PERGAMON

Pattern Recognition 36 (2003) 2031–2040

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Eigen-deformations for elastic matching based handwritten character recognition

Seiichi Uchida*, Hiroaki Sakoe

Faculty of Information Science and Electrical Engineering, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi 812-8581, Japan

Received 7 June 2002; accepted 13 January 2003

Abstract

Deformations in handwritten characters have category-dependent tendencies. In this paper, the estimation and the utilization of such tendencies called eigen-deformations are investigated for the better performance of elastic matching based handwritten character recognition. The eigen-deformations are estimated by the principal component analysis of actual deformations automatically collected by the elastic matching. From experimental results it was shown that typical deformations of each category can be extracted as the eigen-deformations. It was also shown that the recognition performance can be improved significantly by using the eigen-deformations for the detection of overfitting, which is the main cause of the misrecognition in the elastic matching based recognition methods.

Keywords: Handwritten character recognition; Elastic matching; Eigen-deformation; Principal component analysis

1. Introduction

Elastic image matching, or deformable template is one of the promising techniques for handwritten character recognition [1]. In elastic image matching based character recognition, an input pattern is nonlinearly fitted to a reference pattern while minimizing their distance (i.e., dissimilarity) as possible. The minimized distance is ideally invariant to the deformations in the input pattern and thus can be directly used as a discriminant function.

Although the elastic matching based recognition methods have shown good performance, they often suffer from misrecognitions due to *overfitting*, which is the phenomenon that the input pattern is closely fitted to the reference pattern of an incorrect category. For example, an input pattern of "R" may be misrecognized as "A" due to the overfitting because they have the same topological structure and their distance tends to be underestimated by the elastic matching.

The purpose of this paper is to improve the recognition accuracy of the elastic matching based recognition method by reducing the misrecognitions due to the overfitting. For this purpose, the following two problems are tackled in this paper:

- the estimation of *eigen-deformations*, and
- the incorporation of the eigen-deformations into the elastic matching based recognition framework.

The eigen-deformations are defined as the intrinsic deformation directions of each character category. For example, in category "A", global skew transformation and the vertical shift of its horizontal stroke are the candidates for the eigen-deformations. For the estimation of the eigen-deformations, the elastic matching and the principal component analysis (PCA) are used. Specifically, the elastic matching is used to collect actual deformations of handwritten characters and PCA is used to estimate the dominant directions of the collected deformations as the eigen-deformations.

* Corresponding author.
E-mail address: uchida@is.kyushu-u.ac.jp (S. Uchida).

0031-3203/03/$30.00 © 2003 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.
doi:10.1016/S0031-3203(03)00039-6

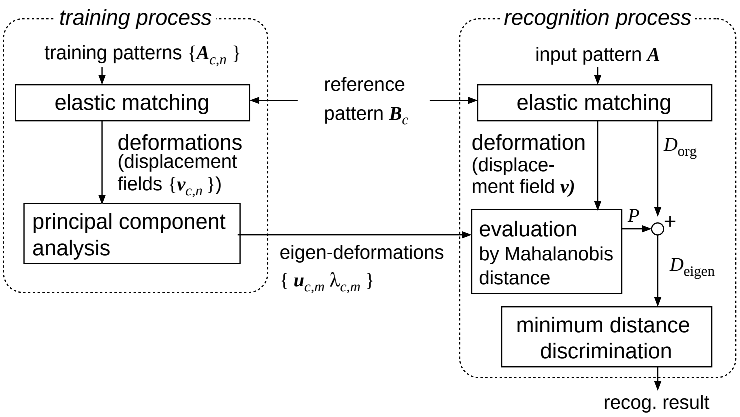

Fig. 1. Diagram of the present recognition method.

The key idea of the incorporation of the eigen-deformation into the recognition framework is that the overfitting can be considered as the deviation from the eigen-deformations. This idea will be natural because any intra-category deformation should be well-approximated by a combination of the eigen-deformations of the category. As the simplest way to reduce the effect of the overfitting based on the idea, a posterior evaluation for measuring the derivation of the result of the elastic matching is newly added to the conventional framework.

Fig. 1 shows the diagram of the present method. The present method consists of two processes, i.e., the training process for the estimation of the eigen-deformations, and the recognition process using the estimated eigen-deformations. Note that the present method can employ any elastic matching technique in the training and the recognition processes. Conversely, the performance of the present recognition method depend on the elastic matching employed. This dependency will be shown through the experimental result of Section 4.

In conventional elastic matching based handwritten character recognition methods [2–9], category-independent and intuitive deformation characteristics (such as monotonicity, continuity, and smoothness) are often considered, but category-dependent and well-grounded deformation characteristics are not. HMM-based techniques [10,11] are promising on their ability to learn the actual deformations but their Markovian property limits manageable deformations to local ones.

The idea of the eigen-deformations was motivated by the Point Distribution Models (PDM) proposed by Cootes et al. [12,13] where the deformations of a contour are manually collected as the displacement of landmarks on the contour and then subjected to PCA. Several extensions of the PDM have been proposed by several researchers. For example, Shen and Davatzikos [14] have introduce an automatic displacement collection scheme into the PDM. Naster et al. [15] have proposed a similar method where two-dimensional (planar) deformations are collected automatically. Compared to those works, the present method has the following features: (a) The eigen-deformations are applied for character recognition. (b) The eigen-deformations are analyzed category-wise. (c) The dependency between the employed elastic matching technique and the estimated eigen-deformations is emphasized.

2. Elastic matching

As shown in Fig. 1, elastic matching plays two important roles in the present method. Thus, in this section, the elastic matching is outlined in advance to the detailed discussion of the present method. The elastic matching formulated here is somewhat a general one. This is because the present recognition method can employ any elastic matching techniques.

Consider two $I \times I$ character images $\boldsymbol{A} = \{\boldsymbol{a}_{i,j} \mid i,j = 1,2,\ldots,I\}$ and $\boldsymbol{B} = \{\boldsymbol{b}_{x,y} \mid x,y = 1,2,\ldots,I\}$, where $\boldsymbol{a}_{i,j}$ and $\boldsymbol{b}_{x,y}$ are feature vectors of pixel (i,j) and (x,y), respectively. The elastic matching between these images are defined as an optimization problem with respect to a 2D–2D mapping $\boldsymbol{F}$ described as a $2I^2$-dimensional integer-valued vector, i.e.,

$$\boldsymbol{F} = ((x_{1,1}, y_{1,1}), \ldots, (x_{i,j}, y_{i,j}), \ldots, (x_{I,I}, y_{I,I})), \qquad (1)$$

where $(x_{i,j}, y_{i,j})$ is the mapping, or the corresponding pixel of (i,j) (Fig. 2). In this paper, it is assumed that $\boldsymbol{B}$ is a reference (i.e., standard) character image. [1] The mapping $\boldsymbol{F}$ is assumed to be not a one-to-one mapping and therefore some pixels on $\boldsymbol{B}$ will be "skipped".

[1] In the preliminary version of this paper [16], the direction of the mapping was opposite. Namely, $\boldsymbol{A} = \{\boldsymbol{a}_{i,j}\}$ was the reference pattern. The direction is changed in this paper since later experimental results have revealed the superiority of the new direction in recognition rates.

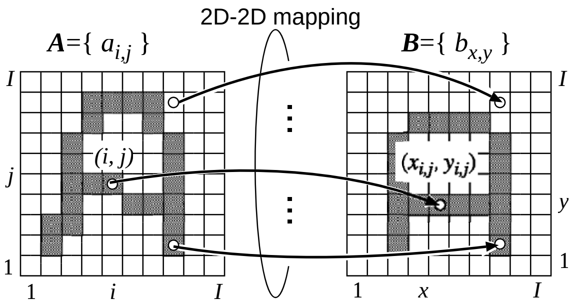

Fig. 2. A 2D–2D pixel mapping for elastic image matching.

The objective function of the optimization problem of $\boldsymbol{F}$ is described as

$$J_{\boldsymbol{A},\boldsymbol{B}}(\boldsymbol{F}) = \sum_{i=1}^{I}\sum_{j=1}^{I} \delta_{i,j}(x_{i,j}, y_{i,j}), \tag{2}$$

where $\delta_{i,j}(x, y)$ provides a distance (dissimilarity) between two pixel feature vectors $\boldsymbol{a}_{i,j}$ and $\boldsymbol{b}_{x,y}$. On the minimization of Eq. (2), several constraints are imposed on $\boldsymbol{F}$ to specify the flexibility (i.e., the range of compensable deformations) of the elastic matching.

The solution of the minimization problem of Eq. (2) is useful for character recognition at the following two points. Firstly, the minimal value of $J_{\boldsymbol{A},\boldsymbol{B}}(\boldsymbol{F})$, denoted as $D_{\mathrm{org}}(\boldsymbol{A},\boldsymbol{B})$, is ideally a deformation invariant distance between $\boldsymbol{A}$ and $\boldsymbol{B}$. Thus, the quantity $D_{\mathrm{org}}(\boldsymbol{A},\boldsymbol{B})$ can be directly used as a discriminant function. Secondly, the planar deformation in $\boldsymbol{A}$ can be derived from the optimized $\boldsymbol{F}$ as an $M(=2I^2)$-dimensional vector called *displacement field*:

$$\boldsymbol{v} = ((1 - x_{1,1}, 1 - y_{1,1}), \ldots, (i - x_{i,j}, j - y_{i,j}), \ldots, (I - x_{I,I}, I - y_{I,I})). \tag{3}$$

The conventional recognition methods based on D_{org} often suffer from the misrecognitions due to *overfitting*, which is the phenomenon that an input pattern is closely fitted to the reference pattern of an incorrect category and then the input is misclassified into the incorrect category. It has been reported in Ref. [6] that about a half of misrecognitions were due to the overfitting. There are mainly two types of overfitting. The first type is the overfitting between two topologically similar patterns (e.g., input “M”→ reference “H”, “T”→“Y”, “X”→“K”). The second type is more delicate and local one where a part of the reference pattern is skipped by the non-one-to-one mapping $\boldsymbol{F}$ (e.g., “F”→“P”, “E”→“B”).

3. The present recognition method using eigen-deformations

In this section, a detailed description of the present recognition method where the elastic matching mentioned above is utilized is provided. As illustrated in Fig. 1, the present method consists of two parts: the training process for the estimation of the eigen-deformations and the recognition process using the estimated eigen-deformations.

3.1. Estimation of eigen-deformations

Let $\boldsymbol{A}_{c,n}$ $(n = 1, \ldots, N,\ c = 1, \ldots, C)$ denote the nth training pattern of category c and $\boldsymbol{B}_c$ denote the reference pattern of category c. As noted in Section 2, the deformation in $\boldsymbol{A}_{c,n}$ can be provided as the displacement field $\boldsymbol{v}_{c,n}$ by the elastic matching between $\boldsymbol{A}_{c,n}$ and $\boldsymbol{B}_c$. Thus, by performing the elastic matching between the reference pattern and each of N training patterns, N displacement fields $\boldsymbol{V}_c = \{\boldsymbol{v}_{c,1}, \ldots, \boldsymbol{v}_{c,n}, \ldots, \boldsymbol{v}_{c,N}\}$ are automatically collected.

The eigen-deformations of category c are defined as the principal axes of the distribution of the displacement fields $\boldsymbol{V}_c$ in M-dimensional space. Thus, the eigen-deformations can be estimated by applying PCA to $\boldsymbol{V}_c$. Namely, letting $\boldsymbol{\Sigma}_c$ be the $M \times M$ covariance matrix of the displacement fields $\boldsymbol{V}_c$, its eigenvectors $\{\boldsymbol{u}_{c,m} \mid m = 1, \ldots, M\}$ and eigenvalues $\{\lambda_{c,m} \mid m = 1, \ldots, M,\ \lambda_{c,m} \geqslant \lambda_{c,m+1}\}$ are the eigen-deformations and their contributions of category c.

3.2. Recognition using eigen-deformations

Unlike the conventional recognition methods where the distance $D_{\mathrm{org}}(\boldsymbol{A},\boldsymbol{B}_c)$ is directly used for discrimination of an unknown input $\boldsymbol{A}$, the following distance is newly used in the present recognition method:

$$D_{\mathrm{eigen}}(\boldsymbol{A},\boldsymbol{B}_c) = (1 - \alpha)D_{\mathrm{org}}(\boldsymbol{A},\boldsymbol{B}_c) + \alpha P(\boldsymbol{A},\boldsymbol{B}_c), \tag{4}$$

where α is a constant $(0 \leqslant \alpha \leqslant 1)$. The function $P(\boldsymbol{A},\boldsymbol{B}_c)$ is an posterior evaluation term defined as

$$P(\boldsymbol{A},\boldsymbol{B}_c) = (\boldsymbol{v} - \bar{\boldsymbol{v}}_c)^{\mathrm{T}}\Sigma_c^{-1}(\boldsymbol{v} - \bar{\boldsymbol{v}}_c) = \sum_{m=1}^{M} \frac{\langle \boldsymbol{v} - \bar{\boldsymbol{v}}_c, \boldsymbol{u}_{c,m}\rangle^2}{\lambda_{c,m}}, \tag{5}$$

where $\boldsymbol{v}$ is the displacement field given by the elastic matching between $\boldsymbol{A}$ and $\boldsymbol{B}_c$ and $\bar{\boldsymbol{v}}$ is the average vector of $\boldsymbol{V}_c$. Thus, the function $P(\boldsymbol{A},\boldsymbol{B}_c)$ provides the Mahalanobis distance between $\bar{\boldsymbol{v}}_c$ and $\boldsymbol{v}$. In the case of overfitting, the displacement field $\boldsymbol{v}$ will not agree with the eigen-deformations of the category c and thus $P(\boldsymbol{A},\boldsymbol{B}_c)$ becomes large. Accordingly $D_{\mathrm{eigen}}(\boldsymbol{A},\boldsymbol{B}_c)$ also becomes large and the input $\boldsymbol{A}$ will not be discriminated into category c.

It is well known that the numerical and the estimation errors of higher-order eigenvalues are amplified in Eq. (5). Thus, in practice, a modified Mahalanobis distance [17,18] is employed, where the higher-order eigenvalues $\lambda_{c,m}$ ($m=$

$M'+2,\ldots,M$) are replaced by $\lambda_{c,M'+1}$, i.e.,

$$P(\boldsymbol{A},\boldsymbol{B}_c) \sim \sum_{m=1}^{M'} \frac{1}{\lambda_{c,m}} \langle \boldsymbol{v}-\bar{\boldsymbol{v}}_c, \boldsymbol{u}_{c,m}\rangle^2 + \sum_{m=M'+1}^{M} \frac{1}{\lambda_{c,M'+1}} \langle \boldsymbol{v}-\bar{\boldsymbol{v}}_c, \boldsymbol{u}_{c,m}\rangle^2 = \frac{1}{\lambda_{c,M'+1}} \|\boldsymbol{v}-\bar{\boldsymbol{v}}_c\| + \sum_{m=1}^{M'} \left(\frac{1}{\lambda_{c,m}} - \frac{1}{\lambda_{c,M'}}\right) \langle \boldsymbol{v}-\bar{\boldsymbol{v}}_c, \boldsymbol{u}_{c,m}\rangle^2, \quad (6)$$

where the parameter M' is to be determined experimentally.

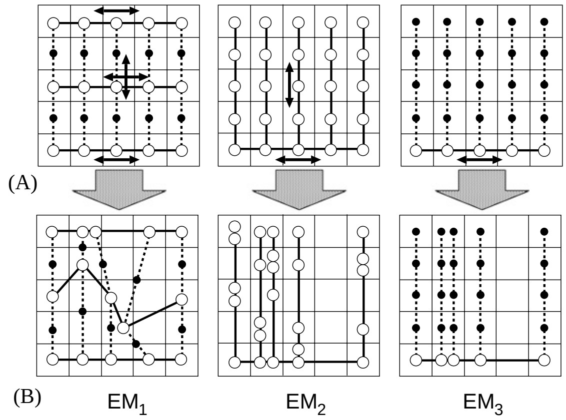

Fig. 3. Three elastic matching techniques EM_1–EM_3.

4. Experimental results

4.1. Database

Experiments were conducted on 26 character categories of capital English alphabets. For each category, 1100 handwritten character samples of database ETL6, offered by Electrotechnical Laboratory, were prepared. The following preprocessing procedure was performed on each sample: Firstly, four-dimensional directional features [19,20] were extracted at each pixel and used as pixel features together with one-dimensional gray-level feature. Thus, $\boldsymbol{a}_{i,j}$ and $\boldsymbol{b}_{x,y}$ were five-dimensional vectors. Then each sample was linearly scaled so that its character region fits to the 16×16 region centered on a 20×20 region (i.e., $I=20$).

After this preprocessing, the first 100 samples of each category were simply averaged to create the reference pattern of the category and the next 500 samples were used as training samples in the eigen-deformation estimation experiment of 4.3. The remaining 500 samples were used as test samples in the recognition experiment of 4.4.

According to the above feature extraction procedure, the distance δ in Eq. (2) was defined as

$$\delta_{i,j}(x,y) = |a^I_{i,j} - b^I_{x,y}| + \eta \sum_{d=1}^{4} |a^d_{i,j} - b^d_{x,y}|, \quad (7)$$

where superscripts "I" and "d" correspond to the intensity feature and the four directional features, respectively. The weighting coefficient η was fixed at 0.4 according to the result of a preliminary recognition experiment based on simple rigid matching.

4.2. Elastic matching techniques

In the following experiments, three different elastic matching techniques EM_1–EM_3 were subjected to observe the dependency of the experimental results on the characteristics of the elastic matching. Fig. 3 illustrates how the pixels on $\boldsymbol{A}$ are mapped into $\boldsymbol{B}$ by these elastic matching techniques. The unfilled circles on $\boldsymbol{A}$ represent the pixels whose mapping is to be optimized. The filled circles on $\boldsymbol{A}$ represent the pixels whose mapping linearly depends on the mapping of the pixels of the unfilled circles. The dashed line connecting two adjacent pixels indicates this dependency. The solid line connecting two adjacent pixels indicates that monotonicity and continuity constraints are imposed to preserve the left–right and upper–lower relation and the neighboring relation of these pixels.[2] For the details of these techniques, see Appendix A.

In all of these elastic matching techniques, boundary constraints and warp range constraints are commonly employed. The boundary constraints are imposed on the mapping to ensure the boundary of $\boldsymbol{A}$ is mapped into the boundary of $\boldsymbol{B}$. The warp range constraints are imposed on the mapping to limit the maximum compensable displacement and defined as

$$|i-x_{i,j}| \leqslant w, \quad |j-y_{i,j}| \leqslant w, \quad (8)$$

where w is a nonnegative integer. By increasing w, the flexibility becomes wider and thus the risk of overfitting also becomes higher.

Although the dimension of the displacement field $\boldsymbol{v}$ is $2I^2(=800)$ in Eq. (3), the actual dimension is far lower. This is because in the above elastic matching techniques there are many pixels whose mapping is linearly dependent on the mapping of other pixels or fixed by the boundary constraints. In fact, the actual dimension M of $\boldsymbol{v}$ in EM_1, EM_2, and EM_3 are 74, 378, and 18, respectively. Such low dimensionality is very favorable to obtain reliable eigen-deformations from a limited number of training samples (500, in our case). Conversely, highly flexible elastic matching techniques providing high-dimensional (near 800-dimensional) displacement fields, such as monotonic two-dimensional warping

[2] The monotonicity and continuity constraints employed here are discrete ones, such as $x_{i,j} = x_{i-1,j} + \{0,1,2\}$, and thus allow one-pixel skip.

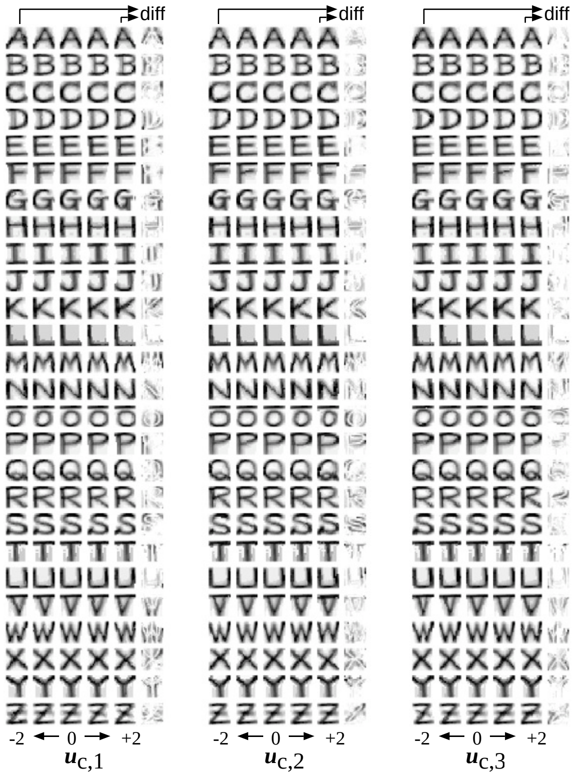

Fig. 4. Reference patterns deformed by the first three eigen-deformations $\boldsymbol{u}_{c,1}, \boldsymbol{u}_{c,2}$, and $\boldsymbol{u}_{c,3}$ ($EM_1, w=2$) with the multiplication by $k\sqrt{\lambda_{c,m}}$, $k=-2,-1,0,1,2$.

[11] and monotonic and continuous two-dimensional warping [21,22], are not favorable and thus were not subjected to our experiment.

4.3. Result (1): estimation of eigen-deformations

Firstly an eigen-deformation estimation experiment was performed using EM_1. For each character category c, 500 training patterns ($\{\boldsymbol{A}_{c,n}\}$) were matched to their reference pattern ($\boldsymbol{B}_c$) and the resulting 500 displacement fields ($\{\boldsymbol{v}_{c,n}\}, M=74$) were subjected to PCA to obtain 74 eigen-deformations $\{\boldsymbol{u}_{c,m}\}$.

Fig. 4 shows the reference patterns deformed by the first three eigen-deformations $\boldsymbol{u}_{c,1}$, $\boldsymbol{u}_{c,2}$, and $\boldsymbol{u}_{c,3}$ with the multiplication by $k\sqrt{\lambda_{c,m}}$ ($k=-2,-1,0,1,2$). From this, it can be observed that these eigen-deformations represent the deformations which frequently appear in actual handwritten characters. For example, the first three eigen-deformations of "A" represent global slant transformation, the vertical shift of the horizontal stroke, and the change of global roundness, respectively. Similarly, the first three eigen-deformations of "B" represent the horizontal shift of the vertical stroke, the change in the prominence of the upper and lower parts, and their mixture, respectively.

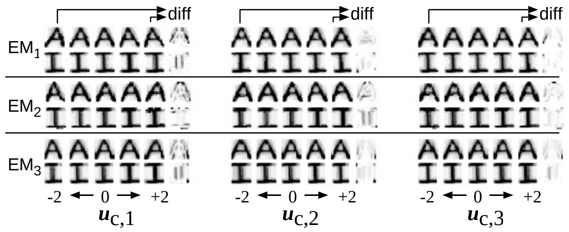

Fig. 5. The eigen-deformations estimated by using three different elastic matching techniques EM_1–EM_3.

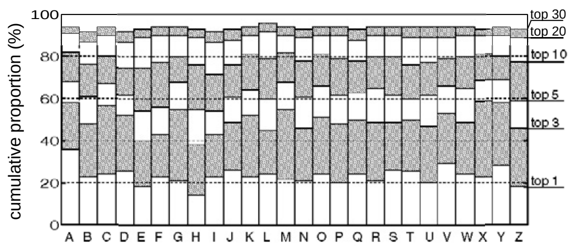

Fig. 6. Cumulative proportion of estimated eigen-deformations ($EM_1, w=2$).

Then similar experiments were performed on other two elastic matching techniques EM_2 and EM_3 and their results were compared to those of EM_1. Fig. 5 shows the eigen-deformations of category "A" and "I" of the three elastic matching techniques. This result shows the obvious dependency of the eigen-deformations on the employed elastic matching technique. In category "A", the eigen-deformation representing the vertical shift of the horizontal stroke can be found in EM_1 and EM_2 (as $\boldsymbol{u}_{c,2}$) and cannot be found in EM_3. This is because EM_3 cannot compensate deformations in the vertical direction. In category "I", the rotations (slant) of the vertical stroke can be found in EM_1 (as $\boldsymbol{u}_{c,2}$) and cannot be found in EM_2 and EM_3. This is because only EM_1 can compensate rotations among those three elastic matching techniques.

It was also experimentally shown that a small number of the eigen-deformations are enough to approximate all deformations in handwritten characters for any elastic matching. Fig. 6 shows the cumulative proportion $cp_c(s) = \sum_{m=1}^{s} \lambda_{c,m} / \sum_{m=1}^{M} \lambda_{c,m}$ of each category c for EM_1 at $w=2$. In all categories, the cumulative proportion exceeds 50% with the top 3–5 eigen-deformations and 80% with the top 10–20 eigen-deformations. Thus, the distribution of displacement fields is not isotropic in $M(=74)$-dimensional space and can be approximated by a small number of eigen-deformations. As summarized in Table 1, similar results were observed for the other two elastic matching techniques, EM_2 and EM_3.

Table 1
The average number of eigen-deformations required for the cumulative proportions over 50% and 80% (at $w=2$). The parenthesized number is its ratio to the number of all eigen-deformations, M

	EM_1	EM_2	EM_3
50%	3.3 (4.4%)	6.6 (1.7%)	1.4 (7.8%)
80%	11 (14.9%)	33 (8.7%)	3.7 (20.1%)

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

Recognition experiments using the present method were conducted. As noted in Section 4.1, 500 test samples independent of the training samples were subjected to the experiment. The same elastic matching technique was used at both the training process for estimating the eigen-deformations and the recognition process. Namely, if EM_1 was used at the training process, EM_1 was also used at the recognition process. The maximum displacement w was changed in the range of $[0,7]$. As noted in Section 4.2, larger w allows larger deformations and thus allows more serious overfittings. When $w=0$, EM_1–EM_3 are reduced to simple rigid matching. The parameters α and M' were manually optimized at each w and EM_1–EM_3.

The recognition results were observed from several viewpoints to confirm (a) the effect of the eigen-deformations on the improvement in recognition rates (Section 4.4.1), (b) the effect of the eigen-deformations on the reduction of the misrecognitions due to the overfitting (Section 4.4.2), and (c) the superiority of the present method over other ways to reduce the overfitting (Section 4.4.3).

4.4.1. Improvement in recognition rate

Fig. 7 shows the recognition rates attained by D_{eigen} and D_{org} of EM_1–EM_3 as the function of the maximum displacement w. It can be observed that for all of EM_1–EM_3 the recognition rates by D_{eigen} are always higher than those by D_{org}. From this observation, the usefulness of the eigen-deformation in handwritten character recognition is confirmed.

The best recognition rate 99.47% was attained by D_{eigen} of EM_1 at $w=6$. This was 0.35% improvement from the best recognition rate 99.12% by D_{org} of EM_1 at $w=3$. This improvement will be not trivial because about 40.0%(=0.35/(100 − 99.12)) of misrecognized samples by D_{org} were successfully avoided. The improvement of 0.35% corresponds to 46 samples and consists of 56 improved samples and 10 deteriorated samples.

4.4.2. Effect on the reduction of overfitting

Fig. 8 shows several examples of the above 56 improved samples. These examples indicate that most of their misrecognitions by D_{org} will be due to the overfitting. This is because incorrect reference patterns $\boldsymbol{B}_c$ are unnaturally deformed and very closely fitted to those test samples $\boldsymbol{A}$ by the elastic matching, as also shown in Fig. 8. Thus, their improvement means that the present distance D_{eigen} can penalize such serious overfitting by using the posterior evaluation term $P(\boldsymbol{A}, \boldsymbol{B}_c)$.

In addition to the above observation, the effect of the eigen-deformations on the reduction of the misrecognitions

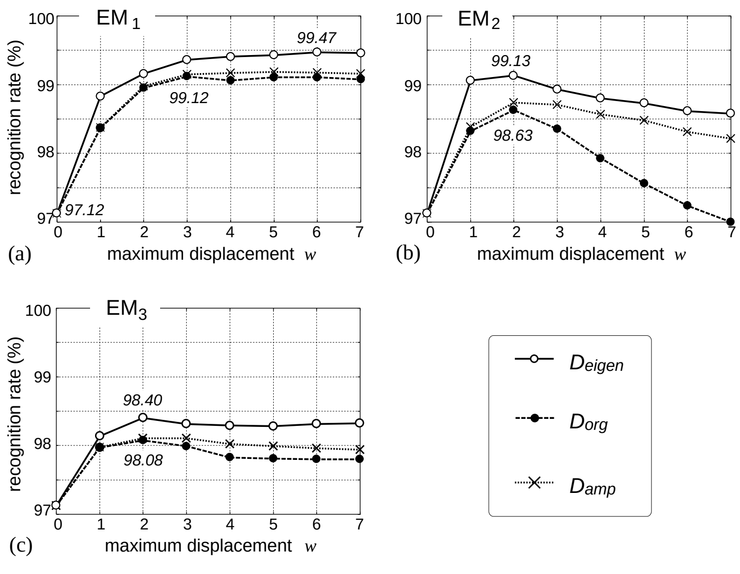

Fig. 7. Recognition rates of 13,000 handwritten characters.

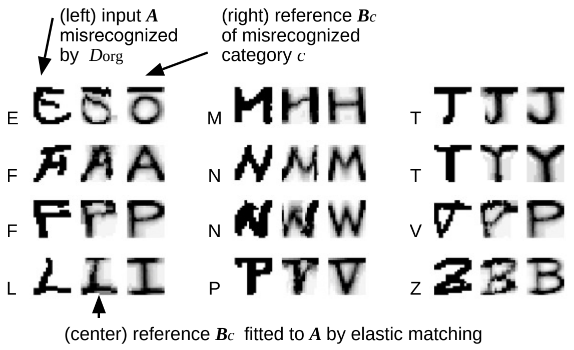

Fig. 8. Samples $\boldsymbol{A}$ misrecognized by D_{org}. For each $\boldsymbol{A}$, the reference pattern $\boldsymbol{B}_c$ of the misrecognized category and $\boldsymbol{B}_c$ fitted to $\boldsymbol{A}$ are also plotted. All these samples $\boldsymbol{A}$ were correctly recognized by D_{eigen}.

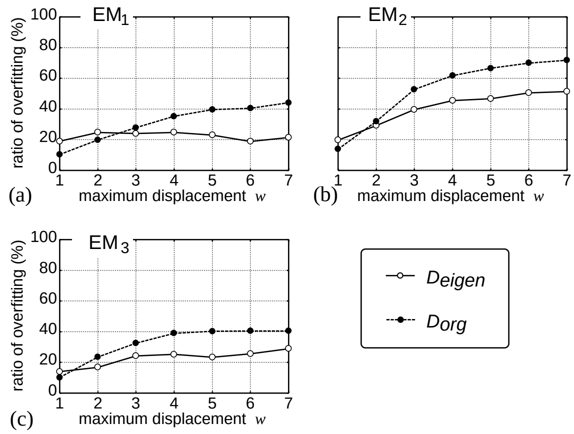

Fig. 10. Ratio of overfitting to all misrecognitions.

due to the overfitting can be confirmed from the significant improvement achieved by D_{eigen} of EM_2 (e.g., the improvement of 1.58% at $w = 7$). This is because in EM_2 the intra-column (vertical) matching is performed independently on each column and therefore the overfitting becomes serious especially at large w and thus the improvement suggests that the misrecognitions due to overfitting are reduced by D_{eigen}.

The effect can be also shown by the category-wise observation of the recognition rates. In EM_1, the major misrecognition by D_{org} was "F"→"P". This misrecognition is clearly due to the overfitting that pixels around the leftmost part of "P" are skipped (or thinned) in the matching. Using D_{eigen}, this misrecognition was reduced from 10 to 2 samples. The reduction of the same misrecognition "F"→"P" was also observed in the results of EM_3. In EM_2, the major misrecognition by D_{org} was "M"→"H". This misrecognition was typical one due to the overfitting between topologically similar patterns. By D_{eigen}, this misrecognition was reduced from 12 to 4.

Another quantitative evaluation of the expected effect can be confirmed by classifying 13,000 test samples into the four cases of Fig. 9. These four cases are different on the positional relation of the reference and the input patterns. Among the four cases, (c) and (d) represent the misrecognition by the elastic matching and (d) represents the misrecognition due to the overfitting more obviously than (c). Thus the number of samples falling into (d) can be considered as an index of the risk of overfitting. Fig. 10 plots the ratio of misrecognized samples of (d) to all misrecognized samples (i.e., (d)/((c) + (d))). The distance D_{eigen} provides almost always smaller ratios (i.e., lower risks) than the distance D_{org} and their difference becomes larger along with w. Thus, this result qualitatively shows the effect of D_{eigen} on the reduction of overfitting. Note that EM_2 shows higher ratios than EM_1 and EM_3 because EM_2 essentially has a higher risk of overfitting than EM_1 and EM_3 as noted above.

4.4.3. *Superiority over other ways to reduce overfitting*

Instead of D_{eigen} where both the amplitude and the direction of the displacement field $\boldsymbol{v}$ are evaluated using the eigen-deformations, the following distance can be used:

$$D_{\mathrm{amp}} = (1 - \beta) D_{\mathrm{org}}(\boldsymbol{A}, \boldsymbol{B}_c) + \beta \|\boldsymbol{v} - \bar{\boldsymbol{v}}_c\|, \tag{9}$$

where β is a nonnegative weighting coefficient. In the distance D_{amp}, only the amplitude of the displacement field $\boldsymbol{v}$ is evaluated. Fig. 7 shows that D_{amp} can achieve higher recognition rates than D_{org} at large w and thus has the effect on the reduction of overfitting. However, for each of EM_1–EM_3, the recognition rates of D_{amp} are always lower than those of D_{eigen}. In addition, the best recognition rate of D_{amp} is almost the same as that of D_{org}. Therefore, these facts show that the effect of D_{eigen} is superior than that of D_{amp}.

Another recognition experiment where "global" eigen-deformations were used commonly in all categories instead of $\boldsymbol{v}_{c,m}$ was conducted. The global eigen-deformations were

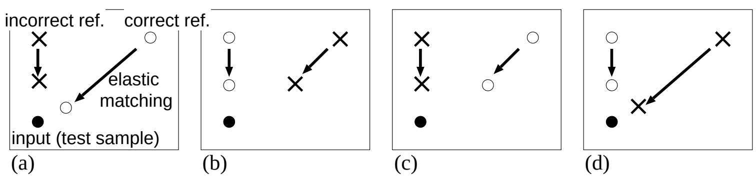

Fig. 9. A classification of the positional relations between input and reference patterns.

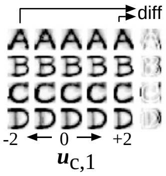

Fig. 11. Global eigen-deformations.

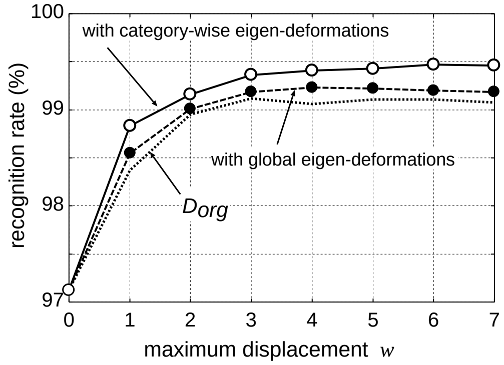

Fig. 12. Comparison between category-wise eigen-deformations and global eigen-deformations in terms of recognition rates.

estimated by PCA of the displacement fields of all categories, i.e., $\{\boldsymbol{V}_1, \ldots, \boldsymbol{V}_c, \ldots, \boldsymbol{V}_C\}$. In Fig. 11 the global eigen-deformations provided by EM_1 at $w = 2$ are shown. The first and the second global eigen-deformations represent the slant transformation. This means that the slant transformation is the deformation common in all categories. In Fig. 12, the recognition rates by the global eigen-deformations are compared with the rates by the category-wise eigen-deformations. Although the comparison with D_{org} shows the effect of the global eigen-deformations on the improvement in recognition accuracy, their effect is far weaker than that of the category-wise eigen-deformations. This result shows that the category-wise eigen-deformations are superior to the global eigen-deformations and thus the intrinsic deformation characteristics are different in every category.

4.5. Computational complexity

The recognition process based on D_{eigen} requires more computations than the conventional recognition process based on D_{org}. Their difference is $O(MM')$, which is the computational complexity to obtain $P(\boldsymbol{A}, \boldsymbol{B}_c)$. Fortunately, this difference is not large because the actual dimension M is often small (as discussed in Section 4.2) and M' $(< M)$ can be also small.

The reduction of the computational complexity will be necessary for larger tasks since the elastic matching itself is often costly. (For example, EM_1 requires $O(w^4 I^2)$ computations for the globally optimal matching between a pair of character images.) The most general complexity reduction technique is rough classification for reducing the number of candidate categories. If the employed elastic matching requires prohibitive computations, it will be necessary to use some approximation strategy for the elastic matching algorithm, such as coarse-to-fine strategies, local search techniques, and iterative techniques, at the cost of the optimality of the matching result.

5. Conclusion

For the better performance of elastic matching based handwritten character recognition, the estimation and the utilization of eigen-deformations were investigated. The eigen-deformations are the intrinsic deformations within each character category and can be estimated by the principal component analysis (PCA) of actual deformations automatically collected by the elastic matching. In the present elastic matching based recognition method, the estimated eigen-deformations are utilized in a posterior process to detect overfitting, which is often the main cause of the misrecognition.

From experimental results, the following points were shown:

- The estimated eigen-deformations represent the deformations which frequently appear in handwritten characters and depend on the characteristics of the elastic matching employed to collect actual deformations.
- The distribution of the actual deformations is not isotropic and therefore can be approximated by several principal axes, i.e., the eigen-deformations.
- The highest recognition rate of 13,000 English capital alphabet samples by the present recognition method was 99.47%. This was 0.35% improvement compared to the recognition result without the eigen-deformations. Equivalently, the number of misrecognized samples was successfully reduced to 60% by using the eigen-deformations. Such improvement could be constantly observed regardless of the elastic matching techniques.
- The eigen-deformations have a notable effect on the reduction of misrecognitions due to overfitting.
- Not only the amplitude but also the direction of the deformation are important to detect overfitting.
- The category-wise eigen-deformations are more effective than the global eigen-deformations estimated from the deformations of all categories. Thus, the intrinsic deformation characteristics are different in each category.

Acknowledgements

The authors thank the Electrotechnical Laboratory for providing us ETL6. This work was supported in part by The

Ministry of Education, Culture, Sports, Science and Technology in Japan under a Grant-in-Aid for Scientific Research No. 14780293. A preliminary version of this paper has been presented in the proceeding of ICPR2002.

Appendix A. Three elastic matching techniques used in experiments

The detailed description of three elastic matching techniques EM_1–EM_3 is as follows:

- EM_1—Piecewise linear two-dimensional warping (PL2DW) [6,23]. The mapping of pixels on each column of $\boldsymbol{A}$ is approximated by the linear interpolation of the mapping of several pixels, called pivots, pre-specified on the column. Monotonicity and continuity constraints are imposed on the mapping of neighboring pivots for topology preservation. In the experiments, three pivots were regularly placed at $j = 1$ (bottom), 10 (center), and 20 (top) on each column.
- EM_2—Elastic matching where each column of $\boldsymbol{A}$ is matched to a column of $\boldsymbol{B}$. In this inter-column, or column-wise matching, 1D–1D intra-column matching is also performed between matched two columns. The inter-column and the intra-column matchings are for compensating horizontal and vertical deformations, respectively. From the viewpoint of the flexibility, this elastic matching technique is equivalent to the one proposed by Kuo et al. [10].
- EM_3—Elastic matching where each column of $\boldsymbol{A}$ is matched to a column of $\boldsymbol{B}$. The 1D–1D intra-column matching in this inter-column matching is rigid. Thus, vertical deformations cannot be compensated. The flexibility of this technique is equivalent to the one proposed by Nakano et al. [5].

For each of three techniques, there is an efficient dynamic programming based algorithm. See their literatures [5,10,23] for details.

References

[1] Ø.D. Trier, A.K. Jain, T. Taxt, Feature extraction methods for character recognition—a survey, Pattern Recognition 29 (4) (1996) 641–662.
[2] A.K. Jain, D. Zongker, Representation and recognition of handwritten digits using deformable templates, IEEE Trans. Pattern Anal. Mach. Intell. 19 (12) (1997) 1386–1391.
[3] Y. Mizukami, A handwritten Chinese character recognition system using hierarchical displacement extraction based on directional features, Pattern Recognition Lett. 19 (7) (1998) 595–604.
[4] M. Nakagawa, et al., An off-line character recognition method employing model-dependent pattern normalization by an elastic membrane model, Proceedings of the Fifth ICDAR, 1999, pp. 495–498.
[5] Y. Nakano, et al., Improvement of Chinese character recognition using projection profiles, Proceedings of the First IJCPR, 1973, pp. 172–178.
[6] M.A. Ronee, S. Uchida, H. Sakoe, Handwritten character recognition using piecewise linear two-dimensional warping, Proceedings of the Sixth ICDAR, 2001, pp. 39–43.
[7] S. Uchida, H. Sakoe, Handwritten character recognition using monotonic and continuous two-dimensional warping, Proceedings of the Fifth ICDAR, 1999, pp. 499–502.
[8] T. Wakahara, Shape matching using LAT and its application to handwritten numeral recognition, IEEE Trans. Pattern Anal. Mach. Intell. 16 (6) (1994) 618–629.
[9] T. Wakahara, Y. Kimura, A. Tomono, Affine-invariant recognition of gray-scale characters using global affine transformation correlation, IEEE Trans. Pattern Anal. Mach. Intell. 23 (4) (2001) 384–395.
[10] S. Kuo, O. Agazzi, Keyword spotting in poorly printed documents using pseudo 2-d hidden Markov models, IEEE Trans. Pattern Anal. Mach. Intell. 16 (8) (1994) 842–848.
[11] E. Levin, R. Pieraccini, Dynamic planar warping for optical character recognition, Proceedings of the ICASSP, Vol. 3, 1992, pp. 149–152.
[12] T.F. Cootes, et al., Active shape models—their training and application, Comput. Vision Image Understanding 61 (1) (1995) 38–59.
[13] A. Lanitis, C.J. Taylor, T.F. Cootes, Automatic interpretation and coding of face images using flexible models, IEEE Trans. Pattern Anal. Mach. Intell. 19 (7) (1997) 743–756.
[14] D. Shen, C. Davatzikos, An adaptive-focus deformable model using statistical and geometric information, IEEE Trans. Pattern Anal. Mach. Intell. 22 (8) (2000) 906–913.
[15] C. Naster, B. Moghaddam, A. Pentland, Flexible images: matching and recognition using learned deformations, Comput. Vision Image Understanding 65 (2) (1997) 179–191.
[16] S. Uchida, M.A. Ronee, H. Sakoe, Using eigen-deformations in handwritten character recognition, Proceedings of the 16th ICPR, Vol. 1, 2002, pp. 572–575.
[17] F. Kimura, et al., Modified quadratic discriminant functions and the application to Chinese character recognition, IEEE Trans. Pattern Anal. Mach. Intell. 9 (1) (1987) 149–153.
[18] M. Kurita, et al., Handprinted Kanji and Hirakana character recognition using weight direction index histograms and quasi-Mahalanobis distance, IEICE Japan, Technical Report, PRL82-79, 1982 (in Japanese).
[19] S. Mori, K. Yamamoto, M. Yasuda, Research on machine recognition of handprinted characters, IEEE Trans. Pattern Anal. Mach. Intell. 6 (4) (1984) 386–405.
[20] T. Saito, H. Yamada, K. Yamamoto, An analysis of handprinted Chinese characters by directional pattern matching approach, Trans. IEICE J65-D (5) (1982) 550–557 (in Japanese).
[21] S. Uchida, H. Sakoe, A monotonic and continuous two-dimensional warping based on dynamic programming, Proceedings of the 14th ICPR, Vol. 1, 1998, pp. 521–524.
[22] S. Uchida, H. Sakoe, An efficient two-dimensional warping algorithm, IEICE Trans. Inform. Systems E82-D (3) (1999) 693–700.
[23] S. Uchida, H. Sakoe, Piecewise linear two-dimensional warping, Proceedings of the 15th ICPR, Vol. 3, 2000, pp. 538–541.

About the Author—SEIICHI UCHIDA received B.E., M.E., and Dr. Eng. degrees from Kyushu University in 1990, 1992 and 1999, respectively. From 1992 to 1996, he joined SECOM Co., Ltd., Tokyo, Japan where he worked on speech processing. Now, he is an associate professor at Faculty of Information Science and Electrical Engineering, Kyushu University. His research interests include pattern analysis and speech processing. Dr. Uchida is a member of IEEE, IEICE, IPSJ, ITE, and ASJ.

About the Author—HIROAKI SAKOE received the B.E. degree from Kyushu Institute of Technology in 1966, and the M.E. and D.E. degrees from Kyushu University in 1968 and 1987, respectively. In 1968, he joined NEC Corporation and engaged in speech recognition research. In 1989, he left NEC Corporation to become a Professor of Kyushu University. His research interests include speech recognition and pictorial pattern analysis. He received 1979 IEEE ASSP Senior Award, 1980 IEICE Achievement Award and 1983 IECE Paper Award. He also received Kamura Memorial Prize from Kyushu Institute of Technology.