Handwritten character recognition using monotonic and continuous two-dimensional warping

Seiichi Uchida and Hiroaki Sakoe
Graduate School of Information Science and Electrical Engineering, Kyushu University
6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, 812-8581 Japan
{uchida, sakoe}@is.kyushu-u.ac.jp

Abstract

In this paper, a handwritten character recognition experiment using a monotonic and continuous two-dimensional warping algorithm is reported. This warping algorithm is based on dynamic programming and searches for the optimal pixel-to-pixel mapping between given two images subject to two-dimensional monotonicity and continuity constraints. Experimental comparisons with rigid matching and local perturbation show the performance superiority of the monotonic and continuous warping in character recognition.

1. Introduction

One of the central problems in handwritten character recognition is how to deal with the deformations of characters, such as translation, rotation, scaling, and meaningless nonlinear deformation. One promising approach to the problem will be the use of an image distance given by two-dimensional warping (2DW). The 2DW is defined as pixel-to-pixel mapping with a minimum residual error between given two images. This minimum residual error can be considered to be an image distance, or dissimilarity, remaining still after fitting one image to the other image, and is expected to be stable against the above deformations.

The previously reported 2DW techniques for character recognition are partitioned into two classes ; parametric 2DW [1, 2, 3, 4] and nonparametric 2DW. The latter, which potentially has more flexibility than the former, can be partitioned into three classes according to warp optimization strategies, i.e., iterative methods, local perturbation, and DP.

For nonparametric 2DW based character recognition, Mizukami et al. [5] have proposed an iterative algorithm based on a variational method, which has been successfully applied to contour detection [6]. In the iterative algorithm proposed by Yanagida et al. [7], 2DW is obtained by falling an elastic membrane made from an input character image into the bottom of a potential field made from a reference character image in a step-by-step manner until convergence.

Independent and local search for the best mapping of each pixel or sub-image, referred to as *local perturbation*, may be the simplest method for the optimization of nonparametric mapping [8, 9, 10]. While the local perturbation requires far less complexity than the other two class of optimization strategies, the local perturbation often yields excessive deformation even when its search area is limited to a relatively small range.

Several DP algorithms have been developed for handwritten character recognition based on nonparametric 2DW [11, 12, 13, 14, 15], motivated by successful applications of DP to the time warping problem in speech recognition (see e.g., [16]). DP has the following advantages : 1) global optimality of its solution, 2) wide varieties of applicable constraints and criterion functions, 3) computational stability, etc. Since the warps given by these previous DP algorithms are inherently one-dimensional, they lack in flexibility to fit actual deformation, especially rotation.

The authors have investigated the general framework for DP-based *monotonic and continuous 2DW* [17, 18]. The DP algorithm searches for the optimal 2DW subject to two-dimensional monotonicity and continuity constraints employed to preserve topological structure in images.

In this paper, experiments on handwritten character recognition using the monotonic and continuous 2DW is reported. Experimental results on a 3680 Hiragana character image subset of ETL8B show the significant superiority of the 2DW over rigid template matching in recognition accuracy. In addition, the validity of the monotonic and continuous constraints on character recognition is quantitatively analyzed in compari-

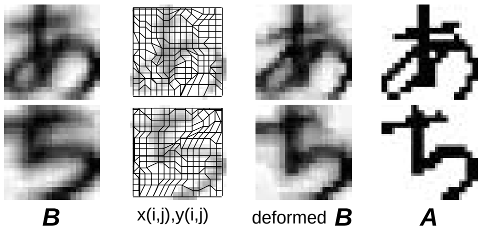


Figure 1. Examples of monotonic and continuous 2DW.

son with recognition results given by local perturbation.

2. Monotonic and continuous 2DW

2.1. Problem formulation and DP algorithm

Consider two character images $\boldsymbol{A} = \{\boldsymbol{a}(i,j) \mid i,j = 1,\ldots,N\}$ and $\boldsymbol{B} = \{\boldsymbol{b}(x,y) \mid x,y = 1,\ldots,N\}$ where pixel values $\boldsymbol{a}(i,j)$ and $\boldsymbol{b}(x,y)$ may be vectors. The optimal 2DW between $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined by the warping function $x = x(i,j)$, $y = y(i,j)$ which minimizes the following criterion function

$$\sum_{i=1}^{N}\sum_{j=1}^{N} \delta(\boldsymbol{a}(i,j),\ \boldsymbol{b}(x(i,j),y(i,j))) \tag{1}$$

where $\delta(\cdot,\cdot)$ is a pixel distance function. The warping function is constrained by the two-dimensional monotonicity and continuity constraints defined as

$$0 \le x(i,j) - x(i-1,j) \le 2 \tag{2}$$
$$0 \le y(i,j) - y(i,j-1) \le 2 \tag{3}$$
$$|x(i,j) - x(i,j-1)| \le 1 \tag{4}$$
$$|y(i,j) - y(i-1,j)| \le 1 \tag{5}$$

and the boundary conditions defined as

$$x(1,j) = y(i,1) = 1 \tag{6}$$
$$x(N,j) = y(i,N) = N. \tag{7}$$

These constraints guarantee the 2DW to approximately preserve the topological structure in images. Figure 1 shows two examples of the 2DW.

Let $D(\boldsymbol{A},\boldsymbol{B})$ denote the minimum value of (1). From the viewpoint of pattern matching, the quantity $D(\boldsymbol{A},\boldsymbol{B})$ is a distance between $\boldsymbol{A}$ and optimally deformed $\boldsymbol{B}$.

The minimization problem of (1) subject to the constraints (2)-(7) can be solved by the following DP algorithm :

[Initialization] for all $\boldsymbol{xy} \in \boldsymbol{XY}(1,N)$

$$g(1,N,\boldsymbol{xy}) = \sum_{j=1}^{N} \delta(\boldsymbol{a}(1,j),\ \boldsymbol{b}(x(1,j),y(1,j)))$$

[Recursion] for all $i(>1)$, j, and $\boldsymbol{xy} \in \boldsymbol{XY}(i,j)$

$$\begin{aligned} g(i,j,\boldsymbol{xy}) \;=\; & \delta(\boldsymbol{a}(i,j),\ \boldsymbol{b}(x(i,j),y(i,j))) \\ & + \min_{\overline{\boldsymbol{xy}} \in \overline{\boldsymbol{XY}}(\boldsymbol{xy})} \begin{cases} g(i-1,j,\overline{\boldsymbol{xy}}) & \text{if } j=1 \\ g(i,j-1,\overline{\boldsymbol{xy}}) & \text{otherwise} \end{cases} \end{aligned}$$

[Termination]

$$D(\boldsymbol{A},\boldsymbol{B}) = \min_{\boldsymbol{xy} \in \boldsymbol{XY}(N,N)} g(N,N,\boldsymbol{xy})$$

where $\boldsymbol{XY}(i,j)$ is the set of possible mapppings of N pixels $[(i-1,j+1),\ldots,(i-1,N),(i,1),\ldots(i,j)]$, and $\overline{\boldsymbol{XY}}(\boldsymbol{xy})$ is the a subset of $\boldsymbol{XY}(i,j-1)$ whose element $\overline{\boldsymbol{xy}}$ satisfies the constraints (2) and (5) for $\boldsymbol{xy} \in \boldsymbol{XY}(i,j)$. For more details, see [17, 18].

2.2. Practical improvements

Since the time complexity of the above DP algorithm are $O(N^3 9^N)$, one has to resort to a polynomial time approximation algorithm at the cost of the optimality. Thus, an approximation algorithm incorporated with beam search technique has been proposed in [17, 18]. In this approximation algorithm, the R smallest cumulative costs $g(i,j,\boldsymbol{xy})$ are taken into account as the active search paths for optimal 2DW, and the remainder is pruned off, at each step i,j. The time complexity of the algorithm are $O(N^3 R)$.

In our experiments, penalty functions (,or stabilizing functional for smoothness) [17] and warp range limitation were employed to avoid unnatural 2DW. Warp range limitation is a simple technique for rejecting large deformation and defined as constraints on the warping function, i.e.,

$$|i - x(i,j)| \le w, \qquad |j - y(i,j)| \le w \tag{8}$$

where w is a positive integer.

3. Experiments and results

3.1. Database and preprocessing

Recognition experiments were conducted on 46 character categories of Japanese Hiragana alphabets (Fig.2). Character image samples used in the experiments were a subset of the handwritten character database offered by the Electrotechnical Laboratory,

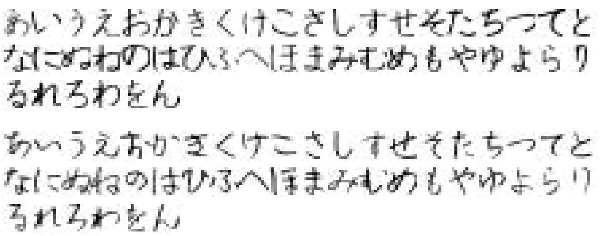


Figure 2. Handwritten Hiragana character set examples from ETL8B. (46 characters × 2.)

ETL8B, which includes 160 samples for each character category.

Each character image sample was preprocessed in the following manner. First, character size was linearly normalized so that its circumscribed rectangular became 64×64. Second, directional feature, defined as the local direction of stroke contour, was detected. Since the directions were quantized into four directions, each pixel was specified by a five dimensional feature vector including intensity level. Third, each sample was scaled so that $N = 16$. Finally, Gaussian blurring and histogram equalization was performed. The reference image $\boldsymbol{B}$ of each category was created by simply averaging 80 preprocessed samples of the category. The remaining 80 samples of each category were used as input images $\boldsymbol{A}$.

We used the pixel distance function defined as

$$\delta(\boldsymbol{a}(i,j),\ \boldsymbol{b}(x,y)) = |a_I(i,j) - b_I(x,y)| + \eta \sum_{k=1}^{4} |a_D^k(i,j) - b_D^k(x,y)|$$

where $a_I(i,j)$ and $b_I(x,y)$ are intensity levels, $a_D^k(i,j)$ and $b_D^k(x,y)$ are directional features, and η is a non-negative weighting coefficient.

By using rigid template matching, the highest recognition rate of 93.9% was attained at $\eta = 0.5$ for the above data set.

3.2. Recognition using monotonic and continuous 2DW

A recognition experiment using the monotonic and continuous 2DW was performed. Based on the experimental result of the previous section, the coefficient η was fixed at 0.5. The approximation algorithm discussed in Section 2.2 was used with the beam size $R = 1000$. Sun Ultra2 (SPECint_95: 12.3, SPECfp_95: 20.2) required about 3 seconds to obtain the 2DW between a pair of images. The minimized criterion function value was directly used as a distance between two character images. Each input image $\boldsymbol{A}$ was classified into the category of the reference image $\boldsymbol{B}$ with the minimum value of this image distance.

Results are shown in Table 1. The highest recognition rate 96.8% was attained when the penalty functions and the warp range limitation ($w = 3$) were used. Comparing rigid template matching and the 2DW at their highest recognition rates, i.e., 93.9% and 96.8%, 131 samples misrecognized by rigid template matching are correctly recognized by the 2DW. Figure 3(a) shows four examples of these samples. It can be seen that the reference images are reasonably deformed to fit the input images by the 2DW. On the other hand, 27 samples correctly recognized by rigid template matching are misrecognized by the 2DW. Figure 3(b) shows four examples. It can be seen that the misrecognition by the 2DW is mainly due to excessive deformation of the reference images belonging to different categories rather than inaccurate warping of the reference images belonging to the same category.

3.3. Comparison with local perturbation

The effect of the monotonicity and continuity constraints on character recognition can be quantitatively analyzed in comparison with local perturbation. The image distance given by local perturbation is defined as the minimum of the criterion function (1) subject to the boundary conditions (6), (7) and the range limitation (8). Since there is no mutual dependence between individual pixel mappings, the minimization of (1) can be decomposed into N^2 independent minimum selections. It is clear that the difference between the monotonic and continuous 2DW and the local perturbation is whether the monotonicity and continuity constraints are present or not.

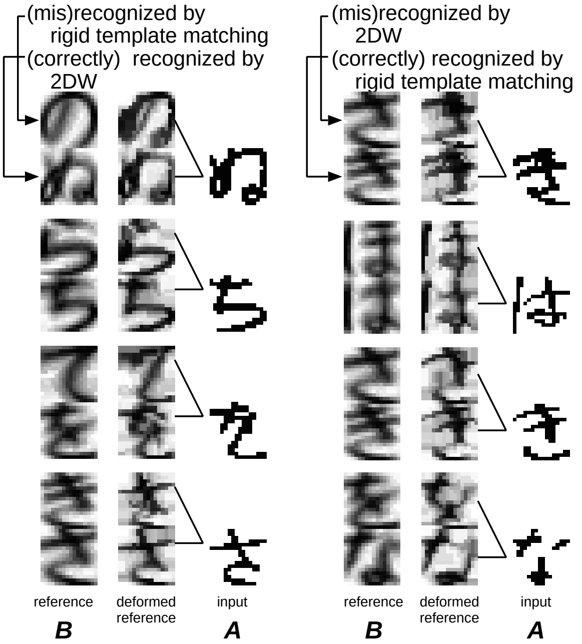
The recognition rates by the local perturbation are shown in Table 1. It can be seen that the 2DW consistently attains higher recognition accuracies than the local perturbation for all warp ranges w. In addition, as the warp range w increases, the recognition accuracy of the local perturbation is significantly degraded, while that of the 2DW is improved . This fact indicates that the monotonicity and continuity constraints suppress unrealistic warp even when the warp range is kept wide enough to fit actual deformation.

4. Conclusions

We experimentally investigated handwritten Hiragana character recognition using the monotonic and continuous two-dimensional warping (2DW) algorithm

Table 1. Recognition rates of the monotonic and continuous 2DW and local perturbation.

warp range w	0 (=rigid matching)	1	2	3	5	∞
Monotonic and continuous 2DW (with penalty)	93.9	95.6	96.5	96.8	96.3	96.2
Monotonic and continuous 2DW (without penalty)		94.8	96.0	95.9	95.2	94.5
Local perturbation		94.3	92.8	90.0	80.6	–



(a) (b)

Figure 3. Examples of samples misrecognized by rigid matching and correctly recognized by monotonic and continuous 2DW (a) , and correctly recognized by the rigid matching and misrecognized by the 2DW (b). For simplicity, directional features of each sample are omitted.

based on DP. The results shows that reference character images are reasonably fitted to input character images by the 2DW, thus attaining higher recognition accuracy than rigid template matching. Experimental comparisons with local perturbation show the validity of the two-dimensional monotonicity and continuity constraints imposed on warping.

Future work will focus on introducing nonuniform elasticity to each reference character image in order to suppress excessive deformation. Further reduction of computational complexity is also to be investigated.

Acknowledgement: The authors thank the Electrotechnical Laboratory for providing us ETL8B. This work was supported in part by the Ministry of Education, Science, Sports and Culture in Japan under a Grant-in-Aid for Scientific Research No.10680385.

References

[1] T. Wakahara. Shape matching using LAT and its application to handwritten numeral recognition. *IEEE Trans. PAMI*, 16(6):618–629, 1994.

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

[3] M. Yasuda, K. Yamamoto, and H. Yamada. Effect of the perturbed correlation method for optical character recognition. *Pattern Recog.*, 30(8):1315–1320, 1997.

[4] T.M. Ha and H. Bunke. Off-line, handwritten numeral recognition by perturbation method. *IEEE Trans. PAMI*, 19(5):535–539, 1997.

[5] Y. Mizukami and K. Koga. A handwritten character recognition system using hierarchical displacement extraction algorithm. *Proc. 13th ICPR*, Vol. 3 of 4, 160–164, 1996.

[6] M. Kass, A. Witkin, and D. Terzopoulos. Snakes : Active contour models. *Proc. ICCV*, 259–268, 1987.

[7] T. Yanagida, T. Nagasaki, and M. Nakagawa. An off-line character recognition method based on an elastic membrane model. *IEICE Tech. Report*, PRMU97-223, 1998. (in Japanese)

[8] S. Meguro and M. Umeda. An extraction of shape deviations in handwritten characters by hierarchical pattern matching. *IEICE Tech. Report, PRL77-70*, 1977. (in Japanese)

[9] H. Yamada, T. Saito, and S. Mori. An improvement of correlation method – locally maximized correlation. *Trans. IEICE*, J64-D(10):970–976, 1981. (in Japanese)

[10] Y. Izui, H. Harashima, and H. Miyakawa. Handprinted Chinese characters recognition by hierarchical modification of dictionary. *Trans. IEICE*, J68-D(3):361–368, 1985. (in Japanese)

[11] Y. Nakano, K. Nakata, and A. Nakajima. Improvements of printed Chinese character recognition using peripheral distributions and their spectra. *Trans. IEICE*, 57-D(1):15–22, 1974. (in Japanese)

[12] N. Tanaka, M. Shiono, H. Sanada, and Y. Tezuka. Recognition of handprinted Kanji characters by dynamic directional matching method. *Trans. IEICE*, J68-D(1):56–63, 1985. (in Japanese)

[13] J. Tsukumo. Handprinted Kanji character recognition based on flexible template matching. *Proc. 11th ICPR*, 483–486, 1992.

[14] E. Levin and R. Pieraccini. Dynamic planar warping for optical character recognition. *Proc. ICASSP*, III 149–152, 1992.

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

[16] H. Sakoe and S. Chiba. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. ASSP*, 26(1):43–49, 1978.

[17] S. Uchida and H. Sakoe. A monotonic and continuous two-dimensional warping based on dynamic programming. *Proc. 14th ICPR*, Vol. 1 of 2, 521–524, 1998.

[18] S. Uchida and H. Sakoe. An efficient two-dimensional warping algorithm. *IEICE Trans. Info. & Syst.*, E82-D(3):693–700, 1999.