}_{c,n}, \ldots, \boldsymbol{v}_{c,N}\}$.

It should be noted that although the displacement fields can be corrected by using any elastic matching technique, they are affected by the property of the technique, such as objective function, constraints, and optimization scheme employed.

2.2. Analysis of collected displacement fields

Eigen-deformations of class c are defined as the principal axes of the distribution of displacement fields in $\boldsymbol{V}_c$. Thus, eigen-deformations can be extracted by applying PCA to $\boldsymbol{V}_c$. Namely, letting $\boldsymbol{\Sigma}_c$ be the covariance matrix of the displacement fields $\boldsymbol{V}_c$, its M-dimensional eigenvectors, denoted as $\{\boldsymbol{u}_{c,1}, \ldots, \boldsymbol{u}_{c,m}, \ldots, \boldsymbol{u}_{c,M}\}$, are the eigen-deformations of class c. The eigenvalues of $\boldsymbol{\Sigma}_c$, denoted as $\{\lambda_{c,1}, \ldots, \lambda_{c,m}, \ldots, \lambda_{c,M}\}$ ($\lambda_{c,m} \geq \lambda_{c,m+1}$), represent the contribution of each eigen-deformation.

3. Recognition using eigen-deformation

Let $D_0(\boldsymbol{A}_c, \boldsymbol{B})$ denote the minimum of the objective function (1) for reference $\boldsymbol{A}_c$ and input $\boldsymbol{B}$. Compared to conventional elastic matching based recognition where the distance $D_0(\boldsymbol{A}_c, \boldsymbol{B})$ is directly used for discrimination, the following distance is used here:

$$D_{\text{eigen}}(\boldsymbol{A}_c, \boldsymbol{B}) = (1-\alpha) D_0(\boldsymbol{A}_c, \boldsymbol{B}) + \alpha P(\boldsymbol{A}_c, \boldsymbol{B}) \quad (2)$$

where α is a constant ($0 \leq \alpha \leq 1$) and

$$\begin{aligned} P(\boldsymbol{A}_c, \boldsymbol{B}) &= (\boldsymbol{v} - \overline{\boldsymbol{v}}_c)^T \Sigma_c^{-1} (\boldsymbol{v} - \overline{\boldsymbol{v}}_c) \\ &= \sum_{m=1}^{M} \frac{\langle \boldsymbol{v} - \overline{\boldsymbol{v}}_c, \boldsymbol{u}_{c,m} \rangle^2}{\lambda_{c,m}} \end{aligned} \quad (3)$$

where $\boldsymbol{v}$ is the displacement field given by the elastic matching for $D_0(\boldsymbol{A}_c, \boldsymbol{B})$ and $\overline{\boldsymbol{v}}_c = \sum_n \boldsymbol{v}_{c,n}/N$. The *a posteriori* evaluation term $P(\boldsymbol{A}_c, \boldsymbol{B})$ is Mahalanobis distance between $\overline{\boldsymbol{v}}_c$ and $\boldsymbol{v}$. In the case of overfitting, displacement field $\boldsymbol{v}$ does not agree with the eigen-deformations of class c and then $P(\boldsymbol{A}_c, \boldsymbol{B})$ becomes large. Thus $D_{\text{eigen}}(\boldsymbol{A}_c, \boldsymbol{B})$ also becomes large and $\boldsymbol{B}$ will not be discriminated into class c.

It is well known that the numerical and estimation errors of higher-order eigenvalues are amplified in (3). Thus, in practice, a modified Mahalanobis distance [12], [13] is employed where the higher-order eigenvalues $\lambda_{c,m}$ ($m = M'+1, \ldots, M$) are replaced by $\lambda_{c,M'+1}$.

4. Experimental results

4.1 Database

Experiments were conducted on 26 character classes of capital English alphabets. For each class, 1100 handwritten character samples of database ETL6, offered by Electrotechnical Laboratory, were prepared. Four-dimensional directional features [14] [15, p.390] were used for pixel features as well as one-dimensional gray-level feature. Accordingly, as the feature vector distance $\delta(\cdot,\cdot)$, weighted L1 distance (cf. [7]) was used in the following experiments. Each sample was normalized so that its size $I = 20$. The

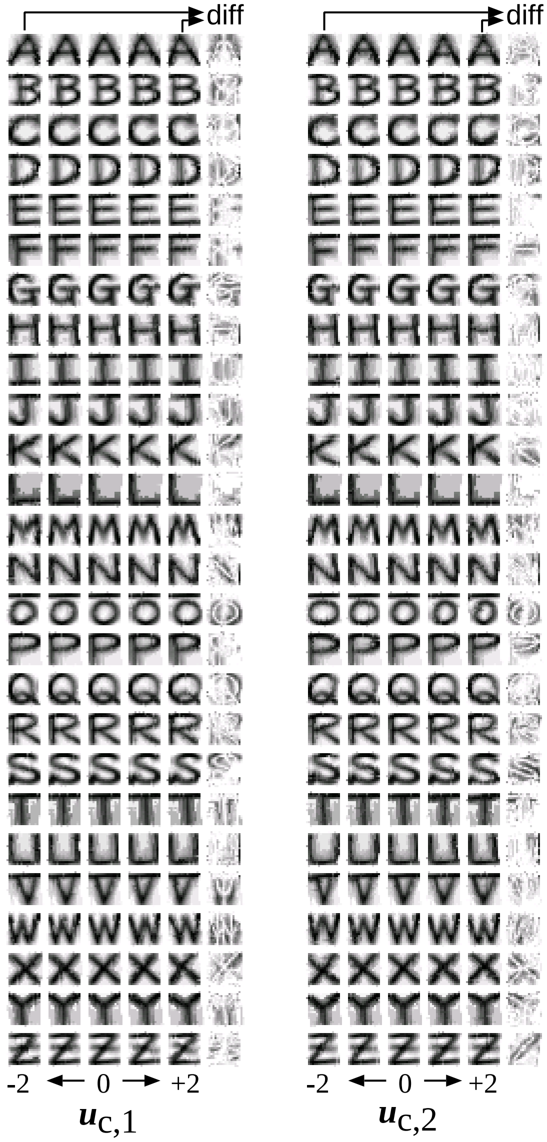

Figure 4. Reference patterns deformed by extracted eigen-deformations.

first 100 samples were simply averaged as the reference pattern and next 500 samples were used as training samples in the eigen-deformation extraction experiment of **4.3**. The remaining 500 samples were used as test samples in the recognition experiment of **4.4**.

4.2. Elastic matching

In our experiment, piecewise-linear two-dimensional warping (PL2DW) [7], [16] was used as an elastic matching technique. **Figure 3** illustrates PL2DW. In PL2DW, the entire mapping from $\boldsymbol{A}$ into $\boldsymbol{B}$ is approximated by the mapping of several pixels, called pivots, specified on each column of $\boldsymbol{A}$. The mapping of non-pivot points is determined by linear interpolation of the pivot mapping. In the experiment, three pivots were simply placed at $j = 1$(bottom),

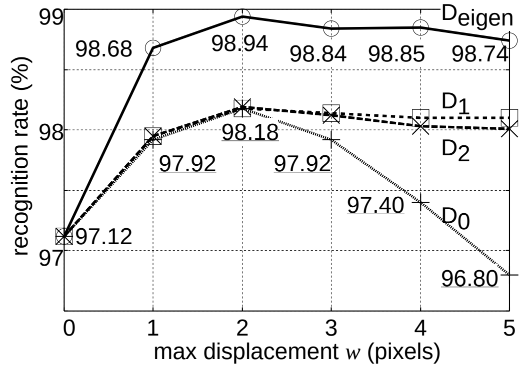

Figure 5. Recognition rate as a function of displacement range w.

10(center), and I(top) on each column.

The ranges of pivot displacement, i.e., $|i - x_{i,j}|$ and $|j - y_{i,j}|$ are limited not to exceed w (cf. thick-lined square in **Fig. 3**). Larger w results in both wider deformation compensability and higher probability of overfitting. In the experiment of **4.4**, the effect of the present framework will be examined while changing this parameter w.

It should be noted that the dimension of displacement fields $\boldsymbol{v}_{c,n}$ is reduced by the use of PL2DW since the displacements of non-pivotal pixels are linearly dependent on the displacements of pivots. In our case, the dimension is 74 and far less than $2I^2 = 800$. This reduction is favorable from the viewpoint of statistical reliability using a limited number of training patterns.

4.3. Result(1): eigen-deformation extraction

An eigen-deformation extraction experiment was performed by using PL2DW. For each character class, PL2DW ($w = 2$) was performed between one reference pattern ($\boldsymbol{A}_c$) and each of 500 training patterns ($\boldsymbol{B}_{c,n}$) and the resulting 500 deformations ($\boldsymbol{v}_{c,n}, M = 74$) were subjected to PCA to obtain 74 eigen-deformations $\boldsymbol{u}_{c,m}$.

Figure 4 shows the reference patterns deformed by two dominant eigen-deformations, $\boldsymbol{u}_{c,1}$ and $\boldsymbol{u}_{c,2}$ at $w = 2$. Foe each class, five patterns were synthesized by setting the displacement of each pixel (i, j) of $\boldsymbol{A}_c$ as $(\overline{\boldsymbol{v}}_c - 2\sqrt{\lambda_{c,m}}\boldsymbol{u}_{c,m})$, $(\overline{\boldsymbol{v}}_c - \sqrt{\lambda_{c,m}}\boldsymbol{u}_{c,m})$, $\overline{\boldsymbol{v}}_c$, $(\overline{\boldsymbol{v}}_c + \sqrt{\lambda_{c,m}}\boldsymbol{u}_{c,m})$, and $(\overline{\boldsymbol{v}}_c + 2\sqrt{\lambda_{c,m}}\boldsymbol{u}_{c,m})$, respectively. From this result, it can be observed that most of these dominant eigen-deformations represent the deformations which frequently appear in handwritten characters. For example, the first eigen-deformation of “A” represents the global slant transformation. Similarly, the first and the second eigen-deformations of “B” are the horizontal translation of its vertical stroke and the change in the prominence of its upper part, respectively.

In almost all classes, the cumulative proportion $\sum_{m=1}^{s} \lambda_{c,m} / \sum_{m=1}^{M} \lambda_{c,m}$ exceeds 50% at $s = 3 \sim 5$ and 80% at $s = 10 \sim 20$. Thus, the distribution of displacement fields is not isotropic in M-dimensional space and can be approximated by several eigen-deformations.

4.4. Result(2): character recognition using eigen-deformations

A recognition experiment was performed using D_{eigen} and D_0. As mentioned in **4.1**, the 500 test samples independent of the training samples were subjected. PL2DW with the same w was used in both the eigen-deformation extraction and the recognition processes.

Figure 5 shows the recognition rates by D_0 and D_{eigen} as functions of the displacement range w. The parameters α and M' were optimized at each measurement. When $w = 0$, both the distances are equivalent to the distance given by simple rigid matching. In **Fig. 5**, it is shown that the highest rate 98.94% was attained by the present technique (D_{eigen}) at $w = 2$. This means that D_{eigen} attained 42% reduction of misrecognitions by D_0 at the best rate (98.18%).

The categorywise inspection of the best results shows that the number of samples “M” recognized as “H” was reduced from 30 to 13. It is reported in [7] that these misrecognitions were typical ones due to the overfitting in PL2DW. Thus, their reduction clarifies the effect of eigen-deformations on the the suppression of overfitting.

This effect is confirmed by the change of recognition rates according to w. **Figure 5** shows that the recognition rate by D_0 decreases drastically and finally becomes lower than that by rigid matching as w increases. This degradation is mainly due to the increase of overfitting by allowing larger displacement. The recognition rate by D_{eigen}, however, only shows slight degradation even at $w = 5$.

The usefulness of eigen-deformations can also be shown by comparing with the following distances:

$$\begin{aligned} D_1(\boldsymbol{A}_c, \boldsymbol{B}) &= (1-\beta) D_0(\boldsymbol{A}_c, \boldsymbol{B}) + \beta \|\boldsymbol{v}\|, \\ D_2(\boldsymbol{A}_c, \boldsymbol{B}) &= (1-\gamma) D_0(\boldsymbol{A}_c, \boldsymbol{B}) + \gamma \|\boldsymbol{v} - \overline{\boldsymbol{v}}_c\|. \end{aligned}$$

In these distances, the amplitude of the deformation in $\boldsymbol{B}$ is evaluated but the direction of the deformation is not. **Figure 5** shows that they have effects on the suppression of overfitting but those effects are weaker than the effect of D_{eigen}. Namely, although these two distances also are stable to the increase of w like D_{eigen}, their highest recognition rates were almost the same as that by D_0.

5. Conclusion

The extraction and utilization of eigen-deformations of handwritten characters were investigated with handwritten character recognition. From experimental results, the following points were clarified.

1. The distribution of displacement fields automatically collected by elastic matching is not isotropic and therefore can be approximated by the several principal axes called eigen-deformations.
2. The extracted eigen-deformations represent the deformations which frequently appear in handwritten characters.
3. Misrecognitions due to overfitting, the side effect of elastic matching, can be reduced by using eigen-deformations. In fact, 98.84% recognition rate was attained by using eigen-deformations in *a posteriori* evaluation. This is 42% reduction of misrecognitions.

References

[1] Y. Nakano, et al., “Improvement of Chinese character recognition using projection profiles,” *Proc. IJCPR*, 172-178, 1973.

[2] Y. Mizukami, “A handwritten Chinese character recognition system using hierarchical displacement extraction based on directional features,” *Pat. Recog. Let.*, 19(7):595-604, 1998.

[3] A. K. Jain and D. Zongker, “Representation and recognition of handwritten digits using deformable templates,” *IEEE Trans. PAMI*, 19(12):1386-1391, 1997.

[4] M. Nakagawa, et al., “An off-line character recognition method employing model-dependent pattern normalization by an elastic membrane model,” *Proc. ICDAR*, 495-498, 1999.

[5] T. Wakahara, et al., “Affine-invariant recognition of gray-scale characters using global affine transformation correlation,” *IEEE Trans. PAMI*, 23(4):384–395, 2001.

[6] S. Uchida and H. Sakoe, “Handwritten character recognition using monotonic and continuous two-dimensional warping,” *Proc. ICDAR*, 499-502, 1999.

[7] M. A. Ronee, et al., “Handwritten character recognition using piecewise linear two-dimensional warping,” *Proc. ICDAR*, 39–43, 2001.

[8] E. Levin and R. Pieraccini, “Dynamic planar warping for optical character recognition,” *Proc. ICASSP*, 3:149–152, 1992.

[9] S. Kuo and O. Agazzi, “Keyword spotting in poorly printed documents using pseudo 2-d hidden Markov models,” *IEEE Trans. PAMI*, 16(8):842–848, 1994.

[10] T. F. Cootes, et al., “Active shape models - their training and application,” *Comput. Vis. Image Und.*, 61(1):38–59, 1995.

[11] A. Lanitis, et al., “Automatic interpretation and coding of face images using flexible models,” *IEEE Trans. PAMI*, 19(7):743–756, 1997.

[12] M. Kurita, et al., “Handprinted Kanji and Hirakana character recognition using weight direction index histograms and quasi-Mahalanobis distance,” *IEICE Japan, Tech. Rep.*, PRL82-79, 1982. (in Japanese)

[13] F. Kimura, et al., “Modified quadratic discriminant functions and the application to Chinese character recognition,” *IEEE Trans. PAMI*, 9(1):149–153, 1987.

[14] T. Saito, et al., “An analysis of handprinted Chinese characters by directional pattern matching approach,” *Trans. IEICE*, J65-D(5):550–557, 1982. (in Japanese)

[15] S. Mori, et al., “Research on machine recognition of handprinted characters,” *IEEE Trans. PAMI*, 6(4):386–405, 1984.

[16] S. Uchida and H. Sakoe, “Piecewise linear two-dimensional warping,” *Proc. ICPR*, 3:538–541, 2000.