# A Handwritten Character Recognition Method Based on Unconstrained Elastic Matching and Eigen-Deformations

Seiichi Uchida and Hiroaki Sakoe
Faculty of Information Science and Electrical Engineering, Kyushu University,
Hakozaki 6-10-1 Fukuoka, 812-8581, Japan
E-mail: {uchida, sakoe}@is.kyushu-u.ac.jp

## Abstract

*A fast elastic matching based handwritten character recognition method is investigated. In the present method, an unconstrained elastic matching technique, where the matching is optimized locally and individually on each pixel, is utilized together with its a posteriori evaluation based on the eigen-deformations of handwritten characters. Our experimental results show that high recognition rates can be attained by the present method with feasible computations.*

## 1. Introduction

Elastic matching, or deformable template is a promising technique for handwritten character recognition because it ideally provides a deformation-invariant distance or similarity between input and reference patterns. The central problem of elastic matching is the optimization of pixel-to-pixel mapping between those two patterns. Thus, the characteristics of elastic matching, such as flexibility and computational complexity, depend on the formulation and the optimization strategy of the mapping.

The purpose of this paper is to present a practical handwritten character recognition method based on a fast elastic matching technique, hereafter called *unconstrained elastic matching*. The unconstrained elastic matching is based on a local and individual (i.e., unconstrained) optimization strategy of the mapping. Specifically, for each pixel on the input pattern, its corresponding pixel on the reference pattern is searched individually within its neighborhood. This optimization strategy provides the superiority of the unconstrained elastic matching in computational efficiency.

Although several unconstrained elastic matching techniques (sometimes referred to as the shift similarity method [7]) were already proposed in the dawn of elastic matching based character recognition [6, 15], they have been substituted by less flexible and often slower elastic matching techniques where the mapping is two-dimensionally constrained (e.g., [4, 5, 8, 10, 11, 12]) or modeled by a two-dimensional (parametric) functions (e.g., [2, 16]). This is probably because the unconstrained elastic matching has the *overfitting* problem that the distance between an input pattern and an incorrect reference pattern is underestimated and thus increasing recognition errors. While any elastic matching technique has the overfitting problem to a certain extent, the unconstrained elastic matching has the most serious and inevitable one due to its local and individual optimization strategy.

In order to revive the unconstrained elastic matching by relieving the overfitting problem, we utilize the fact that there exist intrinsic deformations within each character category. For example, in category “A”, global skews and vertical shifts of its horizontal stroke are often observed. Hereafter, such intrinsic deformations are called *eigen-deformations*. It can be considered that any overfitting results from a mapping deviated from the eigen-deformations. For example, the mapping which fits “A” to “R” will not be expressed by a combination of the eigen-deformations of “A”.

The present recognition method is illustrated in **Fig.1**. The method consists of two processes: a training process and a recognition process. The eigen-deformations estimated in the training process using the unconstrained elastic matching are utilized in the recognition process to evaluate the result of the same unconstrained elastic matching between input and reference patterns. Note that the method of **Fig.1** is general in the sense that any other elastic matching technique can be employed. For example, for higher recognition rates at the cost of speed, one can choose a constrained or parametric elastic matching techniques instead. (The result of a recognition experiment with a constrained elastic matching technique will be precisely reported in [14].)

The idea of the eigen-deformations is motivated by Cootes et al.[1] and Naster et al.[9] but different from their

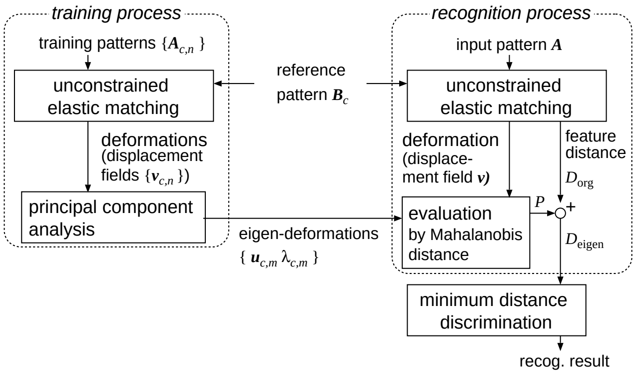


**Figure 1. Diagram of the present recognition method.**

work in (i) applying the eigen-deformations for a character recognition task, and (ii) emphasizing the dependency between the employed elastic matching and the resulting eigen-deformations. Especially for (ii), our experiment will show that the eigen-deformations estimated from the unconstrained matching are not compatible with those from the constrained matching.

This paper is organized as follows. In Section 2, the unconstrained elastic matching and its overfitting problem are described. In Section 3, the details of the present recognition method of **Fig. 1** are given. Finally, experimental results are provided in Section 4 to show that the present recognition method using the fast unconstrained elastic matching and the eigen-deformations can attain high recognition rates with feasible computations.

## 2. Unconstrained Elastic Matching

Elastic matching between two $I \times I$ images $\boldsymbol{A} = \{\boldsymbol{a}(i,j)\}$ and $\boldsymbol{B} = \{\boldsymbol{b}(x,y)\}$ is generally formulated as the following minimization problem

$$D_{\rm org}(\boldsymbol{A},\boldsymbol{B}) = \min_{\{(x_{i,j},y_{i,j})\}} \sum_{i=1}^{I}\sum_{j=1}^{I} \delta(\boldsymbol{a}(i,j),\boldsymbol{b}(x_{i,j},y_{i,j})) \tag{1}$$

where variables $(x_{i,j}, y_{i,j})$ are the mapping of the pixel $(i,j)$ and $\delta(\cdot,\cdot)$ is a distance function between two feature vectors. The value $D_{\rm org}(\boldsymbol{A},\boldsymbol{B})$ can be considered as a deformation invariant distance between $\boldsymbol{A}$ and $\boldsymbol{B}$ and thus can be immediately used as a discriminant function in character recognition tasks.

Although elastic matching techniques usually employ several constraints on the mapping to control the flexibility, the unconstrained elastic matching technique employs no constraints except for boundary constraints and the following maximum displacement limitation

$$|x_{i,j} - i| \le w, \qquad |y_{i,j} - j| \le w, \tag{2}$$

where $w$ is a positive integer specifying maximum displacement. Since these constraints are unary, that is, do not limit the relation between two or more pixel mappings, the minimization problem of (1) can be decomposed into $I^2$ individual subproblems, i.e.,

$$D_{\rm org}(\boldsymbol{A},\boldsymbol{B}) = \sum_{i=1}^{I}\sum_{j=1}^{I}\left[\min_{(x_{i,j},y_{i,j})} \delta(\boldsymbol{a}(i,j),\boldsymbol{b}(x_{i,j},y_{i,j}))\right]. \tag{3}$$

The computational complexity for the unconstrained matching is $O(I^2w^2)$ since each of $I^2$ subproblems requires search efforts of $O(w^2)$. Thus, the unconstrained elastic matching is far faster than other fully two-dimensional elastic matching techniques [5, 12] which require computations exponentially increasing with $I$. Furthermore, this complexity is comparable to or also lower than those of pseudo two-dimensional elastic matching techniques [4, 8, 11].

In spite of its computational efficiency and sufficient flexibility, the unconstrained elastic matching is rarely employed recently in handwritten character recognition. This is probably because of its overfitting problem. For example, the unconstrained elastic matching may closely fit characters with topologically different shapes, such as “A” and “H”. In addition, the unconstrained elastic matching often undergoes a serious overfitting where a stroke is partially or entirely skipped by the gap (i.e., discontinuity) in the mapping. For example, a character of “L” may be easily fitted to “U” by skipping the right vertical part of “U”.

## 3. Recognition Using Eigen-Deformations

In this section, the training process and the recognition process of the present recognition method shown in **Fig.1** are described in detail. The aim of the present method is the relief of the overfitting problem in the unconstrained elastic matching by newly incorporating *a posteriori* evaluation based on the eigen-deformations.

### 3.1. Training process

The training process where the eigen-deformations are estimated in a statistical way is decomposed into two parts: automatic collection of the deformations of handwritten characters using the unconstrained elastic matching, and principal component analysis (PCA) of the collected deformations to estimate the eigen-deformations.

Let $\{\boldsymbol{A}_{c,n} \mid n = 1,\ldots,N\}$ denote the set of training patterns of category $c \in \{1,\ldots,C\}$ and $\boldsymbol{B}_c$ denote the reference pattern of the same category $c$ created by simply

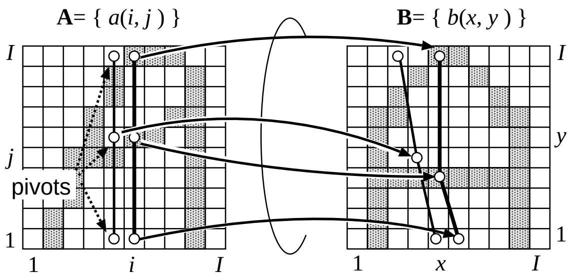


**Figure 2. Piecewise linearization of elastic matching.**

averaging several samples. The deformation in $\boldsymbol{A}_{c,n}$ 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 obtained as the result of the unconstrained elastic matching between $\boldsymbol{A}_{c,n}$ and $\boldsymbol{B}_c$. Hereafter, this vector $\boldsymbol{v}_{c,n}$ is called the *displacement field* and its dimension is denoted as $M$. (Here, $M = 2I^2$.) Thus, applying the unconstrained elastic matching to $N$ training patterns and one reference pattern, $N$ displacement fields are automatically collected as $\boldsymbol{V}_c = \{\boldsymbol{v}_{c,1}, \ldots, \boldsymbol{v}_{c,n}, \ldots, \boldsymbol{v}_{c,N}\}$.

The eigen-deformations of category $c$ are defined as the principal components of the distribution of displacement fields in $\boldsymbol{V}_c$. Letting $\boldsymbol{\Sigma}_c$ be the covariance matrix of the displacement fields in $\boldsymbol{V}_c$, the principal components of $\boldsymbol{V}_c$, i.e., the eigen-deformations of category $c$ are the $M$-dimensional eigenvectors of $\boldsymbol{\Sigma}_c$ and denoted as $\{\boldsymbol{u}_{c,1}, \ldots, \boldsymbol{u}_{c,m}, \ldots, \boldsymbol{u}_{c,M}\}$. 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.

### 3.2. Recognition process

In the recognition process of the present method, the following distance is used as the discrimination function of unknown input $\boldsymbol{A}$ instead of the direct use of $D_{\text{org}}(\boldsymbol{A}, \boldsymbol{B}_c)$,

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

where $\alpha$ is a constant ($0 \leq \alpha \leq 1$). The term $P(\boldsymbol{A}, \boldsymbol{B}_c)$ is *a posteriori* evaluation term of the resulting displacement field $\boldsymbol{v}$ of the unconstrained elastic matching to obtain $D_{\text{org}}(\boldsymbol{A}, \boldsymbol{B}_c)$, and defined as the Mahalanobis distance, i.e.,

$$\begin{aligned} P(\boldsymbol{A}, \boldsymbol{B}_c) &= (\boldsymbol{v} - \overline{\boldsymbol{v}}_c)^T \Sigma_c^{-1} (\boldsymbol{v} - \overline{\boldsymbol{v}}_c) \\ &= \sum_{m=1}^{M} \frac{\langle \boldsymbol{v} - \overline{\boldsymbol{v}}_c, \boldsymbol{u}_{c,m} \rangle^2}{\lambda_{c,m}}, \end{aligned} \quad (5)$$

where $\overline{\boldsymbol{v}}_c$ is the average displacement. In case of overfitting, since the displacement field $\boldsymbol{v}$ will not agree with the eigen-deformations $\{\boldsymbol{u}_{c,m}\}$, the distance $P(\boldsymbol{A}, \boldsymbol{B}_c)$ becomes large. Thus, $D_{\text{eigen}}(\boldsymbol{A}, \boldsymbol{B}_c)$ becomes also large and $\boldsymbol{A}$ will not be discriminated into incorrect category $c$.

It is well known that the numerical and the estimation errors of higher-order eigenvalues are amplified in (5). Thus, in our experiment, a modified Mahalanobis distance [3] was employed where the higher-order eigenvalues $\lambda_{c,m}$ ($m = M' + 1, \ldots, M$) are replaced by $\lambda_{c,M'+1}$. In the following experiments, the parameter $M'$ was manually optimized as well as $\alpha$.

## 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 [17] were prepared. Firstly, four-dimensional directional features [7] together with one-dimensional gray-level feature were extracted as pixel features. Secondly, the character region of each sample was linearly scaled to $16 \times 16$ and added pixel margins to be $I = 20$. Finally, for each category $c$, one reference pattern $\boldsymbol{B}_c$ was created by simply averaging the first 100 samples. The next 500 samples were used as training samples $\{\boldsymbol{A}_{c,n}\}$ in the eigen-deformation estimation of **4.5**. The remaining 500 samples were used as test samples $\boldsymbol{A}$ in the recognition experiment of **4.6**.

According to the above feature extraction procedure, the distance $\delta$ in (3) was defined as a weighted L1 distance between two five-dimensional vectors, i.e.,

$$\begin{aligned} \delta(\boldsymbol{a}(i,j), \boldsymbol{b}(x,y)) &= |a_I(i,j) - b_I(x,y)| \\ &\quad + \eta \sum_{d=1}^{4} |a_d(i,j) - b_d(x,y)| \end{aligned}$$

where subscripts "$I$" and "$d$" denote the intensity feature and the four directional features, respectively. The weighting coefficient $\eta$ was fixed at $0.4$ according to a result of an preliminary experiment.

### 4.2. Piecewise linearization of matching

In our experiments, a piecewisely linearized version of the unconstrained elastic matching was used. As shown in **Fig.2**, 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. In the experiments, three pivots were regularly placed at $j = 1$(bottom), $10$(center), and $20$(top) on each column.

This linearization is meaningful to obtain reliable eigen-deformations from a limited number of the training patterns

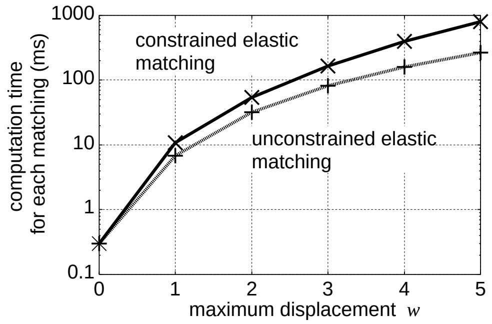


**Figure 3. Computation time versus maximum displacement $w$ on PC (Pentium III, 750MHz).**

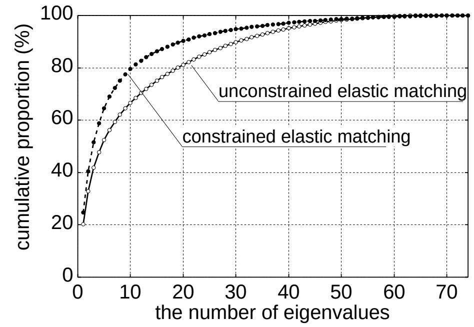


**Figure 5. Cumulative proportions of eigen-deformations.**

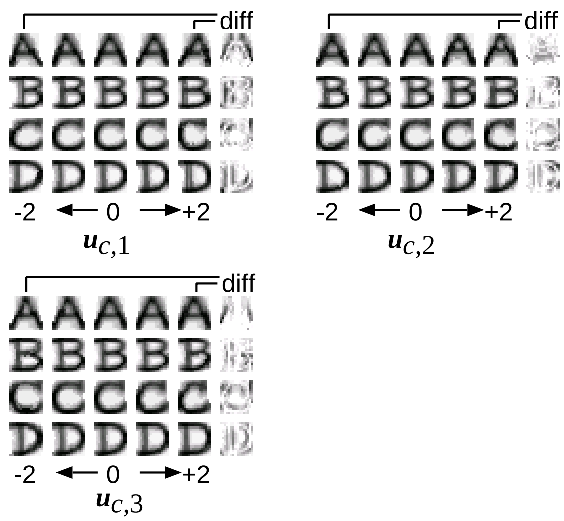


**Figure 4. First three eigen-deformations $\boldsymbol{u}_{c,1}, \boldsymbol{u}_{c,2}$, and $\boldsymbol{u}_{c,3}$ multiplied by $k\sqrt{\lambda_{c,m}}$, $k = -2, -1, 0, 1, 2$.**

since the mappings of non-pivot pixels are linearly dependent on pivot mappings and thus the dimension $M$ can be reduced from $O(I^2)$ to $O(I)$. Specifically, in our case, $M$ is reduced from $800$ to $74$ (where the boundary conditions are also considered). Furthermore, this linearization is meaningful to improve the fundamental recognition accuracy of the unconstrained elastic matching, since the chance of overfitting in the vertical direction can be reduced. In fact, the recognition rates without the linearization were far below 90% at $w \geq 2$, whereas those with the linearization were constantly over 90% (as we shall see in **4.6**).

As for computational complexity, the linearization slightly increases the computations for the matching from $O(I^2w^2)$ to $O(I^2w^3)$ but decreases the computations for the evaluation term $P(\boldsymbol{A}, \boldsymbol{B}_c)$ from $O(I^4)$ to $O(I^2)$.

### 4.3. Constrained elastic matching for comparison

As noted in Section 1, the unconstrained elastic matching can be replaced by any other elastic matching techniques in the present method. In the following experiments, the *constrained* elastic matching technique proposed in [10, 13] was employed for comparison. This constrained elastic matching is perfectly the same as the above linearized unconstrained elastic matching except for several constraints on the mapping of adjacent pivots to preserve the topological structure (continuity and monotonicity) in the horizontal direction.

### 4.4. Comparison of computation time

Theoretically, the unconstrained elastic matching is faster than the constrained one because their computational complexities are $O(I^2w^3)$ and $O(I^2w^4)$, respectively. This superiority is confirmed by the experimental result of **Fig.3** where the matching time on a PC (Pentium III, 750MHz) is shown as a function of maximum displacement $w$. The unconstrained matching is always faster than the constrained one and their difference becomes larger according to $w$. Note that this superiority will become overwhelming one when more pivots are used. In the extreme case of $I$ pivots on each column (i.e., no piecewise linearization), the complexities of the unconstrained and the constrained matching techniques are $O(I^2w^2)$ and $O(I^3 9^I)$, respectively.

### 4.5. Estimation of eigen-deformation

The eigen-deformations was experimentally estimated according to the process described in **3.1**. For each category $c$ and for each $w \in \{1, 2, \ldots, 5\}$, the piecewise linear unconstrained elastic matching was performed between one reference pattern ($\boldsymbol{B}_c$) and each of 500 training patterns ($\boldsymbol{A}_{c,n}$) and the resulting 500 displacement fields ($\boldsymbol{v}_{c,n}$) were subjected to PCA to obtain 74 eigen-deformations

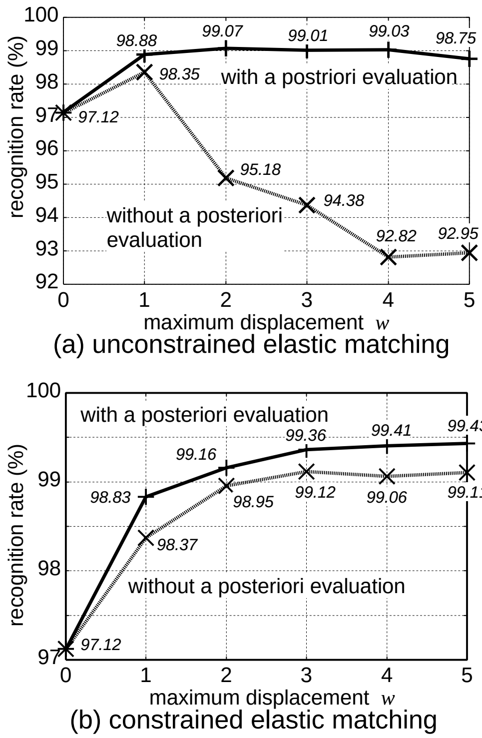


**Figure 6. Recognition rates for 500× 26 test samples.**

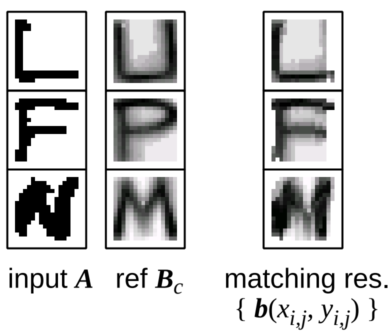


**Figure 7. Three input samples $A$ misrecognized to $B_c$ by the conventional distance $D_{\mathrm{org}}(\boldsymbol{A}, \boldsymbol{B}_c)$. These samples were successfully recognized using the present distance $D_{\mathrm{eign}}(\boldsymbol{A}, \boldsymbol{B}_c)$.**

$\boldsymbol{u}_{c,m}$. **Figure 4** shows the several reference patterns deformed by estimated eigen-deformations at $w = 2$. For example, the skew transformation was extracted as the first eigen-deformation of "A".

**Figure 5** shows the cumulative proportions of eigen-deformations obtained by the unconstrained and the constrained elastic matching techniques, where each value is the average of 26 categories. This result indicates that in both techniques the cumulative proportion is saturated with a small number of eigen-deformations and thus there are certain deformation tendencies in handwritten characters. From this result, it can be also observed that there is a difference between two techniques in the speed of the saturation. Since a slow saturation means that the displacement fields are scattered in the $M$ dimensional space, this difference indicates that the unconstrained elastic matching technique is more unstable than the constrained elastic matching.

### 4.6. Recognition results

Recognition experiments were conducted using 500 test samples for each category. **Figure 6** (a) shows the recognition rates attained by the linearized unconstrained elastic matching with or without *a posteriori* evaluation using the above eigen-deformations. The recognition rates without the *a posteriori* evaluation (i.e., the discrimination based on $D_{\mathrm{org}}(\boldsymbol{A}, \boldsymbol{B}_c)$) were drastically degraded along with the increase of $w$. Especially for $w \geq 2$, the recognition errors of "L" $\rightarrow$ "U" were very serious and the recognition rates were lower than that of the rigid matching ($w = 0$). As noted in Section 2, these recognition errors were due to the overfitting of the unconstrained elastic matching. Other recognition errors over 10 samples at $w = 2$ were "F"→ "P" and "N"→ "M". **Figure 7** shows three input samples misrecognized without the *a posteriori* evaluation. Their matching results also shown in this figure clearly express the overfitting to incorrect references.

On the other hand, using the unconstrained elastic matching with the *a posteriori* evaluation (i.e., the discrimination based on $D_{\mathrm{eigen}}(\boldsymbol{A}, \boldsymbol{B}_c)$), recognition rates were significantly improved and constantly around 99%. The input samples of **Fig.7** were correctly recognized. There was no conspicuous recognition error over 10 samples at all $w$ except for 16 samples of "L" $\rightarrow$ "U" at $w = 5$. It is worth to note that the recognition rates were not highly sensitive to the parameter $\alpha$ of (4) since their changes with respect to $\alpha$ were unimodal for all $w$.

The effect of the *a posteriori* evaluation on the reduction of the overfitting was confirmed in another way. **Figure 8** shows the effect of elastic matching on the positional relations between input and reference patterns. Among these four cases, the case (d) where the input pattern is correctly recognized without elastic matching and misrecognized with elastic matching is most obviously expressing the overfitting. In our experimental result at $w = 2$ of **Fig.6**(a), the number of samples falling into (d) was reduced from $504$ to $43$ by *a posteriori* evaluation. Moreover, the ratio of the recognition errors in (d) to the entire recognition errors (i.e., (d)/ $((\mathrm{c}) + (\mathrm{d}))$ ) was reduced from 80% to 36%.

**Figure 6**(b) shows the recognition rates attained by

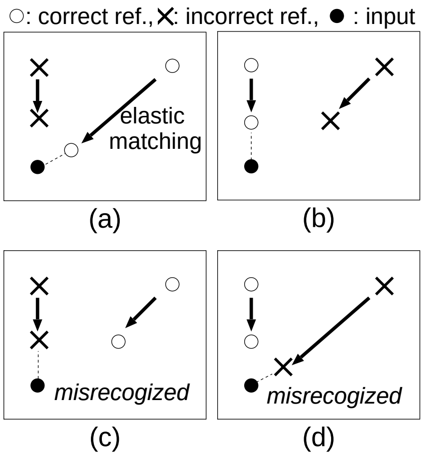


**Figure 8. The effect of elastic matching on the positional relations between input and reference patterns.**

the *constrained* elastic matching. In this case, the eigen-deformations were estimated also using the same constrained elastic matching. The comparison between **Fig.6**(a) and (b) shows that the unconstrained matching with the *a posteriori* evaluation can attain recognition rates comparable to the constrained matching, and the present method is a reasonably practical technique, considering the superiority in computational efficiency.

Another experiment was conducted where the eigen-deformations estimated from the *constrained* matching were used in the recognition process with the *unconstrained* matching. The resulting recognition rates were unstable and always lower than those with the eigen-deformations from the unconstrained matching. Especially at $w = 2$, which is the worst case among $w \in \{1, 2, \ldots, 5\}$, the recognition rate was about 95% and this means that *a posteriori* evaluation could not improve the recognition rate. Thus, this result indicates that the eigen-deformations from the unconstrained matching are not compatible with those from the constrained matching.

## 5. Conclusion

A practical handwritten character recognition method based on a fast and classical elastic matching technique, called *unconstrained elastic matching*, was presented. Although the unconstrained elastic matching has been substituted by slower constrained ones due to its serious overfitting problem, our experimental results encourage its revival. Specifically, the results showed that the unconstrained elastic matching together with *a posteriori* evaluation based on eigen-deformations can attain the recognition rates comparable to the constrained one, with less computations.

## References


[1] T.F.Cootes, et al.,"Active shape models - their training and application," *Comput. Vis. Image Und.*, vol. 61, no. 1, pp.38-59, 1995.

[2] A.K.Jain and D.Zongker, "Representation and recognition of handwritten digits using deformable templates," *IEEE Trans. PAMI*, vol. 19, no. 12, pp. 1386-1391, 1997.

[3] F.Kimura, et al., "Modified quadratic discriminant functions and the application to Chinese character recognition," *IEEE Trans. PAMI*, vol. 9, no. 1, pp. 149-153, 1987.

[4] S.Kuo and O.Agazzi, "Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models," *IEEE Trans. PAMI*, vol. 16, no. 8, pp. 842-848, 1994.

[5] E.Levin and R.Pieraccini, "Dynamic planar warping for optical character recognition," *Proc. ICASSP*, vol. 3, pp. 149-152, 1992.

[6] S.Meguro and M.Umeda, "An extraction of shape deviations in handwritten characters in hierarchical pattern matching," *Tech. Rep. Inst. Electron. & Commun. Eng. Jpn.*, PRL77-70, 1978. (in Japanese)

[7] S.Mori, et al., "Research on machine recognition of handprinted characters," *IEEE Trans. PAMI*, vol. 6, no. 4, pp. 386–405, 1984.

[8] Y.Nakano, et al., "Improvement of Chinese character recognition using projection profiles," *Proc. IJCPR*, pp. 172-178, 1973.

[9] C.Naster, et al., "Flexible images:matching and recognition using learned deformations" *Comput. Vis. Image Und.*, vol.65, no.2, pp.179-191, 1997.

[10] M.A.Ronee, et al., "Handwritten character recognition using piecewise linear two-dimensional warping," *Proc. ICDAR*, pp. 39-43, 2001.

[11] J.Tsukumo, "Handprinted Kanji character recognition based on flexible template matching," *Proc. ICPR*, pp. 483-486, 1992.

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

[13] S.Uchida and H.Sakoe, "Piecewise linear two-dimensional warping," *Proc. ICPR*, vol. 3, pp. 538-541, 2000.

[14] S.Uchida and H.Sakoe, "Using eigen-deformations in handwritten character recognition," to be published in *Proc. ICPR*, 2002.

[15] H.Yamada, et al.,"An improvement of correlation method — locally maximized correlation —," *Trans. Inst. Electron. & Commun. Eng. Jpn.*, vol. J64-D, no. 10, pp. 970-976, 1981. (in Japanese)

[16] T.Wakahara, et al., "Affine-invariant recognition of gray-scale characters using global affine transformation correlation," *IEEE Trans. PAMI*, vol. 23, no. 4, pp. 384-395, 2001.

[17] `http://www.etl.go.jp/~etlcdb/#English`