

# A Survey of Elastic Matching Techniques for Handwritten Character Recognition

Seiichi UCHIDA<sup>†a)</sup>, *Member* and Hiroaki SAKOE<sup>†</sup>, *Fellow*

**SUMMARY** This paper presents a survey of elastic matching (EM) techniques employed in handwritten character recognition. EM is often called deformable template, flexible matching, or nonlinear template matching, and defined as the optimization problem of two-dimensional warping (2DW) which specifies the pixel-to-pixel correspondence between two subjected character image patterns. The pattern distance evaluated under optimized 2DW is invariant to a certain range of geometric deformations. Thus, by using the EM distance as a discriminant function, recognition systems robust to the deformations of handwritten characters can be realized. In this paper, EM techniques are classified according to the type of 2DW and the properties of each class are outlined. Several topics around EM, such as the category-dependent deformation tendency of handwritten characters, are also discussed.

**key words:** *elastic matching, handwritten character recognition, deformation, optimization, survey*

## 1. Introduction

One of the main problems of handwritten character recognition is how to deal with geometric deformations of characters. In [37, p. 106], the geometric deformations of handwritten characters are classified into the following four types:

- fluctuation of stroke thickness due to noise and inappropriate binarization,
- linear deformations, such as translation, scaling, shear, and rotation,
- nonlinear and topology-preserving deformations, such as the deviation from original geometric balance, and
- deformations changing topology, such as the disappearance of loops.

The use of features invariant to those deformations will be a popular solution of the problem. For example, the horizontal projection profile [34] is a classical feature invariant to horizontal shifts of character patterns. The use of shape normalization techniques is another solution [21], [24]. For example, linear scaling of the bounding box of a character is the simplest normalization technique and enough to realize scale-invariant recognition. Line density equalization proposed by Yamada et al. [65] is a nonlinear shape normalization technique and can adjust nonlinear deformations.

An alternative to those solutions is *elastic matching*

Manuscript received October 9, 2004.

<sup>†</sup>The authors are with the Faculty of Information Science and Electrical Engineering, Kyushu University, Fukuoka-shi, 812-8581 Japan.

a) E-mail: uchida@is.kyushu-u.ac.jp  
DOI: 10.1093/ietisy/e88-d.8.1781

(EM), which is also called deformable template [50], flexible matching [32], or nonlinear template matching [31]. EM has been employed in not only handwritten character recognition but also many other image pattern matching problems, such as face recognition, fingerprint recognition, gesture recognition, medical image analysis, automatic image morphing, computer vision (e.g., stereo), and motion analysis.

In EM, one character image  $A$  is treated like a “rubber sheet” and fitted to another character image  $B$  as close as possible. Formally speaking, EM is defined as an optimization problem with respect to a linear or nonlinear pixel-to-pixel mapping, called two-dimensional warping (2DW), from  $A$  to  $B$ . The distance evaluated under the optimized 2DW is invariant to the deformations compensable by 2DW. Thus, by using the distance as a discriminant function, recognition systems robust to the deformations of handwritten characters can be realized.

The advantages of EM over invariant features and normalization techniques are as follows:

- EM is adaptive and thus generally possesses higher ability to compensate various deformations than invariant features and normalization techniques.
- The optimized 2DW itself describes the deformation of subjected character. This fact shows that EM possesses useful properties of structural analysis techniques.
- EM can be linked to statistical and stochastic frameworks. Active shape models [6], [45], [56] and (pseudo-)2D HMMs [1], [20], [22], [38] are two good examples.

The characteristics of EM mainly depend on two factors: (i) the formulation of 2DW, and (ii) the optimization strategy of 2DW. The first factor affects the range of compensable deformations. Thus, 2DW should be formulated by considering the deformation characteristics of handwritten characters. For example, when we can assume that handwritten characters mainly undergo rubber-sheet-like deformations, topology-preserving 2DW is a natural choice. The second factor affects the accuracies of the results of EM, namely, the accuracies of the minimized distance and the optimized 2DW. Generally speaking, optimization strategies for globally optimal solutions will provide more accurate results than those for sub-optimal solutions. Note that these two factors mutually affect the computational complexity of EM.

The purpose of this paper is to overview various EM

techniques employed in handwritten character recognition and to grasp unsolved problems around EM. In Sect. 2, EM is outlined by its mathematical formulation and general properties. In Sect. 3, EM techniques are classified according to the formulation of 2DW. The properties of each class, such as computational complexity and other practical issues, are also discussed in this section. In Sect. 4, several topics around EM are discussed. Some of those topics have not been well investigated in previous researches. Finally, in Sect. 5 conclusions are presented along with future research topics.

Note that, throughout this paper, we mainly concern the EM techniques for two “planer” patterns; other types of EM developed for handwritten character recognition are beyond the scope of this paper. Thus, EM between a 1D pattern and a planer pattern (e.g., [40], [44]) and EM between two 1D patterns (e.g., contour matching proposed in [63], [66] and thinned pattern matching proposed in [7]) are not included in this survey.

## 2. Outline of EM

### 2.1 Problem Formulation of EM

Consider two  $N \times N$  handwritten character image patterns  $\mathbf{A} = \{a_{i,j}\}$  and  $\mathbf{B} = \{b_{x,y}\}$ , where  $a_{i,j}$  and  $b_{x,y}$  are pixel feature vectors at pixel  $(i, j)$  on  $\mathbf{A}$  and  $(x, y)$  on  $\mathbf{B}$ , respectively. Let  $F$  denote a 2D-2D mapping from  $\mathbf{A}$  to  $\mathbf{B}$ , i.e.,  $F : (i, j) \mapsto (x, y)$  (Fig. 1).

EM is formulated as the minimization problem of the following objective function with respect to  $F$ :

$$J_{A,B}(F) = D(\mathbf{A}, \mathbf{B}_F), \quad (1)$$

where  $D(\cdot, \cdot)$  is a simple “rigid” distance metric (e.g., Euclidean distance, or absolute distance) between two character image patterns and  $\mathbf{B}_F$  is the character pattern obtained by fitting  $\mathbf{B}$  to  $\mathbf{A}$ , i.e.,  $\mathbf{B}_F = \{b_{F(i,j)}\}$ .

Let  $\tilde{F}$  denote the 2DW  $F$  minimizing  $J_{A,B}(F)$  of (1). Then, EM distance  $D_{EM}(\mathbf{A}, \mathbf{B})$  is defined as

$$D_{EM}(\mathbf{A}, \mathbf{B}) = \min_F J_{A,B}(F) = J_{A,B}(\tilde{F}). \quad (2)$$

Clearly,  $D_{EM}(\mathbf{A}, \mathbf{B}) = D(\mathbf{A}, \mathbf{B}_{\tilde{F}})$ . Thus, EM distance is the rigid distance between  $\mathbf{A}$  and  $\mathbf{B}_{\tilde{F}}$ , i.e., the rigid distance between  $\mathbf{A}$  and  $\mathbf{B}$  after fitting  $\mathbf{B}$  to  $\mathbf{A}$  as close as possible.

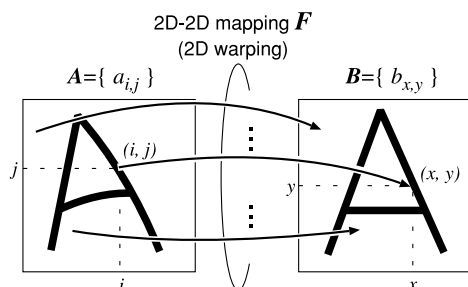


Fig. 1 2DW defined between two handwritten character images.

### 2.2 Properties of EM Distance

The EM distance (2) is different from rigid distance metrics at the following three properties.

(a) EM distance is deformation-invariant.

The EM distance  $D_{EM}(\mathbf{A}, \mathbf{B})$  is invariant to the deformations which are compensable by  $F$ . For example, if the 2DW  $F$  is defined as affine transformation,  $D_{EM}(\mathbf{A}, \mathbf{B})$  is theoretically invariant to any affine transformed version of  $\mathbf{B}$ . This property suggests that EM-based recognizers are robust to geometric deformations within a category and thus expected to attain better recognition rates than rigid distance-based recognizers.

Unfortunately, this property also suggests that similar-shaped characters of different categories may be misrecognized by EM-based recognizers. For example, if  $F$  is affine transformation,  $D_{EM}(\text{“9”}, \text{“6”})$  is, theoretically, very close to  $D_{EM}(\text{“9”}, \text{“9”})$  because affine transformation can compensate 180° rotation. Thus, the input “9” may be misrecognized to the different category “6”. This is the phenomenon called *overfitting*. Generally speaking, there is a trade-off between the ability of compensating deformations within a category and the risk of overfitting.

(b) EM distance is anisotropic.

The conventional Euclidean distance is isotropic in the pattern space and the set of patterns equidistant from a certain pattern form hypersphere in the pattern space. In contrast, the EM distance  $D_{EM}(\mathbf{A}, \mathbf{B})$  is generally anisotropic and the set of patterns equidistant from a certain pattern do not form hypersphere.

This fact can be confirmed from the experimental result of Fig. 2 [26]. The small dots are actual handwritten numeral patterns  $\{A_k\}$  of a category (displayed in the 2D subspace spanned by their first two principal axes), and the black triangle is their centroid  $\mathbf{R}$ . When Euclidean distance is used, the centroid  $\mathbf{R}$ , which minimizes  $\sum_k D(A_k, \mathbf{R})$ , is placed around the center of pattern distribution (Fig. 2 (a)). In contrast, when the EM distance is used, the centroid  $\mathbf{R}$ , which minimizes  $\sum_k D_{EM}(A_k, \mathbf{R})$ , is not placed around the center (Fig. 2 (b)).

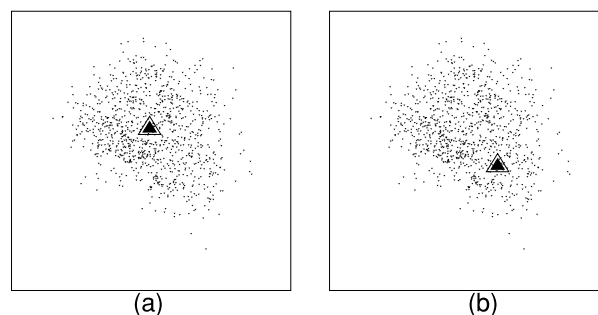


Fig. 2 The centroids of handwritten character patterns under (a) Euclidean distance and (b) EM distance. The character patterns and the centroids are displayed in 2D subspace.

(c) EM distance is often asymmetric.

The EM distance defined by (2) is asymmetric, i.e.,  $D_{EM}(A, B) \neq D_{EM}(B, A)$ , because only  $B$  is deformed by  $F$  for the optimal matching between  $A$  and  $B$ . Thus, the EM distance (2) is, formally, not a distance metric. This asymmetric property, however, is not crucial; in many recognition tasks, the EM distance (2) is successfully used as a discriminant function.

There are several ways to obtain a symmetric EM distance. Among them, the sum of two asymmetric distances,  $D_{EM}(A, B) + D_{EM}(B, A)$ , is the simplest one. Bi-directional EM where not only  $B$  but also  $A$  are deformed by 2DW is another and a more elaborated way.

### 2.3 Comparison with Shape Normalization

Shape normalization, such as line density equalization [65], is closely related to EM in the sense that both of shape normalization and EM have the same purpose of providing a deformation-invariant distance. They, however, differ at the ability of compensating deformations. As shown in Fig. 3 (a), EM shifts  $B$  to a pattern close to  $A$ . In contrast, as shown in Fig. 3 (b), shape normalization shifts  $A$  to a pattern having some ideal property, and shifts  $B$  independently to another pattern having the same property. Thus, EM is potentially more powerful in compensating deformations. Unfortunately, this fact does not guarantee better performance of EM over shape normalization; as noted in Sect. 2.2 (a), this fact also indicates that EM has more risks of providing

underestimated distance between two character patterns of different categories, namely, EM has more risks of overfitting.

Note that EM and shape normalization are not rival. That is, they can be utilized in a collaborative manner. In fact, a simpler 2DW is enough if reasonable shape normalization is available. Tsukumo's EM technique [51] is a very good example where blurring normalization and a simple 2DW are successfully combined.

### 3. Classification of EM Techniques

In this section, EM techniques for handwritten character recognition are classified according to the formulation of 2DW, which is one of the two factors determining the characteristics of EM (as noted in Sect. 1). Figure 4 shows a classification tree. As shown in this figure, the EM techniques can be roughly divided into two classes, i.e., parametric 2DW-based EM and non-parametric 2DW-based EM. In non-parametric 2DW, each variable which controls  $F$  directly represents pixel correspondence. In parametric 2DW, each variable does not represent pixel correspondence but represents a parameter that controls 2DW indirectly.

Each of those two classes is further divided into several classes. As shown in Fig. 4, each class is closely related to several optimization strategies. That is, if a 2DW of a class is chosen, possible optimization strategy is almost determined.

#### 3.1 Parametric 2DW

##### 3.1.1 Linear 2DW

Most of parametric EM techniques for handwritten character recognition assume that the geometric deformations of handwritten characters can be described by some linear transformations. Specifically, the 2DW  $F : (i, j) \mapsto (x, y)$  are formulated as a linear function,

$$(x, y) = (\alpha_1 i + \alpha_2 j + \alpha_3, \alpha_4 i + \alpha_5 j + \alpha_6), \tag{3}$$

where  $\alpha_1, \alpha_2, \dots, \alpha_6$  are real-valued parameters that control

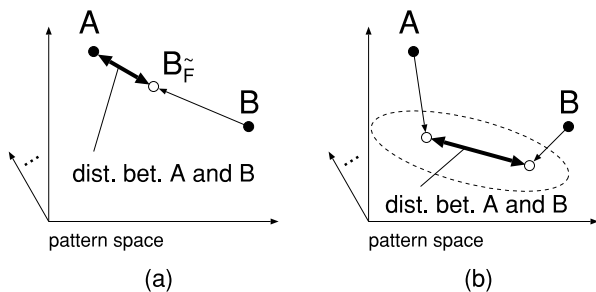


Fig. 3 Distance between  $A$  and  $B$  given by EM (a) and normalization (b).

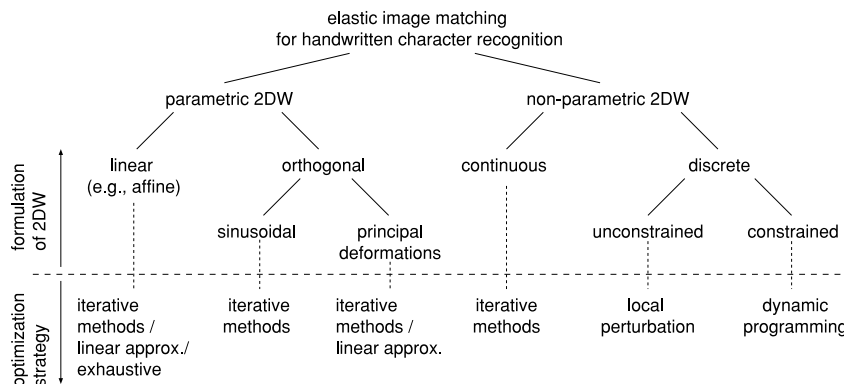


Fig. 4 Classification of EM techniques employed in handwritten character recognition.

**F**. If those six parameters are independent, **F** is affine transformation. Note that  $x$  and  $y$  provided by (3) are generally not integer and therefore several interpolation techniques should be employed to obtain the pixel feature  $\mathbf{b}_{x,y}$ .

Even if 2DW is defined as a linear transformation, its optimization problem becomes nonlinear, since the parameters to be optimized are involved in a nonlinear image pattern function **B**. Thus, the optimization problem is often tackled with iterative solutions or approximate solutions instead of deterministic solutions.

Wakahara and his colleagues [58], [59] have proposed affine transformation-based 2DW techniques, called GAT, for handwritten character recognition. In GAT, 2DW is described by a single global affine transformation. The optimization problem of GAT is approximated as a linear problem by fixing the parameters in the nonlinear part of an objective function at constant values. This approximated problem can be solved by the successive iteration method.

The tangent distance method [15], [47] has been proposed as another linear 2DW-based EM<sup>†</sup> where a nonlinear optimization problem of linear 2DW is approximated as a linear problem by Taylor series expansion. Recently, this idea is successfully linked with statistic framework [17].

There have been several trials for optimizing linear 2DW in an exhaustive manner. Yasuda et al. [67] proposed the perturbed correlation method, where a 2D reference pattern is “perturbed” by affine transformation with discretized parameter values. Each of perturbed patterns is rigidly matched with a 2D input pattern. Since the number of possible parameter values becomes very large, this method requires numerous and repetitive 2D-2D rigid matchings. Recent hardware, however, makes the method computationally tractable one. A similar method can be found in [8].

### 3.1.2 Orthogonal 2DW

In several EM techniques, 2DW is represented as a linear combination of orthogonal functions, i.e.,

$$(x, y) = \sum_{k=1}^K \alpha_k \phi_k(i, j) \quad (4)$$

where,  $\phi_1(i, j), \dots, \phi_k(i, j), \dots, \phi_K(i, j)$  are 2D-2D orthogonal functions, i.e.,  $\langle \phi_k(i, j), \phi_l(i, j) \rangle = 0$  for  $k \neq l$ , and  $\alpha_1, \dots, \alpha_k, \dots, \alpha_K$  are parameters to be optimized.

Jain and Zongker [14] have proposed a 2DW where  $\{\phi_k(i, j)\}$  are 2D orthogonal sinusoids. The optimization of the parameters  $\{\alpha_k\}$  is done by a coarse-to-fine strategy that the parameters of low-frequency sinusoids are firstly determined by the gradient descent method (i.e., an iterative method) and then the parameters of high-frequency sinusoids are determined similarly. Thus, their optimization strategy is doubly iterative.

Active shape model (ASM) proposed by Cootes et al. [6] can be employed for handwritten character recognition as an orthogonal 2DW. In ASM, 2DW is represented by a linear combination of principal deformations of de-

formable patterns. The principal deformations are orthogonal and provided by applying principal component analysis (PCA) to actual deformations collected from training patterns. Inspired by ASM, Shi et al. [45] have proposed a 1D-2D EM technique for the character recognition task. In [45] a linear (e.g., 1D) reference pattern is fitted to a 2D input pattern. The fitting is governed by the principal deformations of the 1D reference pattern and the optimal fitting is searched for by a gradient descent method. Uchida and Sakoe [56] have extended ASM to fully 2D-2D EM, i.e., planar EM and applied into handwritten character recognition. A related idea can be found in [19], where the displacement between the strokes of two skeletonized handwritten characters is evaluated by the Maharanobis distance.

### 3.2 Non-parametric 2DW

Non-parametric 2DWs can be divided into two classes; continuous 2DW and discrete 2DW. The continuous 2DW (Section 3.2.1) is defined as a mapping  $\mathbf{F} : (i, j) \in R^2 \mapsto (x, y) \in R^2$ . The discrete 2DW is defined by  $2N^2$  variables  $((x_{1,1}, y_{1,1}), \dots, (x_{i,j}, y_{i,j}), \dots, (x_{N,N}, y_{N,N}))$  where  $(x_{i,j}, y_{i,j})$  denotes the pixel on **B** corresponding to the pixel  $(i, j)$  on **A**.

The discrete 2DW is further divided into two classes; unconstrained 2DW (Section 3.2.2) and constrained 2DW (Section 3.2.3). In unconstrained 2DW, the mapping of pixel  $(i, j)$ , i.e.,  $(x_{i,j}, y_{i,j})$ , is independent of the mapping of other pixels. On the other hand, in constrained 2DW, the mapping of pixel  $(i, j)$  depends on the mapping of adjacent pixels of  $(i, j)$ .

#### 3.2.1 Continuous 2DW

Non-parametric and continuous 2DW is often assumed as a continuous and differentiable function and optimized by some iterative optimization strategy where 2DW is iteratively updated. In this sense, non-parametric and continuous 2DW is similar to parametric 2DWs.

Deterministic relaxation [43] is an iterative optimization strategy for variational problems. When optimizing non-parametric and continuous 2DW **F** by deterministic relaxation, the Euler-Lagrange equation of the underlying variational problem (1) is firstly discretized to obtain a system of nonlinear equations. Then the sub-optimal 2DW can be obtained by solving the equations by some iterative method, such as the Gauss-Seidel method. One should take care about the following three points on the practical use of this optimization strategy. First, the Euler-Lagrange equation is only a necessary condition for the optimal solution of the variational problem. Second, since the sys-

<sup>†</sup>In principle, the tangent distance method can deal with nonlinear deformations, if the displacement fields by the nonlinear deformations can be defined explicitly. In fact, the method has been extended to deal with several nonlinear deformations [56]. Since the original method [47] was examined with mainly linear transformations, it is classified as a linear 2DW in this paper.

tem of equations is nonlinear, the Gauss-Seidel method cannot guarantee even its convergence. Third, numerical errors will occur in approximated derivatives and interpolated pixel-values. For relaxing these problems, regularization techniques and/or coarse-to-fine strategies are effective. Mizukami et al. [28] have successfully employed the deterministic relaxation in handwritten character recognition while using a regularization technique and a careful coarse-to-fine strategy.

Webster and Nakagawa [61] and Nakagawa et al. [33] have proposed a motion equation-based EM technique. In their techniques, an elastic membrane created from  $\mathbf{B}$  is falling into a potential field created from  $\mathbf{A}$ . The state of the membrane showing  $\mathbf{B}_F$  is iteratively updated by calculating its motion equation until an equilibrium state.

### 3.2.2 Discrete and Unconstrained 2DW

Local perturbation [3], [11], [13], [23], [27], [42], [64] (also called image distortion model in [15], [16]) is the most simple optimization strategy for discrete 2DWs. Local perturbation is based on pixel-independent optimization; that is, for each pixel on a character image, its best corresponding pixel on another character image is searched for locally and independently.

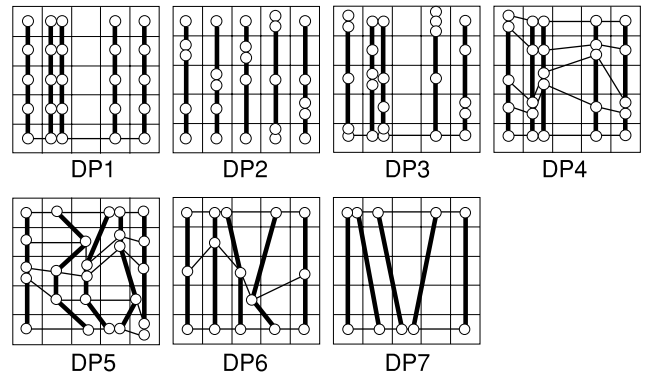
Local perturbation possesses great merit that it requires far less complexity than other optimization strategies. However, local perturbation possesses a weak-point that the resulting 2DW becomes jaggy due to the noise and the ambiguity in pixel features. Thus, careful coarse-to-fine strategies [13], [27], smoothing of local displacements [3], [11], sequential (outside-to-inside) optimization with mild constraints [11], and/or sophisticated pixel features [16], [42] will be indispensable to expect sufficient performance.

### 3.2.3 Discrete and Constrained 2DW

For regulating flexibility, discrete 2DWs often employs constraints between two adjacent pixels. For example, continuity constraints, such as  $|x_{i,j} - x_{i,j-1}| < \Delta_{xi}$  where  $\Delta_{xi}$  is a positive small value, are often imposed on 2DW to exclude large "gaps" from 2DW. Under such constraints, local perturbation cannot guarantee globally optimal 2DW or cannot provide any 2DW satisfying the constraints.

Dynamic programming (DP) is the most popular strategy for optimizing constrained 2DW. Its merits can be summarized as follows:

- *Accuracy*: DP can provide globally optimal 2DW.
- *Versatility*: For example, DP accepts undifferentiable objective functions. Position-dependent constraints and pixel features are also acceptable. Because of this versatility, the incorporation of the various properties of handwritten characters can be easily done.
- *Numerical stability*: DP is purely a combinatorial (i.e., discrete) optimization strategy and thus free from numerical assumptions.



**Fig. 5** Types of DP-based EM. For each type, a possible 2DW is illustrated as a deformed mesh. The link of the mesh indicates that the pixels connected by the link are restricted by some constraints.

HMM, which is a popular framework of EM, can be considered as a stochastic extension of DP. Thus, HMM and DP are not distinguished in this section unless otherwise mentioned.

The previous discrete and constrained 2DWs optimized by DP can be classified into DP1–DP7 of Fig. 5. Type DP1–DP4 are not fully two-dimensional 2DWs. That is, all the pixels on the  $i$ th column (the  $j$ th row) of  $\mathbf{A}$  are mapped together to the (same)  $x$ th column (the  $y$ th row) of  $\mathbf{B}$  and therefore those types cannot compensate truly 2D deformations, such as rotation and slant. Type DP5–DP7 are fully two-dimensional techniques. Note that every type except for DP5 has its transposed (i.e., 90° rotated) version. Also note that as shown in Fig. 5 it is often assumed that all 2DWs are restricted by boundary constraints that any boundary pixel of  $\mathbf{A}$  corresponds to a boundary pixel of  $\mathbf{B}$ .

#### (a) DP1

DP1 is the simplest 2DW and can compensate deformations that all the pixels of each column are shifted equally. The constraints of DP1 are:

$$DP1: \begin{cases} 0 \leq x_{i,j} - x_{i-1,j} \leq \Delta_{xi} \\ x_{i,j} = x_{i,j-1}, \\ y_{i,j} = j. \end{cases}$$

DP1 itself has been employed in word recognition [5], [25], [29], [46] rather than isolated handwritten character recognition. The optimization of DP1 with DP requires  $O(N^3)$  computations.

DP1 is often repeated to compensate complex deformations of handwritten characters and words. Nakano et al. [35] have proposed a DP-based EM technique where the vertical version of DP1 is optimized after the horizontal version is optimized. Hallouli et al. [9] have compared several combinations of vertical and horizontal versions of DP1 in the framework of HMM. In [36], DP1 is repeated four times under different feature vectors having different roles on representing spatial distribution of strokes. Wang et al. [60] firstly use a horizontal DP1 for segmenting a handwritten word into its component characters and secondly use a vertical DP1 for compensating the vertical deformation of each

of those component characters.

(b) *DP2*

DP2 is comprised of independent one-dimensional vertical warpings. DP2 cannot compensate any horizontal deformation. The constraints of DP2 are:

$$\text{DP2: } \begin{cases} x_{i,j} = i, \\ 0 \leq y_{i,j} - y_{i,j-1} \leq \Delta_{y_j} \end{cases}$$

The optimization of DP2 with DP requires  $O(N^3)$  computations.

DP2 is often repeated like DP1. In [12], the vertical version of DP2 is optimized after the horizontal version is optimized. Tsukumo [51] has proposed an EM technique where a character image is firstly decomposed into four images each of which contains  $0^\circ$  ( $\leftrightarrow$ ),  $45^\circ$  ( $\nearrow$ ),  $90^\circ$  ( $\updownarrow$ ), or  $-45^\circ$  ( $\nwarrow$ ) components of the character. Then DP2 is orthogonally applied to each of them. Specifically, for the image of  $0^\circ$  components, the vertical version of DP2 is applied to compensate, for example, the vertical shifts of horizontal strokes. Although the horizontal shifts of horizontal strokes cannot be compensated by his 2DW, they are successfully eliminated in advance by the horizontal blurring operation.

(c) *DP3*

DP3 can be considered as a combination of DP1 and DP2. Among DP1–DP7, DP3 is the most popular one. This may be because DP3 can compensate both vertical and horizontal deformations simultaneously with polynomial-order computations. The HMM version of DP3 is so-called Pseudo-2D HMM [1], [20], [22] and widely used in recognizing handwritten characters [22], machine-printed words [1], [20], [68], and handwritten words [2]. The constraints of DP3 are:

$$\text{DP3: } \begin{cases} 0 \leq x_{i,j} - x_{i-1,j} \leq \Delta_{x_i} \\ x_{i,j} = x_{i,j-1}, \\ 0 \leq y_{i,j} - y_{i,j-1} \leq \Delta_{y_j} \end{cases}$$

The computational amount required for optimizing DP3 is  $O(N^4)$ , i.e., still in a polynomial order.

Recently, DP3 has been extended by Keyser et al. [16]. Their 2DW allows column-wise local perturbation on the 2DW given by DP3. This extended DP3 can provide truly 2D warping with a feasible amount of computations. It should be careful, however, that local perturbation guarantees neither continuity nor monotonicity of 2DW and therefore resulting 2DW may show gaps and fold-overs.

(d) *DP4*

DP4 is a topology-preserving 2DW, while DP2 and DP3 are not topology-preserving 2DWs (because of the lack the constraint between  $y_{i,j}$  and  $y_{i-1,j}$ ). Thus, DP4 can avoid the overfitting of “P” to “b” while it can be happen in DP2 and DP3. The constraints of DP4 are:

$$\text{DP4: } \begin{cases} 0 \leq x_{i,j} - x_{i-1,j} \leq \Delta_{x_i} \\ x_{i,j} = x_{i,j-1}, \\ |y_{i,j} - y_{i-1,j}| \leq \Delta_{y_i}, \\ 0 \leq y_{i,j} - y_{i,j-1} \leq \Delta_{y_j} \end{cases}$$

Although the difference between the constraints of DP4 and those of DP3 seems small, the computational amount of DP4 is far larger than DP3. In fact, the DP-based optimization of DP4 requires  $O(N^3 \Delta_{y_j}^N)$  computations, i.e., exponential-order computations. Simply speaking, this is because we should keep all possible warping of the  $(i-1)$ th column on examining the warping of the  $i$ th column. DP4 is a restricted version of DP5 and therefore its algorithm can be easily derived from DP5 [52], [53].

(e) *DP5*

In DP5, the mappings of 4-adjacent pixels are mutually constrained. Thus, it is easy to regulate the flexibility of 2DW. The constraints of DP5 are:

$$\text{DP5: } \begin{cases} 0 \leq x_{i,j} - x_{i-1,j} \leq \Delta_{x_i} \\ |x_{i,j} - x_{i,j-1}| < \Delta_{x_j} \\ |y_{i,j} - y_{i-1,j}| < \Delta_{y_i} \\ 0 \leq y_{i,j} - y_{i,j-1} \leq \Delta_{y_j} \end{cases}$$

Levin and Pieraccini [22] have proposed “monotonic” 2DW where  $\Delta_{x_i}$ ,  $\Delta_{x_j}$ ,  $\Delta_{y_i}$ , and  $\Delta_{y_j}$  are set at infinity. Thus, their 2DW can preserve upper/lower and left/right relationships and does not care about continuity of character patterns. Inspired by Levin and Pieraccini, Uchida and Sakoe [52], [53] have proposed a monotonic and continuous 2DW where  $\Delta_{x_i} = \Delta_{y_j} = 2$  and  $\Delta_{x_j} = \Delta_{y_i} = 1$ . In some sense, the monotonic and continuous version of DP5 is the most fundamental and widely acceptable 2DW, because it approximates “rubber-sheet” EM.

Unfortunately, the optimization of DP5 (as well as DP4) is an NP-hard problem [18]. Even if character images are small ( $N \sim 20$ ), it is impossible to obtain the globally optimal 2DW. Thus, some approximation should be introduced for the practical use of DP5. In [52], [53], beam search is incorporated into the DP optimization process to obtain a sub-optimal 2DW with fewer computations. One can employ other local search-based approximation algorithms for DP-based EM. In [48], the optimization of DP5 is performed as a sequential and greedy process; after the warping of the  $(i-1)$ th column is determined, the warping of the  $i$ th column is determined satisfying inter-column constraints. Chen and Willson [4] have developed a similar approximation algorithm where the above sequential process is iterated. After the first iteration, the warping of the  $i$ th column is determined considering not only the warping of the  $(i-1)$ th column but also the warping of the  $(i+1)$ th column. In [39], the iteration proceeds alternately in horizontal and vertical directions. Uchida and Sakoe [54] have proposed an approximation algorithm which exploits the fact that the global optimization of DP5 can be done very fast if an image pattern is elongated one.

(f) *DP6 and DP7*

DP6 and DP7 [41], [55] have been proposed as computationally tractable versions of DP5. Both of DP6 and DP7 were derived from DP5 by piecewise-linear approximation. In DP6, each column of  $\mathbf{A}$  is fitted to  $\mathbf{B}$  as a broken line with one corner, called pivot. The correspondence of

the pixels except for the pivot and boundary pixels is determined by linear interpolation. Despite of those piecewise-linear approximations, DP6 still can compensate fully two-dimensional deformations. DP7 undergoes heavier approximation than DP6; it has no pivot and each column of  $A$  is fitted to  $B$  as a straight and slanted line. The computational amounts for the global optimization of DP6 and DP7 are  $O(N^6)$  and  $O(N^4)$ , respectively.

#### (g) Others

Another type of DP-based EM has been proposed in [30], [49], [62], where the matching proceeds diagonally from one corner  $(1, 1)$  to its opposite corner  $(N, N)$  while extending a rectangular matched region. Despite of their theoretical interests and computational feasibility, it seems that the flexibilities of their 2DWs do not match actual deformations of handwritten characters.

### 3.3 Hybrid between Parametric 2DW and Non-parametric 2DW

The boundary between parametric 2DW and non-parametric 2DW is ambiguous. In fact, the following EM techniques are their hybrids having the properties of both parametric 2DW and non-parametric 2DW.

In local affine transformation (LAT) proposed by Wakahara [57], 2DW is described by a set of many locally-effective affine transformations. Thus, LAT is a parametric 2DW in a microscopic sense and simultaneously a non-parametric 2DW in a macroscopic sense.

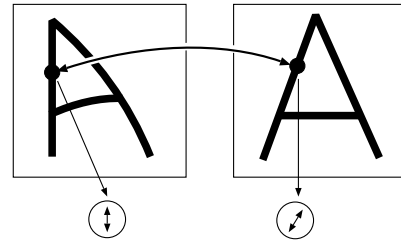
Uchida and Sakoe [56] have proposed an EM technique whose 2DW is defined as a linear combination of eigen-deformations, which are principal deformations in a category of handwritten characters. This 2DW is formulated as (4) and thus a parametric and orthogonal 2DW. The eigen-deformations themselves, however, are estimated from the results of some non-parametric EM. Thus, this technique can be considered as a parametric 2DW on its optimization and as a non-parametric 2DW on the ability of compensating deformations.

## 4. Related Topics

### 4.1 Reference Patterns for EM

The setting of reference patterns is generally important for handwritten character recognition. This importance is the same for EM-based recognizers. Although EM implicitly produces multiple reference patterns from one reference pattern  $B$  by applying  $F$  (as shown in Fig. 3), its producible reference patterns are generally limited (in order to avoid overfitting). Thus, for topology-preserving 2DWs, multiple reference patterns should be set when some category contains topologically different allograph.

Despite of the importance, most of EM-based character recognition techniques do not pay strict attention to this task. In fact, prototypes have been designed manually in



**Fig. 6** Inconsistency in pixel feature. Two pixels depicted by small filled circles are truly corresponding pixels and thus should be matched by 2DW. Their directional features, however, are rather different.

many EM-based recognizers. Otherwise, all training patterns are directly used as prototypes, at the cost of computational complexity (e.g., [14], [47]).

Clustering is the most promising way to setting prototypes automatically. Matsumoto et al. [26] have pointed out that the objective function of clustering should be designed using EM distance instead of conventional Euclidean distance. This is because the prototypes optimized under Euclidean distance-based clustering are not optimal for EM distance-based discrimination. (See also Fig. 2.) In [26], a k-means algorithm based on EM distance is proposed. In [10], several clustering algorithms based on the tangent distance (see Sect. 3.1.1) have been proposed.

### 4.2 Pixel Feature

In EM, pixel feature vectors are required to be less ambiguous; only two pixels corresponding truly should have similar pixel feature vectors. This is because ambiguous pixel feature vectors, such as grey-level feature, may confuse pixel-to-pixel correspondences.

As less ambiguous pixel feature vectors, local context [16] and directional feature [28], [53] are simple and reasonable choices. Those shape-sensitive features, however, face a problem. That is, those feature vectors of corresponding pixels are often not the same (Fig. 6). Thus, strictly speaking, one should modify the feature vectors according to 2DW. Invariant features, such as moments, will relax this problem.

### 4.3 Category-Dependent Deformation Tendency

Most of conventional EM techniques assume that all categories have the same deformation tendency. Although the assumption is approximately correct, we should accept the existence of category-dependent deformation tendency. Figure 7 is a simple example that proves its existence. If an EM technique based on the assumption can compensate the deformations of “M” (Fig. 7 (a)), it may suffer from the overfitting of “H” to “A” (Fig. 7 (b)). This simple example indicates the necessity of category-dependent EM techniques.

This necessity can be confirmed from the experimental result of Table 1, where the recognition results of several DP-based EM techniques of Fig. 5 are shown. This recognition experiment was conducted on 26 character categories of

**Table 1** The number of misrecognized samples. The error rate can be derived by dividing each number by 500. DP1, DP3, DP5, and DP6 are DP-based EM techniques illustrated in Fig. 5. The number is underlined at the best EM of each category.

cat.	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	Total
rigid	6	48	10	14	48	61	20	10	18	33	19	<u>3</u>	108	18	54	17	12	45	29	7	10	31	28	22	3	45	719
DP1	2	9	1	3	15	18	16	9	<u>6</u>	<u>1</u>	7	6	31	15	12	7	6	19	21	<u>5</u>	<u>4</u>	6	3	11	0	17	250
DP3	<u>1</u>	4	<u>0</u>	4	<u>10</u>	13	2	9	9	<u>1</u>	5	12	23	17	4	8	5	4	6	10	11	<u>5</u>	4	5	1	<u>5</u>	178
DP5	<u>1</u>	6	<u>0</u>	4	12	21	4	3	12	2	1	12	<u>6</u>	10	4	31	5	8	6	14	5	6	3	2	5	7	190
DP6	<u>1</u>	<u>2</u>	<u>0</u>	<u>1</u>	11	<u>12</u>	<u>1</u>	<u>2</u>	<u>6</u>	2	0	7	10	<u>9</u>	<u>3</u>	<u>2</u>	<u>4</u>	<u>3</u>	<u>4</u>	10	5	7	<u>2</u>	<u>1</u>	1	9	115



**Fig. 7** Example of category-dependent deformation tendency. (a) In category “M”, two parallel vertical strokes are often slanted to be closer. (b) In category “H”, however, the same deformation is rarely observed.

capital English alphabets. For each category, 600 handwritten character samples of ETL6[69] were prepared. Four-dimensional directional features ( $0^\circ$  ( $\leftrightarrow$ ),  $45^\circ$  ( $\nearrow$ ),  $90^\circ$  ( $\updownarrow$ ), and  $-45^\circ$  ( $\nwarrow$ )) were extracted at each pixel and used as a five-dimensional pixel feature vector together with one-dimensional gray-level feature. Each sample were linearly scaled to  $20 \times 20$ . The first 100 samples of each category were simply averaged to create one reference pattern  $\mathbf{B}$ . The remaining 500 samples were used as test samples  $\mathbf{A}$ . As DP-based EM techniques, DP1 ( $\Delta_{xi} = 2$ ), DP3 ( $\Delta_{xi} = \Delta_{yj} = 2$ ), DP5 ( $\Delta_{xi} = \Delta_{yj} = 2, \Delta_{xj} = \Delta_{yi} = 1$ ), and DP6 were chosen.

The results of Table 1 indicates a simple and important fact the most appropriate flexibility is different in each category. The most flexible 2DW (DP5) could not provide the best result in many categories; on the contrary, the most simple 2DW (DP1) and even rigid matching could provide the best result for several categories. That is, each category has its own range of deformations and excessive/insufficient flexibility often degrades the recognition performance. In this sense, category-dependent EM techniques, such as HMM and ASM-based EM, are more promising than category-independent ones.

## 5. Conclusion and Future Work

Elastic matching (EM) techniques employed in handwritten character recognition tasks were surveyed. Formally, EM is defined as the optimization problem of two-dimensional warping (2DW) which specifies the pixel-to-pixel correspondence between two subjected character image patterns. The image distance evaluated through 2DW is called EM distance and invariant to a certain range of geometric deformations. Thus, by using EM distance as a discriminant function, we can develop recognition systems robust to deformations of handwritten characters.

A classification of EM techniques was done according to the types of 2DW. There are two classes of 2DWs; parametric 2DW and non-parametric 2DW. Each of those two

classes is further divided into several classes. Each class and its optimization strategy are related closely. For example, dynamic programming (DP) is often chosen for the optimization of non-parametric, constrained, and discrete 2DW.

Future research on EM for handwritten character recognition should tackle the following problems:

- *Design of 2DW based on the actual deformations.* As shown in Sect. 4.3, category-dependent deformation tendencies should be incorporated into 2DW by observing the deformations of actual handwritten characters. Statistic and/or stochastic frameworks will be useful. Discriminative learning of 2DW will also be useful to suppress the misrecognitions due to overfitting.
- *Combination of parametric 2DW and non-parametric 2DW.* The deformations of handwritten characters can be decomposed into global deformations and local deformations. Scaling, rotation, translation and projective transformation of an entire character image are examples of global deformations. Independent and partial changes of stroke direction, curvature, and length are the examples of local deformations. EM should compensate both deformations. Since parametric 2DW and non-parametric 2DW are suitable for compensating global deformations and local deformations respectively, their cooperative combination will be promising.
- *Reduction of computational complexity.* Generally speaking, EM requires a fair amount of computations. This fact is crucial especially for characters comprised of many categories, such as Chinese characters. Thus, acceleration of the EM algorithm is necessary. Coarse classification based on a rigid distance is a possible remedy.
- *Combination with shape normalization.* As noted in Sect. 2.3, shape normalization can compensate many deformations and therefore it is useful to narrow the range of 2DW.
- *Reference pattern.* There are few investigations on setting reference patterns for EM-based handwritten character recognition. As noted in Sect. 4.1, conventional clustering techniques based on Euclidean distance are not appropriate for EM. The number of reference pattern per category will depend on the property of 2DW.
- *Feature extraction.* As noted in Sect. 4.2, less ambiguous pixel features are required for accurate 2DW. If such a desirable pixel feature is available, we will be able to use unconstrained 2DW instead of costly constrained 2DW.

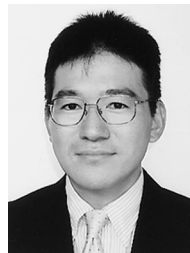


- *EM for handwritten word recognition.* EM techniques employed in word recognition are rather simple like DP1 of Sect. 3.2.3. In handwritten words, not only deformations within individual characters but also deformations between adjacent characters are observed. The compensation of such complex deformations is challenging.
- *Utilization of optimized 2DW.* In handwritten character recognition, only EM distance is often important as a discriminant function and 2DW obtained on minimizing the EM distance is not emphasized. The optimized 2DW, however, is very meaningful because the optimized 2DW between  $A$  and  $B$  represents the deformation of  $B$  relative to  $A$ . From the viewpoint of structural analysis, the utilization of 2DW is very promising to extract various properties of handwritten characters. Automatic generation of active shape models is one example [56]. Writer identification and character synthesis are also promising.

## References

- [1] O. Agazzi, S. Kuo, E. Levin, and R. Pieraccini, "Connected and degraded text recognition using planar hidden Markov models," Proc. ICASSP, pp. V113–116, 1993.
- [2] R. Bippus and V. Märgner, "Script recognition using inhomogeneous P2DHMM and hierarchical search space reduction," Proc. ICDAR, pp. 773–776, 1999.
- [3] D.J. Burr, "A dynamic model for image registration," Comput. Graph. Image Process., vol. 15, pp. 102–112, 1981.
- [4] M.C. Chen and A.N. Willson, Jr., "Motion-vector optimization of control grid interpolation and overlapped block motion compensation using iterated dynamic programming," IEEE Trans. Image Process., vol. 9, no. 7, pp. 1145–1157, 2000.
- [5] W. Cho, S.-W. Lee, and J.H. Kim, "Modeling and recognition of cursive words with hidden Markov models," Pattern Recognit., vol. 28, no. 12, pp. 1941–1953, 1995.
- [6] T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham, "Active shape models—Their training and application," Comput. Vis. Image Underst., vol. 61, no. 1, pp. 38–59, Jan. 1995.
- [7] Y. Fujimoto, S. Kadota, S. Hayashi, M. Yamamoto, S. Yajima, and M. Yasuda, "Recognition of handprinted characters by nonlinear elastic matching," Proc. ICPR, pp. 113–118, 1976.
- [8] T.M. Ha and H. Bunke, "Off-line, handwritten numeral recognition by perturbation method," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 5, pp. 535–539, 1997.
- [9] K. Halloui, L. Likforman-Sulem, and M. Sigelle, "A comparative study between decision fusion and data fusion in Markovian printed character recognition," Proc. ICPR, vol. 3 of 4, pp. 147–150, 2002.
- [10] T. Hastie, P.Y. Simard, and E. Säckinger, "Learning prototype models for tangent distance," Adv. Neural Inf. Process. Syst., vol. 7, pp. 999–1006, 1995.
- [11] T. Hattori, Y. Watanabe, H. Sanada, and Y. Tezuka, "Absorption of local variations in handwritten character by an elastic transformation using vector field," IEICE Trans. Inf. & Syst. (Japanese Edition), vol. J66-D, no. 6, pp. 645–652, June 1983.
- [12] Y. Isomichi and T. Ogawa, "Pattern matching by using dynamic programming," J. IPSJ, vol. 16, no. 1, pp. 15–22, 1975.
- [13] Y. Izui, H. Harashima, and H. Miyagawa, "Handprinted Chinese characters recognition by hierarchical modification of dictionary," IEICE Trans. Inf. & Syst. (Japanese Edition), vol. J68-D, no. 3, pp. 361–368, March 1985.
- [14] A.K. Jain and D. Zongker, "Representation and recognition of handwritten digits using deformable templates," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 12, pp. 1386–1391, 1997.
- [15] D. Keysers, J. Dahmen, T. Theiner, and H. Ney, "Experiments with an extended tangent distance," Proc. ICPR, vol. 2 of 4, pp. 38–42, 2000.
- [16] D. Keysers, C. Gollan, and H. Ney, "Local context in non-linear deformation models for handwritten character recognition," Proc. ICPR, vol. 4 of 4, pp. 511–514, 2004.
- [17] D. Keysers, W. Macherey, H. Ney, and J. Dahmen, "Adaptation in statistical pattern recognition using tangent vectors," IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 2, pp. 269–274, 2004.
- [18] D. Keysers and W. Unger, "Elastic image matching is NP-complete," Pattern Recognit. Lett., vol. 24, no. 1–3, pp. 445–453, 2003.
- [19] F. Kimura, M. Yoshimura, Y. Miyake, and M. Ichikawa, "Unconstrainedly handprinted 'KATAKANA' character recognition by a stroke structure analysis method," IEICE Trans. Inf. & Syst. (Japanese Edition), vol. J62-D, no. 1, pp. 16–23, Jan. 1979.
- [20] S. Kuo and O.E. Agazzi, "Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 16, no. 8, pp. 842–848, 1994.
- [21] S.-W. Lee and J.-S. Park, "Nonlinear shape normalization methods for the recognition of large-set handwritten characters," Pattern Recognit., vol. 27, no. 7, pp. 895–902, 1994.
- [22] E. Levin and R. Pieraccini, "Dynamic planar warping for optical character recognition," Proc. ICASSP, pp. III 149–152, 1992.
- [23] N. Liolios, E. Kavallieratou, N. Fakotakis, and G. Kokkinakis, "A new shape transformation approach to handwritten character recognition," Proc. ICPR, vol. 1 of 4, pp. 584–587, 2002.
- [24] C.L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten digit recognition: Investigation of normalization and feature extraction techniques," Pattern Recognit., vol. 37, no. 2, pp. 265–279, 2004.
- [25] J. Makhoul, R. Schwartz, C. Lapre, and I. Bazzi, "A script-independent methodology for optical character recognition," Pattern Recognit., vol. 31, no. 9, pp. 1285–1294, 1998.
- [26] N. Matsumoto, S. Uchida, and H. Sakoe, "Prototype setting for elastic matching-based image pattern recognition," Proc. ICPR, vol. 1 of 4, pp. 224–227, 2004.
- [27] S. Meguro and M. Umeda, "An extraction of shape derivations in handwritten characters by hierarchical pattern matching," IEICE Technical Report, PRL77-70, 1978.
- [28] Y. Mizukami, "A handwritten Chinese character recognition system using hierarchical displacement extraction based on directional features," Pattern Recognit. Lett., vol. 19, no. 7, pp. 595–604, 1998.
- [29] M. Mohamed and P. Gader, "Handwritten word recognition using segmentation-free hidden Markov modeling and segmentation-based dynamic programming techniques," IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 5, pp. 548–554, 1996.
- [30] R.K. Moore, "A dynamic programming algorithm for the distance between two finite areas," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-1, no. 1, pp. 86–88, 1979.
- [31] S. Mori, C.Y. Suen, and K. Yamamoto, "Historical review of OCR research and development," Proc. IEEE, vol. 80, no. 7, pp. 1029–1058, 1992.
- [32] S. Mori, K. Yamamoto, and M. Yasuda, "Research on machine recognition of handprinted characters," IEEE Trans. Pattern Anal. Mach. Intell., vol. PAMI-6, no. 4, pp. 386–405, 1984.
- [33] M. Nakagawa, T. Yanagida, and T. Nagasaki, "An off-line character recognition method employing model-dependent pattern normalization by an elastic membrane model," Proc. ICDAR, pp. 495–498, 1999.
- [34] K. Nakata, Y. Nakano, and Y. Uchikura, "Recognition of Chinese characters," Proc. Conf. Machine Perception of Patterns and Pictures, pp. 45–52, 1972.
- [35] Y. Nakano, K. Nakata, Y. Uchikura, and A. Nakajima, "Improvement of Chinese character recognition using projection profiles," Proc. Int. Joint Conf. Pat. Recog., pp. 172–178, 1973.

- [36] H. Nishimura, M. Tsutsumi, M. Maruyama, M. Miyao, and Y. Nakano, "Off-line handwritten character recognition using integrated 1D HMMs based on feature extraction filters," Proc. ICDAR, pp.417–421, 2001.
- [37] H. Ogawa, ed., *The new development of pattern recognition and understanding*, The Institute of Electronics, Information, and Communication Engineers, 1994.
- [38] H.-S. Park and S.-W. Lee, "A truly 2-D hidden Markov model for off-line handwritten character recognition," *Pattern Recognit.*, vol.31, no.12, pp.1849–1864, 1998.
- [39] G.M. Quénou, "The orthogonal algorithm for optical flow detection using dynamic programming," Proc. ICASSP, pp.III-249–252, 1992.
- [40] M. Revow, C.K.I. Williams, and G.E. Hinton, "Using generative models for handwritten digit recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.18, no.6, pp.592–606, 1996.
- [41] M.A. Ronee, S. Uchida, and H. Sakoe, "Handwritten character recognition using piecewise linear two-dimensional warping," Proc. ICDAR, pp.39–43, 2001.
- [42] T. Saito, H. Yamada, and K. Yamamoto, "An analysis of hand-printed Chinese characters by directional pattern matching approach," *IEICE Trans. Inf. & Syst. (Japanese Edition)*, vol.J65-D, no.5, pp.550–557, May 1982.
- [43] K. Sakau, A. Amano, and N. Yokoya, "Optimization approaches in computer vision and image processing," *IEICE Trans. Inf. & Syst.*, vol.E82-D, no.3, pp.534–547, March 1999.
- [44] H. Sakoe, "Handwritten character recognition by rubber string matching method," *IEICE Technical Report*, PRL74-20, 1974.
- [45] D. Shi, S.R. Gunn, and R.I. Dampier, "Handwritten Chinese radical recognition using nonlinear active shape models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.25, no.2, pp.277–280, Feb. 2003.
- [46] O. Shiku, A. Nakamura, H. Kuroda, and S. Miyahara, "A method for handwritten Japanese word recognition based on holistic strategy," *Trans. IPSJ*, vol.41, no.4, pp.1086–1095, 2000.
- [47] P.Y. Simard, Y. Le. Cun, and J.S. Denker, "Efficient pattern recognition using a new transformation distance," *Adv. Neural Inf. Proces. Syst.*, vol.5, pp.50–58, 1993.
- [48] M. Sugimura, Y. Iiguni, and N. Adachi, "A 2-dimensional dynamic programming for image matching with Zernike moments," *IEICE Trans. Inf. & Syst. (Japanese Edition)*, vol.J80-D-II, no.1, pp.101–108, Jan. 1997.
- [49] E. Tanaka, "A two dimensional context-dependent similarity measure," *Trans. IECE Japan*, vol.E68, no.10, pp.667–673, 1985.
- [50] Ø.D. Trier, A.K. Jain, and T. Taxt, "Feature extraction methods for character recognition – A survey," *Pattern Recognit.*, vol.29, no.4, pp.641–662, 1996.
- [51] J. Tsukumo, "Handprinted Kanji character recognition based on flexible template matching," Proc. ICPR, pp.483–486, 1992.
- [52] S. Uchida and H. Sakoe, "A monotonic and continuous two-dimensional warping based on dynamic programming," Proc. ICPR, vol.1 of 2, pp.521–524, 1998.
- [53] S. Uchida and H. Sakoe, "Handwritten character recognition using monotonic and continuous two-dimensional warping," Proc. ICDAR, pp.499–502, 1999.
- [54] S. Uchida and H. Sakoe, "An approximation algorithm for two-dimensional warping," *IEICE Trans. Inf. & Syst.*, vol.E83-D, no.1, pp.109–111, Jan. 2000.
- [55] S. Uchida and H. Sakoe, "Piecewise linear two-dimensional warping," Proc. ICPR, vol.3, pp.538–541, 2000.
- [56] S. Uchida and H. Sakoe, "Handwritten character recognition using elastic matching based on a class-dependent deformation model," Proc. ICDAR, vol.1 of 2, pp.163–167, 2003.
- [57] T. Wakahara, "Shape matching using LAT and its application to handwritten numeral recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.16, no.6, pp.618–629, 1994.
- [58] T. Wakahara and K. Odaka, "Adaptive normalization of handwritten characters using global/local affine transformation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.20, no.12, pp.1332–1341, 1998.
- [59] T. Wakahara, Y. Kimura, and A. Tomono, "Affine-invariant recognition of gray-scale characters using global affine transformation correlation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.23, no.4, pp.384–395, 2001.
- [60] W. Wang, A. Brakensiek, A. Kosmala, and G. Rigoll, "Multi-branch and two-pass HMM modeling approaches for off-line cursive handwriting recognition," Proc. ICDAR, pp.231–235, 2001.
- [61] R. Webster and M. Nakagawa, "An on-line/off-line compatible character recognition method based on a dynamic model," *IEICE Trans. Inf. & Syst.*, vol.E80-D, no.6, pp.672–683, June 1997.
- [62] C.-M. Wu, P. Liu, and W.-C. Chang, "Unconstrained-endpoint dynamic space-warping algorithm with experiments in binary English character images," *Int. J. Electron.*, vol.78, no.1, pp.55–66, 1995.
- [63] H. Yamada, "Contour DP matching method and its application to handprinted Chinese character recognition," Proc. ICPR, pp.389–392, vol.1 of 2, 1984.
- [64] H. Yamada, T. Saito, and S. Mori, "An improvement of correlation method – Locally maximized correlation," *IEICE Trans. Inf. & Syst. (Japanese Edition)*, vol.J64-D, no.10, pp.970–976, Oct. 1981.
- [65] H. Yamada, K. Yamamoto, and T. Saito, "A nonlinear normalization method for handprinted Kanji character recognition – Line density equalization," *Pattern Recognit.*, vol.23, no.9, pp.1023–1029, 1990.
- [66] K. Yamamoto and A. Rosenfeld, "Recognition of hand-printed KANJI characters by a relaxation method," Proc. ICPR, vol.1 of 2, pp.395–398, 1982.
- [67] M. Yasuda, K. Yamamoto, and H. Yamada, "Effect of the perturbed correlation method for optical character recognition," *Pattern Recognit.*, vol.30, no.8, pp.1315–1320, 1997.
- [68] C. Yen, S.-S. Kuo, and C.-H. Lee, "Minimum error rate training for PHMM-based text recognition," *IEEE Trans. Image Proc.*, vol.8, no.8, pp.1120–1124, 1999.
- [69] <http://www.is.aist.go.jp/etlcedb/>



**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. Since 2002, he has been 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, IPSJ, ITE, and ASJ.



**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 IEICE Paper Award. He also received Kamura Memorial Prize from Kyushu Institute of Technology. Dr. Sakoe is a member of IEEE, IPSJ, ITE, and ASJ.