Elastic Matching Techniques for Handwritten Character Recognition

Seiichi Uchida

Faculty of Information Science and Electrical Engineering, Kyushu University

Motooka 744, Nishi-ku, Fukuoka-shi, 819--0395 Japan

Tel: +81-92-802-3586
Fax: +81-92-802-3600

uchida@ait.kyushu-u.ac.jp

Keywords: Document Management Systems, Image Processing, Image Analysis

Elastic Matching Techniques for Handwritten Character Recognition

Abstract

This chapter reviews various elastic matching techniques for handwritten character recognition. Elastic matching is formulated as an optimization problem of planar matching, or pixel-to-pixel correspondence, between two character images under a certain matching model, such as affine and nonlinear. Use of elastic matching instead of rigid matching improves the robustness of recognition systems against geometric deformations in handwritten character images. In addition, the optimized matching itself represents the deformation of handwritten characters and thus is useful for statistical analysis of the deformation. This chapter argues that the general property of elastic matching techniques and their classification by matching models and optimization strategies. It also argues various topics and future work related to elastic matching for emphasizing theoretical and practical importance of elastic matching.

INTRODUCTION

In handwritten character recognition, it is important to tackle geometric deformations of characters. The geometric deformations are classified into the following four types: fluctuation of stroke thickness, linear deformations (such as translation, scaling, shear, and rotation), nonlinear and topology-preserving deformations, and deformations changing topology. Those deformations will be caused by many factors; for example, writing material, writer's habit, writing speed, writing style (especially, cursive style), character size, inherent character shape, and noise and geometric transformation at character image acquisition.

The purpose of this chapter is to overview various elastic matching (EM) techniques for handwritten character recognition. EM is also called deformable template (Trier, Jain, & Taxt, 1996), flexible matching (Mori, Yamamoto, & Yasuda, 1984), or nonlinear template matching (Mori, Suen, & Yamamoto, 1992), and has been developed by many researchers to tackle the geometric deformations. 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. For more general surveys, see Glasbey, & Mardia (1998), Jain, Zhong, & Dubuisson-Jolly (1998), Lester & Arridge (1999), Redert, Hendriks, & Biemond (1999), and Zitová, & Flusser (2003).

EM is formulated as an optimization problem of planar matching, or 2D-2D mapping between two character images, $\mathbf{A}$ and $\mathbf{B}$. From another viewpoint, EM treats a character image $\mathbf{A}$ like a ``rubber sheet" and fits it to another character image $\mathbf{B}$ as close as possible. Hereafter, this 2D-2D mapping from $\mathbf{A}$ to $\mathbf{B}$ is called 2D warping (2DW). Note that we can consider EM based on 1D-2D mapping where a 1D-stroke model is fitted to input image, although this chapter mainly concern EM techniques based on 2DW. In a later section, we will briefly review these EM techniques based on 1D-2D mapping.

For handwritten character recognition, EM possesses two merits. The first merit is that the distance evaluated under the optimized 2DW is deformation-invariant. Thus, by using the EM distance as a discriminant function, we can realize character recognition systems robust to the geometric deformations. The range of the invariance depends on the definition of 2DW. That is, the more flexible 2DW becomes, the more invariant the EM distance becomes. Thus, EM has a potential to provide more intuitive and robust recognition frameworks than other deformation-invariant techniques, such as invariant feature (e.g., the horizontal projection profile by Nakata, Nakano, & Uchikura (1972)) and shape normalization (Lee & Park, 1994; Liu, Nakashima, Sako, & Fujisawa, 2004).

The second merit is that the optimized 2DW itself describes the deformation of subjected characters. This fact shows that EM possesses a useful property of structural analysis techniques. Furthermore, EM can be linked to statistical and stochastic frameworks by the merit. Active shape models (Cootes, 1995; Shi, Gunn, & Damper, 2003; Uchida, & Sakoe, 2003a) and (pseudo-) 2D HMMs (Agazzi, Kuo, Levin, & Pieraccini, 1993; Kuo, & Aggazi, 1994; Levin, & Pieraccini, 1992; Park, & Lee, 1998) are two good examples.

The remaining part of this chapter is organized as follows. First, EM is formulated as an optimization problem of 2D-2D mapping, i.e., 2DW. General properties of EM and the EM distance are also described. Second, EM techniques are classified according to their specific formulations of 2DW and optimization strategies. It will be emphasized that there is strong relation between the formulation and the optimization strategy. Third, several related topics are discussed, such as incorporation of category-dependent deformation tendency. Fourth, EM techniques based on 1D-2D mapping are briefly reviewed, which is another type of EM used in handwritten character recognition. Finally, conclusions are presented after listing various future tasks for EM.

OUTLINE OF ELASTIC MATCHING

Formulation of EM

As described before, EM is formulated as an optimization problem of 2D-2D mapping (i.e., 2DW) between two character images, $\mathbf{A}$ and $\mathbf{B}$. Let $\mathbf{a}_{i,j}$ and $\mathbf{b}_{x,y}$ denote pixel values (e.g., intensity values) or pixel feature vectors (e.g., RGB vectors) at pixel (i,j) on $\mathbf{A}$ and (x,y) on $\mathbf{B}$, respectively. While we can deal with the matching between two images of arbitrary sizes, we hereafter assume $N \times N$ images for simplicity.

Let $\mathbf{F}$ denote 2DW from $\mathbf{A}$ to $\mathbf{B}$, i.e., $\mathbf{F}:(i,j) \mapsto (x,y)$. The 2DW $\mathbf{F}$ represents the pixel-to-pixel correspondence between $\mathbf{A}$ and $\mathbf{B}$ as shown in Fig. 1. Using $\mathbf{F}$, we can consider $\mathbf{B}_\mathbf{F} = \{\mathbf{b}_{\mathbf{F}(i,j)}\}$ which is the deformed image of $\mathbf{B}$ by the 2DW $\mathbf{F}$. If $\mathbf{F}$ is a topology-preserving 2DW, $\mathbf{B}_\mathbf{F}$ undergoes rubber-sheet like deformations.

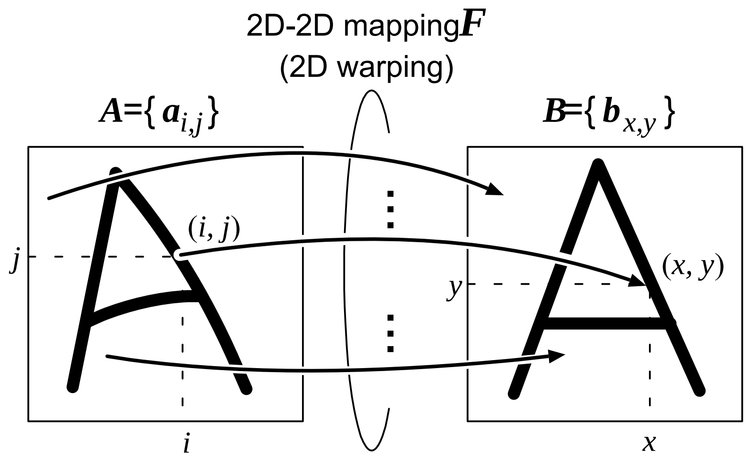

Fig. 1 2DW defined between two handwritten character images .(copyright@2005IEICE)

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

$$J_{\mathbf{A},\mathbf{B}}(\mathbf{F}) = D(\mathbf{A}, \mathbf{B}_{\mathbf{F}}) = \sum_{i=1}^{N} \sum_{j=1}^{N} \left\| \mathbf{a}_{i,j} - \mathbf{b}_{\mathbf{F}(i,j)} \right\|,$$

where $D(\mathbf{A}, \mathbf{B}_{\mathbf{F}})$ is a simple ``rigid'' distance between $\mathbf{A}$ and $\mathbf{B}_{\mathbf{F}}$, and $\|\cdot\|$ denotes a distance metric between two pixel feature vectors, such as Euclidean distance and absolute distance.

The *EM distance* $D_{\mathrm{EM}}(\mathbf{A}, \mathbf{B})$ is obtained as the solution of the above minimization problem, i.e.,

$$D_{\mathrm{EM}}(\mathbf{A}, \mathbf{B}) = \min_{\mathbf{F}} J_{\mathbf{A},\mathbf{B}}(\mathbf{F}) = J_{\mathbf{A},\mathbf{B}}(\tilde{\mathbf{F}}),$$

where $\tilde{\mathbf{F}}$ denote the optimal $\mathbf{F}$ which minimizes $J_{\mathbf{A},\mathbf{B}}(\mathbf{F})$. Clearly, $D_{\mathrm{EM}}(\mathbf{A}, \mathbf{B}) = D(\mathbf{A}, \mathbf{B}_{\tilde{\mathbf{F}}})$. That is, EM distance is the rigid distance between $\mathbf{A}$ and $\mathbf{B}$

after fitting $\mathbf{B}$ to $\mathbf{A}$ as close as possible.

Fig. 2 shows two examples of EM results (Uchida, & Sakoe, 1998), where the 2DW $\tilde{\mathbf{F}}$ is represented as a deformed mesh. If the optimal 2DW $\tilde{\mathbf{F}}$ shows correct correspondence, it represents the deformation of $\mathbf{B}$ relative to $\mathbf{A}$.

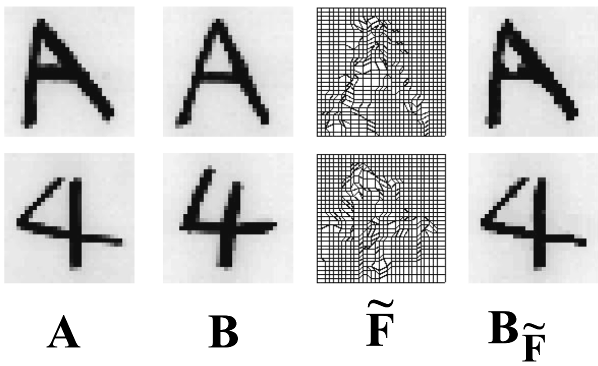

Fig. 2 Examples of EM results.

Recognizers based on the EM distance $D_{\mathrm{EM}}(\mathbf{A},\mathbf{B})$ are expected to be robust to geometric deformations. This is because $D_{\mathrm{EM}}(\mathbf{A},\mathbf{B})$ is invariant to the deformations compensable by $\mathbf{F}$. For example, if the 2DW $\mathbf{F}$ is defined as an affine transformation, $D_{\mathrm{EM}}(\mathbf{A},\mathbf{B})$ is theoretically invariant to any affine transformed version of $\mathbf{B}$.

Although past literatures have reported the usefulness of the EM distance on its robustness to the deformations, the EM-based recognizers often suffer from *overfitting,* which is the phenomenon that the distance between similar-shaped characters of different categories is underestimated. For example, if $\mathbf{F}$ is an affine transformation, $D_{\mathrm{EM}}(\text{"9"},\text{"6"})$ becomes very a small value because the affine transformation can compensate 180° rotation. Thus, the input ``9" may be misrecognized to the different category ``6". Generally, there is a trade-off between the ability of compensating deformations (i.e., the flexibility of $\mathbf{F}$) and the risk of overfitting.

Anisotropy is another important aspect of the EM distance $D_{\mathrm{EM}}(\mathbf{A},\mathbf{B})$. Thus, the set of patterns equidistant from a certain pattern do not form a hypersphere. Similarly, the centroid (the center of gravity) of a set of patterns may not be placed around the center of their distribution in Euclidean space. Fig. 3 (Matsumoto, Uchida, & Sakoe, 2004) shows an experimental result of deriving a centroid for a set of patterns. The small dots are actual handwritten numeral patterns of a category and the black triangle is their centroid. They are displayed in the 2D subspace spanned by their first two principal axes. When the Euclidean distance is used, the centroid is placed around the center of pattern distribution (Fig. 3(a)). In contrast, when the EM distance is used, the centroid is not placed around the center (Fig. 3(b)).

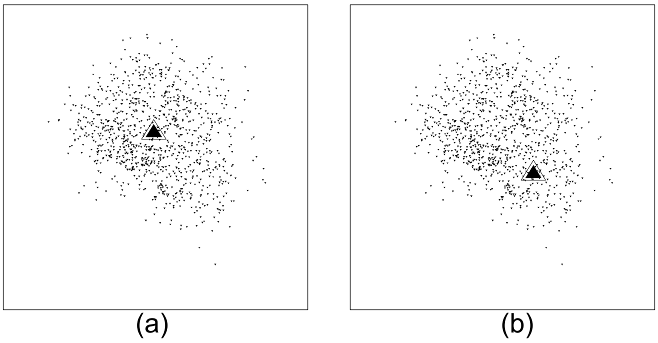

Fig. 3 The centroids under (a) Euclidean distance and (b) EM distance .(copyright@2005IEICE)

The EM distance is often asymmetric, i.e., $D_{\mathrm{EM}}(\mathbf{A},\mathbf{B}) \neq D_{\mathrm{EM}}(\mathbf{B},\mathbf{A})$, and therefore just a pseudo-distance metric. This asymmetric property comes from the fact that the 2DW defined above is not a bijective mapping. If symmetric property is necessary, one may simply use $D_{\mathrm{EM}}(\mathbf{A},\mathbf{B}) + D_{\mathrm{EM}}(\mathbf{B},\mathbf{A})$ instead of $D_{\mathrm{EM}}(\mathbf{A},\mathbf{B})$. A more plausible solution is the use of a bijective 2DW; in this case, not only $\mathbf{B}$ but also $\mathbf{A}$ are deformed by the 2DW and the optimization of the bijective 2DW tends to be complicated one. As reported in past literatures, this asymmetric property is not crucial for the recognition performance.

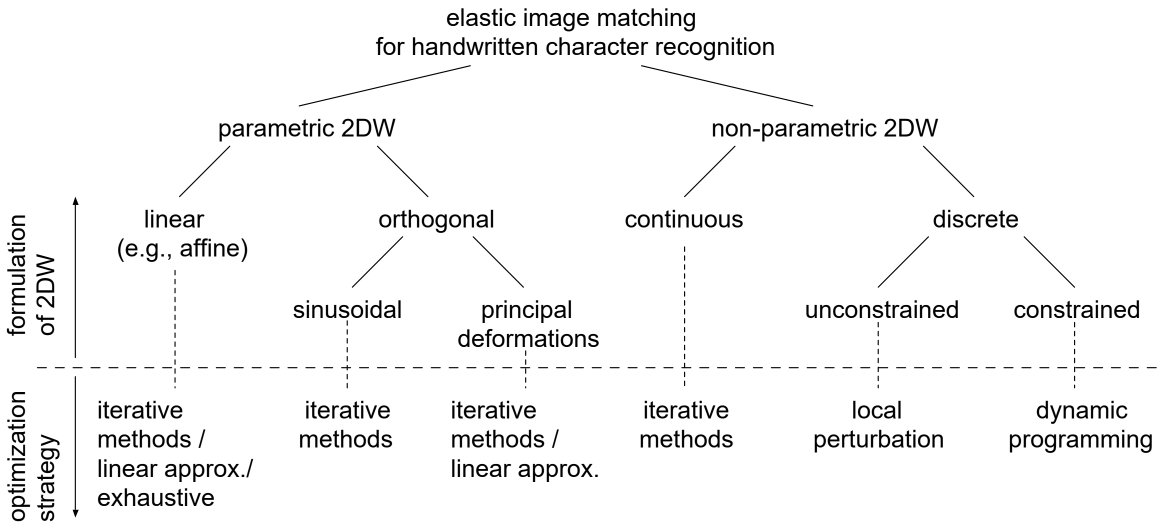

Fig. 4 Classification of planar EM techniques for handwritten character recognition. (copyright@2005IEICE)

CLASSIFICATION OF ELASTIC MATCHING TECHNIQUES

In this section, EM techniques are classified into several types. The classification can be done by two factors, i.e., the specific formulation of 2DW and the optimization strategy of 2DW. Fig. 4 (Uchida & Sakoe, 2005) shows a classification tree based on those factors.

The first factor affects the range of compensable deformations. While 2DW **F** was formulated as a general 2D-2D mapping in the previous section, specific formulations are necessary for individual tasks. As shown in Fig. 4, the EM techniques can be roughly divided into two classes by the 2DW formulation, i.e., parametric 2DW-based EM and non-parametric 2DW-based EM. In parametric 2DW, each variable does not represent pixel correspondence but represents a parameter that controls 2DW indirectly. In non-parametric 2DW, each variable which controls **F** directly represents pixel correspondence.

On the specification of 2DW, it is important to consider 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. Both of parametric 2DW and non-parametric 2DW are further classified into several classes depending on this consideration.

The second factor, that is, the optimization strategy of 2DW, 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.

As shown in Fig. 4, each class by the formulation of 2DW is closely related to several optimization strategies. In other words, possible optimization strategies are almost determined according to the formulation.

We should note that those two factors mutually affect the computational complexity of EM. For example, strategies for globally optimal solutions will require more computations than those for sub-optimal solutions in general. Effect of the formulation of 2DW on computational complexity is more complicated. In the following, the computational complexity of each EM technique is carefully discussed.

Parametric 2DW

Linear 2DW

Among parametric 2DWs, linear 2DW is the simplest and the most common one. Linear 2DW $\mathbf{F}:(i,j)\mapsto(x,y)$ is generally formulated as,

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

where α_1, α_2, …, and α_6 are real-valued parameters which control $\mathbf{F}$. If $\alpha_1=\alpha_5=1$ and $\alpha_2=\alpha_4=0$, this linear 2DW will compensate translation only. If $\alpha_2=\alpha_3=\alpha_4=\alpha_6=0$, the 2DW will compensate horizontal and vertical scaling. If all of the six parameters are controllable, $\mathbf{F}$ becomes the affine transformation.

Although the formulation of linear 2DW is very simple, its optimization problem becomes complicated one. In fact, since the parameters to be optimized are involved in a nonlinear image pattern function $\mathbf{B}$, the optimization problem also becomes nonlinear. Thus, EM techniques based on linear 2DW have employed iterative solutions or approximate solutions.

An example of the iterative solution of linear 2DW is done by Wakahara and his colleagues (Wakahara, & Odaka, 1998; Wakahara, Kimura, & Tomono, 2001). They have employed affine transformation as 2DW. The optimization problem of the parameters α_1, α_2, …, and α_6 is approximated as a linear problem by a successive iteration method.

The tangent distance method (Simard, Cun, & Denker, 1993; Keysers, Dahmen, Theiner, & Ney, 2000) has been proposed as another linear 2DW-based EM where a nonlinear optimization problem of linear 2DW is approximated as a linear problem by Taylor series expansion. Specifically, the tangent distance method assumes a nonlinear manifold which contains all deformed character patterns of a certain category. Then, this manifold is approximated be its tangent plane. Consequently, the EM problem is reduced to the minimum distance problem between the tangent plane and an input pattern, that is, a linear problem. Recently, this idea is successfully linked with statistic framework (Keysers, Macherey, Ney, & Dahmen, 2004).

A more straightforward solution is also possible by using an exhaustive search strategy. Yasuda, Yamamoto, & Yamada (1997) proposed the perturbed correlation method, where a 2D reference pattern is ``perturbed" by a huge number of affine transformations. Each of perturbed patterns is rigidly matched with a 2D input pattern and the best perturbed pattern is searched for. Since the number of possible parameter values becomes very large, this method requires numerous and repetitive 2D-2D rigid matching. Recent hardware, however, makes the method computationally tractable one. A similar method can be found in Ha, & Bunke (1997).

Perspective transformation is another important class of linear transformation. Nowadays, camera-based character recognition is widely investigated (Liang, Doermann, & Li, 2005) and one of its main hurdles is how to compensate the perspective transformation due to an oblique camera angle. In past researches, perspective transformation has not been compensated by 2DWs at individual characters of a document; instead, it has been compensated by a preprocessing technique, called dewarping, for the entire document image. When camera-based character recognition tackles "rumpled" documents, EM should be applied to individual characters for compensating perspective transformation.

Orthogonal 2DW

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

$$(x,y)=\sum_{k=1}^{K}\alpha_k\phi_k(x,y), \tag{1}$$

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

One of the most reasonable choices of $\{\phi_k\}$ will be orthogonal sinusoids. Jain, & Zongker (1997) have proposed a sinusoid-based 2DW and optimized its gain parameters $\{\alpha_k\}$ by a natural coarse-to-fine strategy; the parameters of low-frequency sinusoids are first determined and fixed by the gradient descent method (i.e., an iterative method) and then the parameters of high-frequency sinusoids are determined in a similar way.

The active shape model (ASM) proposed by Cootes et al. (1995) is a statistical deformation model defined as a linear combination of orthogonal deformations, which are provided by applying principal component analysis (PCA) to actual deformations collected from training patterns. Inspired by ASM, Shi et al. (2003) have proposed a 1D-2D EM technique for the character recognition task. In this work, 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 evaluated by the chamfer distance. The optimal fitting that minimizes the chamfer distance is searched for by a gradient descent method. A related idea can be found in Kimura et al.(1970), where the displacement between the strokes of two skeletonized handwritten characters is evaluated by the Maharanobis distance.

Uchida & Sakoe (2003a) have extended ASM to fully 2D-2D EM, i.e., planar EM and applied into handwritten character recognition. Fig. 5 shows the reference pattern of "A" deformed by applying the principal deformations ϕ_1,ϕ_2,ϕ_3 (called the eigen-deformations in Uchida & Sakoe (2003a)) of "A" positively or negatively. The first eigen-deformation represents the slant deformation and the second represents the vertical shift of the horizontal stroke. It is interesting to note that those eigen-deformations were estimated from the deformations collected as the results of non-parametric EM as noted later.

Uchida & Sakoe (2003b) have combined the above ASM technique with the tangent distance method. In this strategy, the eigen-deformations of Fig. 5 are converted into the tangent vectors of Fig. 6 by the Taylor series expansion. Those tangent vectors will span the tangent plane of the manifold which contains all the patterns realized by the ASM. Those tangent vectors are then used as $\{\phi_k\}$ in (1), while they are not orthogonal.

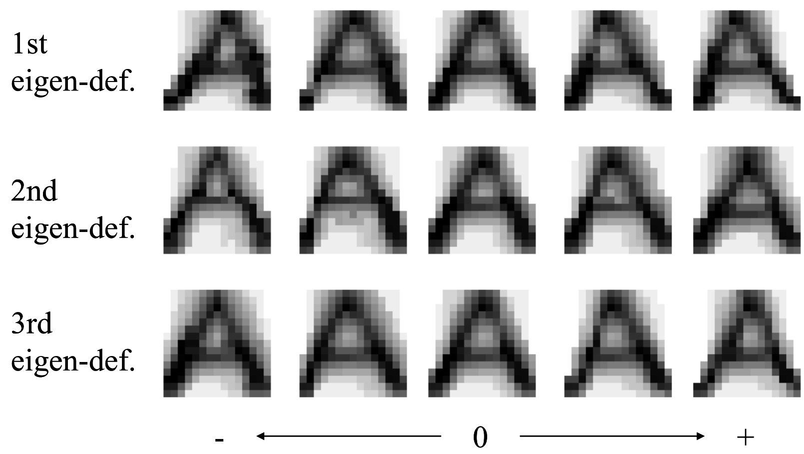

Fig. 5 Top three eigen-deformations of "A."

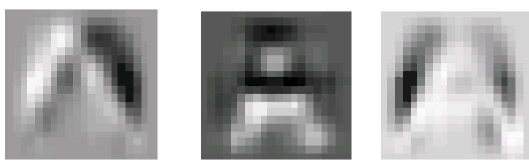

Fig. 6 Tangent vectors for the top three eigen-deformations of Fig. 5.

Non-Parametric 2DW

Non-parametric 2DW is treated as a set of individual pixel-to-pixel correspondences between two images. This implies that non-parametric 2DW is controlled more directly than parametric 2DW and therefore more flexible than parametric 2DW. The non-parametric 2DWs can be divided into two classes: continuous 2DW and discrete 2DW.

Continuous 2DW

Continuous (and non-parametric) 2DW is formulated as a mapping $\mathbf{F}:(i,j)\in\Re^2 \mapsto (x,y)\in\Re^2$, that is a 2D-2D function. Consequently, the optimization problem of the continuous 2DW is often formulated as a variational problem; a continuous 2DW $\mathbf{F}$ is considered as an argument function of the underlying variational problem and $J_{\mathbf{A},\mathbf{B}}(\mathbf{F})$ is the functional to be minimized.

The variational problem is often solved by the deterministic relaxation (Sakaue, Amano, & Yokoya, 1999). Specifically, the Euler-Lagrange equation is firstly derived and discretized to obtain a system of nonlinear equations. Then a sub-optimal 2DW is obtained by solving the equations by some iterative method, such as the Gauss-Seidel method. Mizukami (1998) has successfully applied the deterministic relaxation to handwritten character recognition.

Regularization techniques and/or coarse-to-fine strategies are necessary for solving the variational problem based on the deterministic relaxation. This is because the Euler-Lagrange equation is only a necessary condition for the optimal solution of the variational problem. In addition, since the system of equations is nonlinear, the Gauss-Seidel method cannot guarantee even its convergence. Mizukami (1998) has employed a regularization technique and a careful coarse-to-fine strategy.

Another interesting approach is to relate continuous 2DW to physical phenomena. Webster and Nakagawa (1997) and Nakagawa et al. (1999) 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 membrane showing $\mathbf{B}_{\mathbf{F}}$ is updated iteratively by calculating its motion equation until an equilibrium state. For EM of medical images, we can find different physical formulations (Bajscy & Broit, 1982, Christensen, Rabbitt, & Miller, 1996).

As seen above, continuous 2DW has often been assumed as a differentiable function and optimized by some iterative optimization strategy. In this sense, continuous 2DW is similar to several parametric 2DWs.

Discrete and Unconstrained 2DW

The discrete 2DW is formulated as a set of $2N^2$ variables $((x_{1,1}, y_{1,1}), \ldots, (x_{i,j}, y_{i,j}), \ldots, (x_{N,N}, y_{N,N}))$ where $(x_{i,j}, y_{i,j})$ denotes the pixel on $\mathbf{B}$ corresponding to the pixel (i, j) on $\mathbf{A}$. The discrete 2DW is further divided into two classes: unconstrained 2DW and constrained 2DW. In unconstrained 2DW, the mapping of the pixel (i, j), i.e., $(x_{i,j}, y_{i,j})$, is independent of the mapping of other pixels.

Since there is no constraint among pixels, it is possible to determine the optimal $(x_{i,j}, y_{i,j})$ for each pixel (i, j) independently. This pixel-independent optimization strategy is called local perturbation (Burr, 1981, Hattori, et al., 1983; Izui, et al., 1985; Liolios, et al., 2002; Meguro, & Umeda, 1978; Saito, Yamada, & Yamamoto, 1982; Yamada, Saito, & Mori, 1981), or image distortion model (Keysers, Dahmen, Theiner, & Ney, 2000; Keysers, Gollan, & Ney, 2004). For each pixel (i, j) on $\mathbf{A}$, its best corresponding pixel $(x_{i,j}, y_{i,j})$on $\mathbf{B}$ is searched for locally and independently. The great merit of this simplest optimization strategy is its far less complexity than other optimization strategies.

Local perturbation, however, 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 (Izui, Harashima, & Miyagawa, 1985; Meguro, & Umeda, 1978), smoothing of local displacements (Burr, 1981, Hattori, Watanabe, Sanada, & Tezuka, 1983), sequential (outside-to-inside) optimization with mild constraints (Hattori, Watanabe, Sanada, & Tezuka, 1983), and/or sophisticated pixel features (Saito, et al., 1982; Keysers, Gollan, & Ney, 2004) will be indispensable to expect sufficient performance.

Discrete and Constrained 2DW

In discrete and constrained 2DW, the mapping of the pixel (i, j) is constrained by the mapping of adjacent pixels of (i, j) for regulating flexibility. For example, continuity constraints, such as $\left|x_{i,j} - x_{i-1,j}\right| \le \Delta xi$ where Δxi is a positive small constant, are often imposed on 2DW to exclude large gaps from 2DW. The four parameters Δxi, Δxj, Δyi, and Δyj of Fig. 7 are often used for specifying the constraints, that is, for specifying the flexibility of the 2DW. Monotonicity constraints, such as $x_{i,j} - x_{i-1,j} \ge 0$, are also popular constraints to exclude fold-over from 2DW.

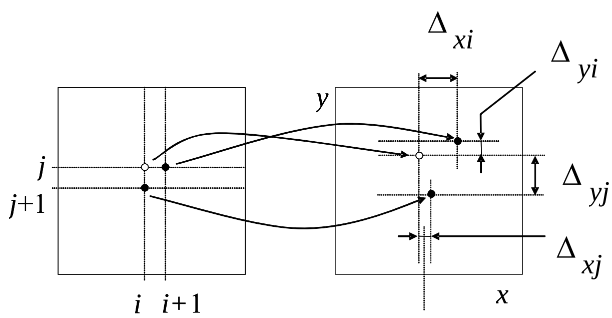

Fig. 7 Parameters for specifying the constraints of 2DW.

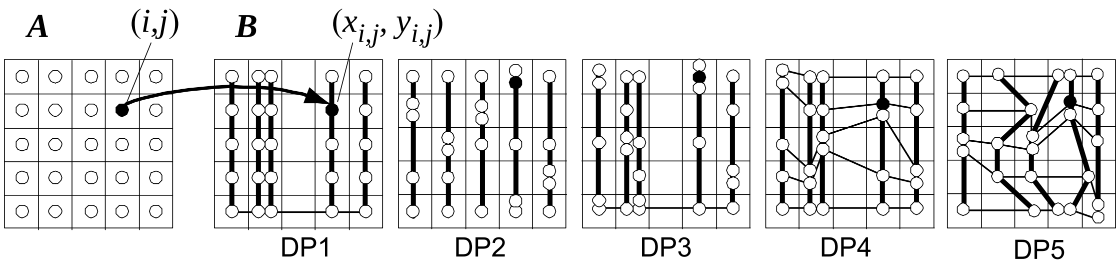

Fig. 8 Several types of discrete and constrained EM.

Table 1 Property of **discrete and constrained EMs**, DP1-DP5.

	Constraints	Computational Complexity by DP	Note
DP1	$\begin{cases} 0 \le x_{i,j} - x_{i-1,j} \le \Delta_{xi}, \\ x_{i,j} = x_{i,j-1}, \\ y_{i,j} = j. \end{cases}$	Polynomial $O\left(N^2\left(N + \Delta_{xi}\right)\right)$ $\approx O\left(N^3\right)$	
DP2	$\begin{cases} x_{i,j} = i, \\ 0 \le y_{i,j} - y_{i,j-1} \le \Delta yj. \end{cases}$	Polynomial $O\left(N^3 \Delta_{yj}\right)$	
DP3	$\begin{cases} 0 \le x_{i,j} - x_{i-1,j} \le \Delta_{xi}, \\ x_{i,j} = x_{i,j-1}, \\ 0 \le y_{i,j} - y_{i,j-1} \le \Delta_{yj}. \end{cases}$	Polynomial $O\left(N^2\left(N^2 \Delta_{yj} + \Delta_{xi}\right)\right)$ $\approx O\left(N^4 \Delta_{yj}\right)$	Often referred as "pseudo-2D"
DP4	$\begin{cases} 0 \le x_{i,j} - x_{i-1,j} \le \Delta_{xi}, \\ x_{i,j} = x_{i,j-1}, \\ \left\lvert y_{i,j} - y_{i-1,j} \right\rvert \le \Delta_{yi}, \\ 0 \le y_{i,j} - y_{i,j-1} \le \Delta_{yj}. \end{cases}$	Exponential $O\left(N^2 \Delta_{yj}^N \left(N + \Delta_{xi} \Delta_{yi}^N\right)\right)$ $\approx O\left(N^2 \Delta_{xi} \left(\Delta_{yi} \Delta_{yj}\right)^N\right)$	
DP5	$\begin{cases} 0 \le x_{i,j} - x_{i-1,j} \le \Delta_{xi}, \\ \left\lvert x_{i,j} - x_{i,j-1} \right\rvert \le \Delta_{xj}, \\ \left\lvert y_{i,j} - y_{i-1,j} \right\rvert \le \Delta_{yi}, \\ 0 \le y_{i,j} - y_{i,j-1} \le \Delta_{yj}. \end{cases}$	Exponential $O\left(N^2 \left(\Delta_{xj} \Delta_{yj}\right)^N \left(N + \left(\Delta_{xi} \Delta_{yi}\right)^N\right)\right)$ $\approx O\left(N^2 \left(\Delta_{xi} \Delta_{xj} \Delta_{yi} \Delta_{yj}\right)^N\right)$	Truly 2D

According to the constraints imposed on the adjacent pixels, previous discrete and constrained 2DWs can be classified into DP1-DP5 of

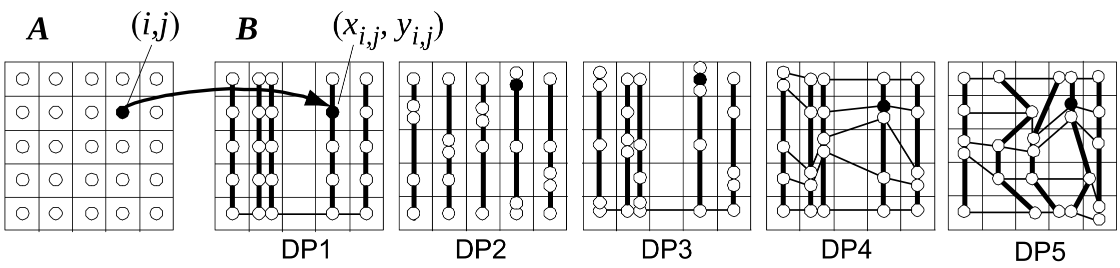

Fig. 8. Each node of the graph on **B** represents the pixel $\left(x_{i,j}, y_{i,j}\right)$ corresponding to

a pixel (i, j) on $\mathbf{A}$. Each link represents the dependency between the mappings of adjacent pixels on $\mathbf{A}$ (e.g., the dependency between $(x_{i,j}, y_{i,j})$ and $(x_{i-1,j}, y_{i-1,j})$). As shown in the figure, 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 $\mathbf{B}$.

Note that a discrete and constrained 2DW other than DP1-DP5 has been proposed in (Moore, 1979; Tanaka, 1985; Wu, et al., 1995). This 2DW is optimized diagonally from one corner $(1,1)$ to its opposite corner (N, N) while extending a rectangular region being processed. Despite of its computational feasibility, its flexibility does not seem to match actual deformations of handwritten characters.

Table 1 summarizes the constraints specifying the 2DWs, DP1-DP5. DP1-DP4 are specified by asymmetric constraints; the constraints in vertical direction is different from those in horizontal direction. Thus, DP1-DP4 are not fully two-dimensional 2DWs and cannot compensate truly 2D deformations, such as rotation and slant. This fact can be understood from the fact that all the pixels on each column of $\mathbf{A}$ are mapped together to the same column of $\mathbf{B}$. In contrast, DP5 is specified by symmetric constraints and thus fully two-dimensional. Note that every type except for DP5 has its transposed (i.e., rotated 90 degrees) version.

Table 1 also summarizes the computational complexity to obtain the optimal 2DW when dynamic programming (DP) is used as the optimization strategy. DP has been employed for many discrete and constrained 2DWs because of its useful properties. DP can guarantee globally optimal 2DW and free from numerical errors. Those properties imply high accuracy of the 2DW optimized by DP. In addition, DP accepts undifferentiable objective functions, position-dependent constraints, and various pixel features. Furthermore, DP framework can be readily extended to be HMM. (Note that HMM and DP are not distinguished in this section unless otherwise mentioned.) Those versatility and extension ability are other useful properties of DP.

DP1 is the simplest 2DW and can compensate simple (and intrinsically 1D) deformations where all the pixels of each column are shifted equally and horizontally. DP1 itself has been employed in word recognition (Cho, Lee, & Kim, 1995; Makhoul, et al., 1998; Mohamed, & Gader, 1996; Shiku, et al., 2000) rather than isolated handwritten character recognition.

While the optimization of DP1 requires fewer computations than others, DP1 cannot compensate any vertical deformation. Thus, DP1 is often repeated to compensate deformations that are more complex. Nakano et al. (1973) have proposed a DP-based EM technique where the vertical version of DP1 is optimized after the horizontal version is optimized. Hallouli et al. (2002) have compared several combinations of vertical and horizontal versions of DP1 in the framework of HMM. In Nishimura, et al. (2001), DP1 is repeated four times under different feature vectors having different roles on representing spatial distribution of strokes. Wang et al. (2001) 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.

DP2 is comprised of independent one-dimensional vertical warpings. DP2 cannot compensate any horizontal deformation. Thus, DP2 is often repeated like DP1. In Isomichi, & Ogawa (1975), the vertical version of DP2 is optimized after the horizontal version is optimized. Tsukumo (1992) has proposed a smart technique where a blurring operation is employed to complement the compensation ability of DP2.

DP3 is the most popular 2DW among discrete and constrained 2DWs. DP3 can compensate both vertical and horizontal deformations simultaneously with polynomial-order computations. The HMM version of DP3 is so-called Pseudo-2D HMM (Agazzi, et al., 1993; Kuo, & Aggazi, 1994; Levin, & Pieraccini, 1992) and widely used in recognizing handwritten characters (Levin, & Pieraccini, 1992), machine-printed words (Agazzi, et al., 1993; Kuo, & Aggazi, 1994; Yen, Kuo, & Lee, 1999), and handwritten words (Bippus, & Margner, 1999).

DP3 has been extended by Keysers, Gollan, & Ney (2004). 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. Large perturbation should be carefully used because it may lose continuity and monotonicity.

DP4 can realize a topology-preserving 2DW since it is constrained in both of vertical and horizontal direction at each pixel. Thus, DP4 can avoid the overfitting of ``P'' to ``b'' by the discontinuity between two strokes. (This overfitting can happen in DP2 and DP3.)

The computational amount of DP4 is an exponential order of N and thus far larger than DP1-DP3. This is because it is impossible to optimize the mapping of each column independently; that is, the mutual relation between the mappings of adjacent columns should be considered during the optimization. DP4 is a restricted version of DP5 and therefore its algorithm can be easily derived from DP5 (Uchida & Sakoe, 1998; Uchida & Sakoe, 1999).

DP5, which is a truly 2D 2DW, has originally proposed by Levin and Pieraccini (1992). In their technique, Δxi, Δxj, Δyi, and Δyj are set at their maximum value, N. Thus, their 2DW can preserve upper/lower and left/right relationships (i.e., vertical and horizontal monotonicity) and does not care about continuity of character patterns. In other words, their 2DW allows large discontinuities and thus is not a topology-preserving 2DW. (Theoretically, their 2DW can map “A” to “H” by losing the continuity around the top of “A.”) Inspired by Levin and Pieraccini (1992), Uchida and Sakoe (1998; 1999) have proposed a topology-preserving and truly 2D 2DW (i.e., truly rubber-sheet EM) where $\Delta xi = \Delta yj = 2$ and $\Delta xj = \Delta yi = 1$.

Although DP5 has a good potential to realize truly rubber-sheet EM, its optimization of DP5 is an NP-hard problem (Keysers, & Unger, 2003) and thus requires exponential-order computations. Even if character images are small ($N \approx 20$), it is impossible to obtain the globally optimal 2DW. Thus, some approximation should be introduced for the practical use of DP5. In Uchida & Sakoe (1998; 1999), 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 Sugimura, et al. (1997), the optimization of DP5 is performed as a sequential (i.e., column-by-column) and greedy manner. In Chen and Willson (2000), this sequential and greedy optimization is iterated to refine the result. In Quenot (1992), the iteration proceeds alternately in horizontal and vertical directions. Uchida and Sakoe (2000a) have proposed an approximation algorithm which exploits the fact that the global optimization by DP can be done very fast if an image pattern is elongated one.

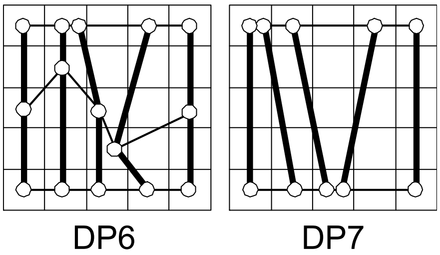

Fig. 9 Piecewise linear 2DW

Hybrid between Parametric 2DW and Non-Parametric 2DW

The boundary between parametric 2DW and non-parametric 2DW is not strict and, in fact, their hybrids have been proposed. A popular way to realize the hybrids is to employ piecewise linearity, or local linearity. DP6 and DP7 (Uchida & Sakoe, 2000b; Ronee, et al, 2001) of Fig. 9 is a piecewise linear version of DP5. In DP6, the mapping of each column (i.e., the mapping of N pixels) is represented as line segments whose control points are fewer than N. DP7 employs more drastic linearization that the column is mapped as a line. The merit of the DP6 and DP7 is that they can compensate 2D deformations (unlike DP1-DP4) with polynomial-order computations. (For DP6 and DP7, the optimization by DP requires $O(N^6)$ and $O(N^4)$ computations.)

Local affine transformation (LAT) proposed by Wakahara (Wakahara, 1994) is also a hybrid based on local linearization. In LAT, 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 (2003b) is also a hybrid 2DW and different from local/piecewise linearization. This 2DW is a parametric and orthogonal 2DW where the principal components of within-category deformations, called eigen-deformations, are used as orthogonal functions. The eigen-deformations themselves, however, are estimated from the results of some non-parametric 2DW. 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.

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 the global deformations. Independent and partial changes of stroke direction, curvature, and length are the examples of the local deformations. EM should compensate both deformations. Since the parametric 2DW and the non-parametric 2DW are suitable for compensating the global deformations and the local deformations respectively, their cooperative combination will be promising.

RELATED TOPICS

Comparison with Shape Normalization

Shape normalization (Lee & Park, 1994; Liu, Nakashima, Sako, & Fujisawa, 2004) is another strategy to providing a deformation-invariant distance. 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 (Yamada, Yamamoto, & Saito, 1990) is a nonlinear shape normalization technique and can adjust nonlinear deformations.

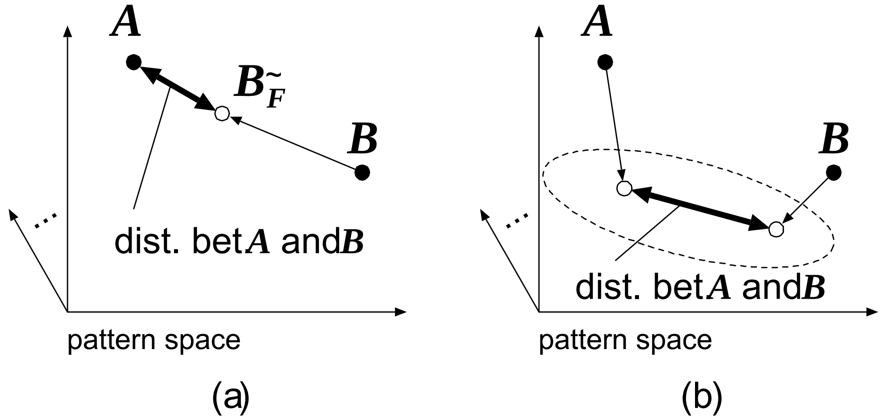

Fig. 10 Distance between A and B given by EM (a) and normalization (b).

Fig. 10 illustrates the difference between shape normalization and EM. As shown in Fig. 10(a), EM shifts $\mathbf{B}$ to a pattern close to $\mathbf{A}$. In contrast, as shown in Fig. 10(b), shape normalization shifts $\mathbf{A}$ to a pattern having some ideal property, and shifts $\mathbf{B}$ independently to another pattern having the same property. In general, EM has larger ability in compensating deformations. 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.

EM and shape normalization can be utilized in a collaborative manner like the collaboration of parametric 2DW and non-parametric 2DW discussed above. For example, shape normalization is applied first to remove global and simple deformations and then EM is applied to remove local and complicated deformations. This collaboration may suppress the overfitting by EM. Tsukumo's EM technique based on DP2 (Tsukumo, 1992) is another good example where blurring normalization is employed to complement the less flexibility of DP2.

Reference Patterns for EM

EM-based recognition can be considered multiple reference-based recognition since EM generates a reference pattern $\mathbf{B_F}$ adaptively to the input pattern $\mathbf{A}$. Thus, the seed of the adaptive generation, i.e., the original reference pattern $\mathbf{B}$ should be carefully designed. Most of EM-based recognizers, however, do not pay much attention to this point; several patterns manually designed/selected are often used as $\mathbf{B}$, or all of training patterns are directly used as a set of $\mathbf{B}$ (e.g., Jain, & Zongker (1997), Simard, et al. (1993)).

Like other recognizers, clustering will be promising to set reference patterns in a sophisticated manner. Actually, any clustering technique (e.g., k-means, ISODATA, LVQ, and GLVQ) can be used for EM. Matsumoto et al. (2004) have pointed out that the objective function of clustering should be designed using the EM distance instead of the conventional Euclidean distance. This is because the reference patterns optimized under a Euclidean distance-based clustering are not optimal for EM distance-based discrimination. (See also Fig. 3.) According to this fact, Matsumoto, Uchida, & Sakoe (2004) have proposed a k-means algorithm based on the EM distance. Another trial can be found in Hastie (1995) where clustering based on the tangent distance has been proposed.

Pixel Feature

In general, less ambiguous pixel feature vector becomes, more accurate the optimized 2DW becomes. In the ultimate case that only a pair of pixels corresponding truly has the same pixel feature vector, local perturbation (the most naïve optimization strategy) is sufficient to obtain optimal 2DW. Conversely, more ambiguous pixel feature vector becomes, less accurate the optimized 2DW becomes. For example, if we try to find 2DW with binary pixel feature (black/white), it is very hard to obtain a reasonable 2DW.

Local context (Keysers, Gollan, & Ney, 2004) and directional feature (Mizukami, 1998; Uchida & Sakoe, 1999) are simple and reasonable choices as less ambiguous pixel feature. Those shape-sensitive features, however, face a problem of inconsistency. As shown in Fig. 11, those feature vectors are often not the same at corresponding pixel pairs. Local contexts, each of which is a subimage, are different at the pixels corresponding truly.

There are two possible remedies for this problem of inconsistency. One is the adaptation of the feature vectors according to 2DW. For example, under the 2DW showing the rotation of angle θ, the directional feature becomes consistent by shifting the original directional feature by θ. The other is the use of invariant features, such as moments.

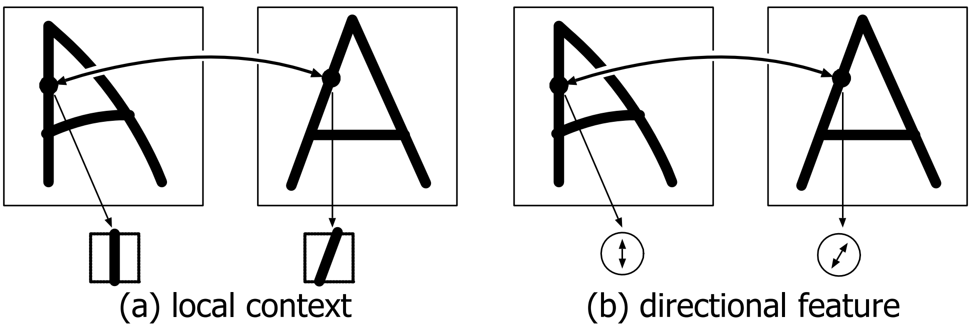

Fig. 11 Inconsistency in pixel features.

Category-Dependent Deformation Tendency

Each category has its own deformation tendency. Fig. 12 is a simple example that proves its existence. In category "M," two parallel vertical strokes are often slanted to be closer (Fig. 12 (a)). The same deformation, however, is rarely observed in "H"(Fig. 12(b)). Consequently, if an EM technique based on the assumption can compensate the deformations of "M," it may suffer from the overfitting of "H" to "A."

Fig. 12 Example of category-dependent deformation tendency.(copyright@2005IEICE)

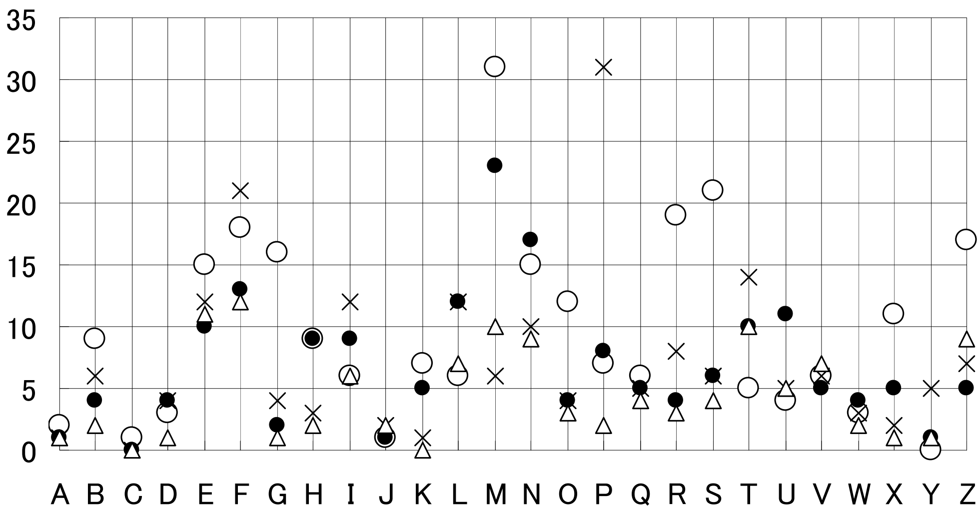

Fig. 13 The number of misrecognized samples by DP1 (○), DP3 (●), DP5 (×), and DP6 (Δ).

The necessity of the category-dependent EM techniques indicated by this simple example will be confirmed by the result of a character recognition experiment. Six hundred handwritten character samples of ETL6 (http://www.is.aist.go.jp/etlcdb/) were prepared for each of 26 categories of English alphabets. Four-dimensional directional features (0, 45, 90, and -45 degrees) were extracted at each pixel. Then they are combined with gray-level feature to be a five-dimensional pixel feature vector. Each sample was linearly scaled to 20×20. 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 $\mathbf{A}$. The EM techniques based on 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 and DP was used as their optimization strategy.

The necessity of the category-dependent EM techniques is proved by the category-wise result of this experiment shown in Fig. 13; that is, 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 rigid 2DW (DP1) 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.

EM TECHNIQUES BASED ON 1D-2D MAPPING

Since any character is a "linear" pattern, it is also natural to use a one-dimensional stroke model instead of a two-dimensional image model. When we use these 1D-stroke models, our task becomes the fitting of the 1D model onto a 2D input pattern. This task is an optimization problem of a 1D-2D mapping function between the 1D model and the 2D input and therefore also a kind of EM.

The 1D-2D EM techniques can be classified by their stroke models. One stroke model will be defined as a sequence of x-y coordinates, and another will be defined as a sequence of line segments. More generally, the model will be defined as a sequence of states each of which represents a local shape of a stroke.

Rubber-string (RS) matching (Sakoe 1974) is one of the most classical 1D-2D EM techniques. As shown in Fig. 14, its stroke model is a sequence of line segments. The direction of the line segment is fixed but the length is flexible. Thus, the optimal fitting problem is reduced to the optimization problem of the lengths of the line segments (and the initial point). In Sakoe (1974), a DP-based algorithm has been proposed for solving the optimization problem. In Sakoe, Ali, & Katayama (1994), RS matching has been extended to increase its ability to represent fine deformations within line segments.

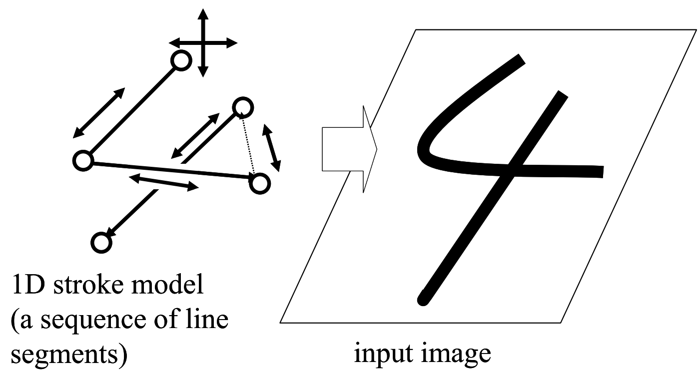

Fig. 14 1D-2D EM technique with a 1D stroke model.

The stroke model by Revow, Williams, & Hinton (1996) is more elaborated one. It is a stochastic model and represents stroke shape variations by a sequence of Gaussian ink generators. The ink generator outputs ink dots according to a probabilistic distribution. The similarity between the model and the input image is evaluated as a likelihood of the input image by the model and calculated through an expectation-maximization estimation of the location and the variance of the ink generators. This approach has been rearranged in a Bayesian framework in Cheung, Yeung, & Chin (1998). Kato, Omachi, & Aso (2000) have introduced a multi-resolution framework for dealing with characters with heavier deformations.

A weak-point of the 1D-2D EM is its embedment problem; the stroke model of the digit "1" may be fitted (or embedded) perfectly to character images with a vertical line, e.g., "4," "5", "7", and "9". Once the stroke model is embedded into an input character image, the fitting is evaluated only around the stroke model and the input character image may be considered as a similar image to the model. Hastie & Tibshirani (1994) called this situation *uncovered*. One possible remedy is a posteriori evaluation of a residual part on the input image. For example, we will have a “c”-shaped residual part on fitting the model “1” to an input image “9,” and therefore we can avoid the misrecognition by penalizing the residual part. Another remedy is the regulation of global deformation of the stroke model. Revow, Williams, & Hinton (1996) employs affine transformation to fit their model roughly to the input image.

We can find 1D-2D EM techniques in other formulations, e.g., contour matching techniques by Yamada (1984) and Yamamoto, & Rosenfeld (1982) and thinned pattern matching proposed by Fujimoto, et al. (1976). The ASM-based method (Shi, et al. 2003) is also a 1D-2D EM technique, where a chamfer matching distance is employed to evaluate the fitting between a stroke model and a binary character image. Kobayashi et al. (2001) has been proposed a SNAKES-like deformable model where optimal fitting which gives minimum energy is searched for in an iterative manner.

FUTURE TASKS

We still have many tasks to utilize EM techniques in practical problems (such as commercial OCR for handwritten characters). The following is a list of those tasks.

Reduction of computational complexity.
Generally speaking, EM requires a fair amount of computations. This fact is crucial especially for a character set 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.

Design of 2DW by category-dependent characteristics.
As proven by the experiment, category-dependent EM is very promising. Statistic and/or stochastic frameworks will help to realize it. Discriminative learning of 2DW will be useful to suppress the misrecognitions due to overfitting. Kernel machines are also promising partners.

Multi-step deformation compensation.

As discussed, we can apply parametric 2DW to compensate global deformations and then non-parametric 2DW to compensate local and complicated deformations. This two-step framework may reduce the degradation by overfitting. Shape normalization also compensates global deformations and therefore it is useful to narrow the range of 2DW.

Feature extraction.

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. Adaptation of pixel features for consistent correspondence is still open problem.

EM for handwritten word recognition.

EM techniques employed in word recognition are rather simple like DP1. 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.

It is important to note that the optimized 2DW represents the deformation of $\mathbf{B}$ relative to $\mathbf{A}$. This fact implies that EM is one of structural analysis techniques for image patterns. The utilization of the optimized 2DW is very promising to extract various properties of handwritten characters.

EM for Camera-based character recognition.

Camera-based character recognition is a recent research trend with many open problems (Liang, Doermann, & Li, 2005). Due to an oblique camera angle, each character undergoes perspective transformation. Due to a non-flat document surface, each character undergoes nonlinear geometric transformation as well as photometric deformation. Past trials to tackle those problems are called *dewarping* and treated as preprocessing techniques. Although they are superficially different from EM techniques reviewed in this paper, the various techniques to optimize 2DW will be applicable to the dewarping problem. In fact, the document dewarping method by Ezaki, Uchida, &

Sakoe (2005) relies on a DP-based 2DW optimization technique.

EM for Kernel Machines.

Recently, EM techniques have been linked with kernel machines, since the distance by EM can be considered as a nonlinear kernel between two patterns. Especially, the DP-based EM techniques are strongly related to string kernels based on the edit distance, or Levenshtein distance, which is often provided by DP. The use of EM will extend the horizon of kernel machines so that they can deal with a set of patterns with different dimensionalities.

CONCLUSION

This chapter reviewed elastic matching (EM) techniques. Since the distance provided by EM is invariant to a certain range of geometric deformations, EM has been employed in handwritten character recognition tasks. EM is defined as the optimization problem of two-dimensional warping (2DW) which specifies 2D-2D mapping between two image patterns. In discrete case, 2DW specifies pixel-to-pixel correspondence between them.

This chapter also showed that EM techniques can classify by two factors; the formulation of 2DW and the optimization strategy of 2DW. Those factors, actually, are strongly related to each other. That is, each kind of 2DW has its appropriate optimization strategy.

As noted in the last section, there remain many open problems and future work in the application of EM to handwritten character/word recognition. Further researches tackling these problems will be very meaningful from not only theoretical but also practical viewpoints.

Acknowledgement: The author greatly thanks to Prof. Hiroaki Sakoe for his thoughtful, sincere, and valuable advices throughout my research carrier.

REFERENCES

Agazzi, O., Kuo, S., Levin, E., & Pieraccini, R. (1993). Connected and degraded text recognition using planar hidden Markov models. Proc. ICASSP, V113-116.

Bippus R. & Märgner, V. (1999). Script recognition using inhomogeneous P2DHMM and hierarchical search space reduction. Proc. ICDAR, 773-776.

Burr, D. J. (1981). A dynamic model for image registration. Computer Graphics, and Image Proc., 15, 102-112.

Bajscy, R. & Broit, C. (1982). Matching of deformed images. Proc. ICPR, 351-353.

Chen, M. C. & Willson, A. N., Jr. (2000). Motion-vector optimization of control grid interpolation and overlapped block motion compensation using iterated dynamic programming. IEEE Trans. Image Proc., 9(7), 1145-1157.

Cheung, K.-W., Yeung, D.-T., & Chin, R. T. (1998). A Bayesian framework deformable pattern recognition with application to handwritten character recognition, IEEE Trans. Pattern Anal. Mach. Intell., 20(12), 1382-1388.

Cho, W., Lee, S.-W., & Kim, J. H. (1995). Modeling and recognition of cursive words with hidden Markov models. Pattern Recognit., 28(12), 1941-1953.

Christensen, G. E., Rabbitt, R. D., & Miller, M. I. (1996). Deformable templates using large deformation kinematics. IEEE Trans. Image Processing, 5(10), 1435-1447.

Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active shape models - their training and application. Comput. Vis. Image Und., 61(1), 38-59, Jan. 1995.

Ezaki, H., Uchida, S., & Sakoe, H. (2005). Dewarping of document image by global optimization. Proc. ICDAR, 1, 302-306.

Fujimoto, Y., Kadota, S., Hayashi, S., Yamamoto, M., Yajima, S., & Yasuda, M. (1976).

Recognition of handprinted characters by nonlinear elastic matching. Proc. ICPR, 113-118.

Glasbey, C. A. & Mardia, K. V. (1998). A review of image-warping methods. J. Applied Statistics, 25(2), 155-171.

Ha, T. M. & Bunke, H. (1997). Off-line, handwritten numeral recognition by perturbation method. IEEE Trans. Pattern Anal. Mach. Intell., 19(5), 535-539.

Halloui, K., Likforman-Sulem, L., & Sigelle, M. (2002). A comparative study between decision fusion and data fusion in Markovian printed character recognition. Proc. ICPR, 3 of 4, 147-150.

Hastie, T., Simard, P. Y., & Säckinger, E. (1995). Learning prototype models for tangent distance. Advances in Neural Information Processing Systems, 7, 999-1006.

Hastie, T. J, & Tibshirani, R. (1994). Handwritten digit recognition via deformable prototypes. AT&T Bell Laboratories Technical Report.

Hattori, T., Watanabe, Y., Sanada, H., & Tezuka, Y. (1983). Absorption of local variations in handwritten character by an elastic transformation using vector field. IEICE Trans., J66-D(6), 645-652. (in Japanese)

Isomichi, Y. & Ogawa, T. (1975). Pattern matching by using dynamic programming. J. Inf. Process. Soc. Japan, 16(1), 15-22. (in Japanese)

Izui, Y., Harashima, H., & Miyagawa, H. (1985). Handprinted Chinese characters recognition by hierarchical modification of dictionary. IEICE Trans., J68-D(3), 361-368. (in Japanese)

Jain, A. K., Zhong, Y., & Dubuisson-Jolly, M.-P. (1998). Deformable template models: A review. Signal Processing, 71(2), 109-129.

Jain, A. K. & Zongker, D. (1997). Representation and recognition of handwritten digits using deformable templates. IEEE Trans. Pattern Anal. Mach. Intell., 19(12), 1386-1391.

Kato, T., Omchi, S., & Aso, H. (2000). Precise hand-printed character recognition using elastic models vianonlinear transformation, Proc. ICPR, 2, 364-367.

Keysers, D., Dahmen, J., Theiner, T., & Ney, H. (2000). Experiments with an extended tangent distance. Proc. ICPR, 2 of 4, 38-42.

Keysers, D., Gollan, C., & Ney, H. (2004). Local context in non-linear deformation models for handwritten character recognition. Proc. ICPR, 4 of 4, 511-514.

Keysers, D., Macherey, W., Ney, H., & Dahmen, J. (2004). Adaptation in statistical pattern recognition using tangent vectors. IEEE Trans. Pattern Anal. Mach. Intell., 26(2), 269-274.

Keysers, D. & Unger, W. (2003). Elastic image matching is NP-complete. Pattern Recog. Lett., 24(1)-3, 445-453.

Kimura, F., Yoshimura, M., Miyake, Y., & Ichikawa, M. (1970). Unconstrainedly handprinted ``KATAKANA" character recognition by a stroke structure analysis method. Trans. IEICE, J62-D(1), 16-23. (in Japanese)

Kobayashi, T., Nakamura, K., Muramatsu, H., Sugiyama, T., & Abe, K. (2001) Handwritten numeral recognition using flexible matching based on learning of stroke statistics, Proc. ICDAR, 612-616.

Kuo, S. & Agazzi, O. E. (1994). Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models. IEEE Trans. Pattern Anal. Mach. Intell., 16(8), 842-848.

Lee, S.-W. & Park, J.- S. (1994). Nonlinear shape normalization methods for the recognition of large-set handwritten characters. Pattern Recognit., 27(7), 895-902.

Lester, H. & Arridge, S. R. (1999). A survey of hierarchical non-linear medical image registration. Pattern Recognit., 32(1), 129-149.

Levin, E. & Pieraccini, R. (1992). Dynamic planar warping for optical character recognition. Proc. ICASSP, III 149-152.

Liang, J., Doermann, D., & Li, H. (2005). Camera-based analysis of text and documents: a survey. Int. J. Doc. Anal. Recog., 7, 84-104.

Liolios, N., Kavallieratou, E., Fakotakis, N., & Kokkinakis, G. (2002). A new shape transformation approach to handwritten character recognition. Proc. ICPR, 1 of 4, 584-587.

Liu, C. L., Nakashima, K., Sako, H., & Fujisawa, H. (2004). Handwritten digit recognition: investigation of normalization and feature extraction techniques. Pattern Recognit., 37(2), 265-279.

Makhoul, J., Schwaltz, R., Lapre, C., & Bazzi, I. (1998). A script-independent methodology for optical character recognition. Pattern Recognit., 31(9), 1285-1294.

Matsumoto, N., Uchida, S., & Sakoe, H. (2004). Prototype setting for elastic matching-based image pattern recognition. Proc. ICPR, 1 of 4, 224-227.

Meguro, S. & Umeda, M. (1978). An extraction of shape derivations in handwritten characters by hierarchical pattern matching. Tech. Rep. of IEICE Japan, PRL77-70. (in Japanese)

Mizukami, Y. (1998). A handwritten Chinese character recognition system using hierarchical displacement extraction based on directional features.
Pattern Recog. Lett., 19(7), 595-604.

Mohamed, M. & Gader, P. (1996). Handwritten word recognition using segmentation-free
hidden Markov modeling and segmentation-based dynamic programming techniques. IEEE Trans. Pattern Anal. Mach. Intell., 18(5), 548-554.

Moore, R. K. (1979). A dynamic programming algorithm for the distance between two finite areas. IEEE Trans. Pattern Anal. Mach. Intell., PAMI-1(1), 86-88.

Mori, S., Suen, C. Y., & Yamamoto, K. (1992). Historical review of OCR research and development. Proc. IEEE, 80(7), 1029-1058.

Mori, S., Yamamoto, K., & Yasuda, M. (1984). Research on machine recognition of handprinted characters. IEEE Trans. Pattern Anal. Mach. Intell., PAMI-6(4), 386-405.

Nakagawa, M., Yanagida, T., & Nagasaki, T. (1999). An off-line character recognition method employing model-dependent pattern normalization by an elastic membrane model. Proc. ICDAR, 495-498.

Nakata, K., Nakano, Y., & Uchikura, Y. (1972). Recognition of Chinese characters. Proc. Conf. Machine Perception of Patterns and Pictures, 45-52.

Nakano, Y., Nakata, K., Uchikura, Y., & A. Nakajima. (1973). Improvement of Chinese character recognition using projection profiles. Proc. Int. Joint Conf. Pat. Recog., 172-178.

Nishimura, H., Tsutsumi, M., Maruyama, M., Miyao, M., & Nakano, Y., (2001). Off-line handwritten character recognition using integrated 1D HMMs based on feature extraction filters. Proc. ICDAR, 417-421.

Park, H.- S. & Lee, S. W. (1998). A truly 2-D hidden Markov model for off-line handwritten character recognition. Pattern Recognit., 31(12), 1849-1864.

Quénot, G. M. (1992). The orthogonal algorithm for optical flow detection using dynamic programming. Proc. ICASSP, III-249-252.

Redert, A., Hendriks, E., & Biemond, J. (1999). Correspondence estimation in image pairs. IEEE SP Mag., 16(3), 29-46.

Revow, M., Williams, C. K. I., & Hinton, G. E. (1996). Using generative models for handwritten digit recognition. IEEE Trans. Pattern Anal. Mach. Intell., 18(6), 592-606.

Ronee, M. A., Uchida, S., & Sakoe, H. (2001). Handwritten character recognition using piecewise linear two-dimensional warping. Proc. ICDAR, 39-43.

Saito, T., Yamada, H., & Yamamoto, K. (1982). An analysis of handprinted Chinese

characters by directional pattern matching approach.
IEICE Trans., J65-D(5), 550-557. (in Japanese)

Sakaue, K., Amano, A., & Yokoya, N. (1999). Optimization approaches in computer vision and image processing. IEICE Trans. Inf. & Syst., E82-D(3), 534-547.

Sakoe, H. (1974). Handwritten character recognition by rubber string matching method. Tech. Rep. of IEICE Japan, PRL74-20. (in Japanese)

Sakoe, H., Ali M. M., & Katayama, Y. (1994). One dimensional-two dimensional dynamic programming algorithm for character recognition, IEICE Trans. Information & Systems, E77-D(9), 1047-1054.

Shi, D., Gunn, S. R., & R. Damper, I. (2003). Handwritten Chinese radical recognition using nonlinear active shape models. IEEE Trans. Pattern Anal. Mach. Intell., 25(2), 277-280.

Shiku, O., Nakamura, A., Kuroda, H., & Miyahara, S. (2000). A method for handwritten Japanese word recognition based on holistic strategy. Trans. Inf. Process. Soc. Japan., 41(4), 1086-1095. (in Japanese)

Simard, P. Y., Cun, Y. Le., & Denker, J. S. (1993). Efficient pattern recognition using a new transformation distance. Advances in Neural Information Processing Systems, 5, 50-58.

Sugimura, M., Iiguni, Y., & Adachi, N. (1997). A 2-dimensional dynamic programming for image matching with Zernike moments. Trans. IEICE, J80-D-II(1), 101-108. (in Japanese)

Tanaka, E. (1985). A two dimensional context-dependent similarity measure. Trans. IECE Japan, E-68(10), 667-673.

Trier, Ø. D., Jain, A. K., & Taxt, T. (1996). Feature extraction methods for character recognition - a survey. Pattern Recognit., 29(4), 641-662.

Tsukumo, J. (1992). Handprinted Kanji character recognition based on flexible template matching. Proc. ICPR, 483-486.

Uchida, S. & Sakoe, H. (1998). A monotonic and continuous two-dimensional warping based on dynamic programming. Proc. ICPR, 1 of 2, 521-524.

Uchida, S. & Sakoe, H. (1999). Handwritten character recognition using monotonic and continuous two-dimensional warping. Proc. ICDAR, 499-502.

Uchida, S. & Sakoe, H. (2000a). An approximation algorithm for two-dimensional warping. IEICE Trans. Info. & Syst., vol.E83-D(1), 109-111.

Uchida, S. & Sakoe, H. (2000b). Piecewise linear two-dimensional warping. Proc. ICPR, 3, 538-541.

Uchida, S. & Sakoe, H. (2003a). Eigen-deformations for elastic matching based handwritten character recognition. Pattern Recognit., 36(9), 2031-2040.

Uchida, S. & Sakoe, H. (2003b). Handwritten character recognition using elastic matching based on a class-dependent deformation model, Proc. ICDAR, 1, 163-167.

Uchida, S., & Sakoe, H. (2005). A survey of elastic matching techniques for handwritten character recognition. IEICE Transactions on Information & Systems, E88-D(8), 1781-1790.

Wakahara, T. (1994). Shape matching using LAT and its application to handwritten numeral recognition. IEEE Trans. Pattern Anal. Mach. Intell., 16(6), 618-629.

Wakahara, T. & Odaka, K. (1998). Adaptive normalization of handwritten characters using global/local affine transformation. IEEE Trans. Pattern Anal. Mach. Intell., 20(12), 1332-1341.

Wakahara, T., Kimura, Y., & Tomono, A. (2001). Affine-invariant recognition of gray-scale characters using global affine transformation correlation. IEEE Trans. Pattern Anal. Mach. Intell., 23(4), 384-395.

Wang, W., Brakensiek, A., Kosmala, A., & Rigoll, G. (2001). Multi-branch and two-pass HMM modeling approaches for off-line cursive handwriting recognition. Proc. ICDAR, 231-235.

Webster, R. & Nakagawa, M. (1997). An on-line/off-line compatible character recognition method based on a dynamic model. IEICE Trans. Inf. & Syst., E80-D(6), 672-683.

Wu, C.- M., Liu, P., & Chang, W.- C. (1995). Unconstrained-endpoint dynamic space-warping algorithm with experiments in binary English character images. Int. J. Electronics, 78(1), 55-66.

Yamada, H. (1984). Contour DP matching method and its application to handprinted Chinese character recognition. Proc. ICPR, 389-392, 1 of 2.

Yamada, H., Saito, T., & Mori, S. (1981). An improvement of correlation method - locally maximized correlation -. Trans. IEICE, J64-D(10), 970-976. (in Japanese)

Yamada, H., Yamamoto, K., & Saito, T. (1990). A nonlinear normalization method for handprinted Kanji character recognition - line density equalization. Pattern Recognit., 23(9), 1023-1029.

Yamamoto, K. & Rosenfeld, A. (1982). Recognition of hand-printed KANJI characters by a relaxation method. Proc. ICPR, 1 of 2, 395-398.

Yasuda, M., Yamamoto, K., & Yamada, H. (1997). Effect of the perturbed correlation method for optical character recognition", Pattern Recognit., 30(8), 1315-1320.

Yen, C., Kuo, S. S., & Lee, C.-H. (1999). Minimum error rate training for PHMM-based text recognition. IEEE Trans. Image Proc., 8(8), 1120-1124.

Zitová, B, & Flusser, J. (2003). Image registration methods: a survey. Image and Vision Computing, 21, 977-1000.

Biography

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, IEICE, IPSJ, and ITE.