Nonuniform Slant Correction Using Dynamic Programming

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

Abstract

Slant correction is an indispensable technique for handwritten word recognition systems. Conventional slant correction techniques estimate the average slant angle of component characters and then correct the slant uniformly. Thus these conventional techniques will perform successfully under the assumption that each word is written with a constant slant. However, it is more widely acceptable assumption that the slant angle fluctuates during writing a word. In this paper, a nonuniform slant correction technique is presented where the slant correction problem is formulated as an optimal estimaiton problem of local slant angles at all horizontal positions. The optimal estimation is governed by a criterion function and several constraints for the global and local validity of the local angles. The optimal local slant angles which maximize the criterion satisfying the constraints are searched for efficiently by a dynamic programming based algrithm. Experimental results show the advantageous characteristics of the present technique over the uniform slant correction techniques.

1. Introduction

Slant correction is an indispensable technique for both holistic and analytical handwritten word recognition. In holistic recognition, slant correction is often employed for coarse shape normalization of the component characters of input words. In analytical recognition, where an input word is segmented into its component characters which are compared to single character reference patterns, slant correction is employed to improve the segmentation peformance which seriously affects recognition accuracy.

In convensional slant correction techniques, the average slant angle of component characters is estimated and then *uniform* correction is performed by shear transformation. Bozinovic and Srihari [1] and Kim and Govindaraju [2] have proposed slant correction techniques where the average slant angle is estimated from the angles of extracted vertical strokes. Guillevic and Suen[3], Kavallieratou et al.[4], and Nicchiotti and Scagliola[5] analyzed a set of projection histograms for the estimation of the average slant angle. Kimura et al.[6], Simoncini and Kovács[7], Ding et al.[8], and Britto et al.[9] utilized statistics of chain-coded stroke contours.

These uniform slant correction techniques will perform successfully under the assumption that each word is written with a constant slant. However, it is a more widely acceptable assumption that the slant angle fluctuates in a word due to various factors such as writer's habit, the inherent shape of each character, and writing position. This assumption raises the necessity to estimate local slant angles and to correct them *nonuniformly*. This necessity was justified by Britto et al.'s result[9] that there are significant gaps between the average and the individual slant angles of component characters, which degrade the performance of analytical recognition systems.

One could propose a naive nonuniform slant correction technique where a word image is divided into small regions with a certain width and then the average slant angle of each region is independently estimated as a local slant angle. It is clear, however, that this naive technique provides poor results since the estimated local slant heavily depends on the shape of character, character fragement, or their mixture contained in the region. This suggests that the nonuniform slant correction shoud be performed consistently by evaluating global goodness of locally estimated slant angles.

In this paper, a nonuniform slant correction technique is presented where the slant correction problem is formulated as an optimal estimaiton problem of local slant angles at all horizontal positions. The optimal estimation is governed by a criterion function and several constraints designed to evaluate global and local goodness of the estimtated local angles. The optimal local slant angles which maximize the criterion function satisfying the constraints are searched for efficiently by a dynamic programming (DP)-based algorithm. Experimental results show the advantageous characteristics of the present technique over the uniform slant correction technique.

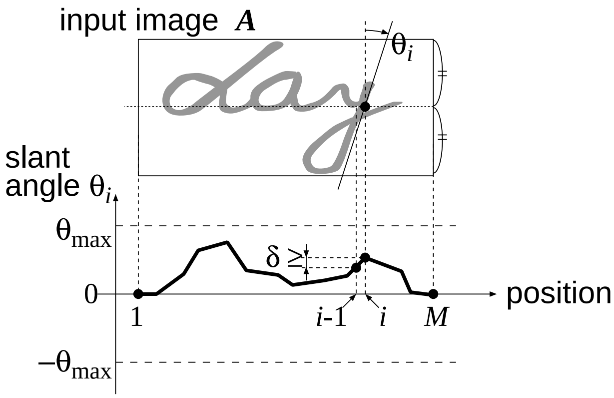

Figure 1. Nonuniform slant.

2. Nonuniform Slant Correction

2.1. Problem Formulation

Let $\boldsymbol{A} = \{a(i,j)\}$ denote a binary-valued word image of size $M \times N$. The nonuniform slant correction technique proposed in this paper is based on optimal estimation of the sequence $\theta_1, \ldots, \theta_i, \ldots, \theta_M$ where θ_i is the local slant angle at position i (Fig.1).

The policy of optimal estimation is the detection of long vertical strokes and the continuous propagation of their slant angles to their neighborhood. This policy is based on the fact that the long vertical strokes show the slant angle more evidently than other strokes.

According to the policy, the optimal estimation problem of $\theta_1, \ldots, \theta_i, \ldots, \theta_M$ is formulated as the maximization problem of the criterion function

$$F(\theta_1, \ldots, \theta_i, \ldots, \theta_M) = \sum_{i=1}^{M} f_i(\theta_i \mid \theta_{i-1}) \qquad (1)$$

with respect to variables $\theta_1, \ldots, \theta_i, \ldots, \theta_M$, subject to a continuity constraint and a range limitation defined as

$$|\theta_i - \theta_{i-1}| \leq \delta, \qquad (2)$$

$$|\theta_i| \in \Theta, \qquad (3)$$

where δ is a positive constant and $\Theta = [-\theta_{\max}, \theta_{\max}]$. The function $f_i(\theta_i \mid \theta_{i-1})$ is defined as the sum of two terms, i.e.,

$$f_i(\theta_i \mid \theta_{i-1}) = s_i(\theta_i) + \rho(\theta_i \mid \theta_{i-1}). \qquad (4)$$

where $s_i(\theta_i)$ evaluates a confidence level that a long vertical stroke with slant θ_i exists at position i and $\rho(\theta_i \mid \theta_{i-1})$ evaluates the smoothness, or continuity between θ_i and θ_{i-1}. By maximizing the criterion (1), slanted long vertical strokes are detected by the effect of $s_i(\theta_i)$ and their slant angles are smoothly propagated to their neighborhood by the effect of $\rho(\theta_i \mid \theta_{i-1})$. Note that $\rho(\theta_i \mid \theta_{i-1}) = 0$ when $i = 1$.

1 **for all** $\theta_1 \in \Theta$ **do** $g_1(\theta_1) := s_1(\theta_1)$;
2 **for** $i = 2$ **to** M **do**
3 **for all** $\theta_i \in \Theta$ **do begin**
4 $g_i(\theta_i) := \max_{\theta_{i-1}} [g_{i-1}(\theta_{i-1}) + f_i(\theta_i \mid \theta_{i-1})]$;
5 $b_i(\theta_i) := \theta_{i-1}$ which gives the maximum at Step 4;
6 **end;**
7 $\overline{\theta}_M := \operatorname{argmax}_{\theta_M \in \Theta} g_M(\theta_M)$;
8 **for** $i = M$ **downto** 2 **do** $\overline{\theta}_{i-1} := b_i(\overline{\theta}_i)$;

Figure 2. DP algorithm for the optimal local slant angles.

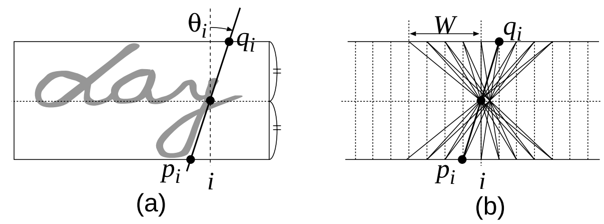

Figure 3. (a) Relation between local slant angle θ_i and correction line specified by p_i and q_i. (b)Possible correction lines at position i.

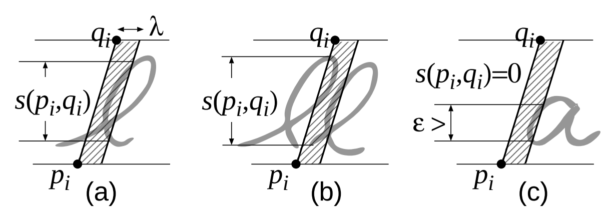

Figure 4. Calculation of the function $s(p_i, q_i)$.

2.2. Solution by DP

Fig.2 shows a DP algorithm for the optimization problem formalized in Section 2.1. Step 4 is so-called DP recursion which is calculated for all possible slant angles θ_i at each position i. The maximum selection in the DP recursion is performed with respect to θ_{i-1} which satisfy the constraints (2) and (3). The optimally estimated θ_i, denoted as $\overline{\theta}_i$, is obtained by the backtracking procedure, Step 7 and 8.

The DP algorithm itself shows its validity in an inductive manner. Assume that the value $g_{i-1}(\theta_{i-1})$ is the maximum (i.e., optimal) cumulated value of $f_k(\theta_k \mid \theta_{k-1})$ up to $i-1$. Then, it is clear from the DP recursion of Fig.2 (Step 4) that $g_i(\theta_i)$ is the maximum cumulated value up to i. Based on this idea, one can inductively prove the validity of the DP algorithm.

2.3. Implementation

2.3.1. Discretization of Slant Angle. The DP algorithm of Fig.2 can not be directly implemented on digital computers because the slant angle is real-valued, i.e., $\theta_i \in \Re$. Thus, instead of θ_i, we use the line segment between pixels $(p_i, 1)$ and (q_i, N) where p_i and q_i are integers and

$$\lfloor (p_i + q_i)/2 \rfloor = i \tag{5}$$

to represent the local slant angle at position i (Fig.3). The notation $\lfloor x \rfloor$ means the maximum integer which does not exceed x. Hereafter, the line segment is called a *correction line*.

According to this discretization, the description of the criterion function is modified. The function $f_i(\theta_i \mid \theta_{i-1})$ becomes $f(p_i, q_i \mid p_{i-1}, q_{i-1})$ which is defined as the sum of $s(p_i, q_i)$, and $\rho(p_i, q_i \mid p_{i-1}, q_{i-1})$. Since variables p_i and q_i carry positional information, subscript i is omitted from f and s in their discretized forms. In the next section, the detailed descriptions of the latter two functions are given.

The discretization also affects the constraints (2) and (3). The constraint (2) is to be modified as:

$$p_i \geq p_{i-1}, \quad q_i \geq q_{i-1}. \tag{6}$$

This constraint is not a continuity but a monotonicity constraint by itself. However, the constraint imposes continuity in a discrete form between θ_i and θ_{i-1} together with (5) (actually, $p_i - p_{i-1} \leq 3$). The constraint (3) becomes as follows (Fig.3(b)):

$$|p_i - i| \leq W, \quad |q_i - i| \leq W. \tag{7}$$

2.3.2. Criterion Function. The function $s(p_i, q_i)$ is designed to measure the confidence level that a slanted long vertical stroke exists on the correction line. Here we use the maximum height of connected black pixels on the correction line between $(p_i, 1)$ and (q_i, N) as the confidence level. Along with this basic definition, the following three points are considered on calculating $s(p_i, q_i)$.

1. For detecting curved strokes, it is assumed that the correction line has a width λ (Fig.4(a)).
2. For computational simplicity, the connectivity of black pixels is judged only for the vertical direction (Fig.4(b)).
3. For suppressing the effects of short vertical strokes and horizontal strokes, $s(p_i, q_i)$ is set at 0 if the maximum height is smaller than a positive constant ϵ (Fig.4(c)).

The function $\rho(p_i, q_i \mid p_{i-1}, q_{i-1})$ is designed to evaluate the smoothness between θ_i and θ_{i-1} and defined as

$$\rho(p_i, q_i \mid p_{i-1}, q_{i-1}) = -\alpha\Gamma_1 - \beta\Gamma_2 \tag{8}$$

where both Γ_1 and Γ_2 are nonnegative functions and α and β are nonnegative weights. It is clear that the maximum value of $\rho(p_i, q_i \mid p_{i-1}, q_{i-1})$ is 0. The value of Γ_1 equals to zero if $p_i = p_{i-1} + 1$ and $q_i = q_{i-1} + 1$, that is, two consecutive correction lines have same slant angle. Otherwise, the value of Γ_1 equals to the number of black pixels on the correction line between $(p_i, 1)$ and (q_i, N). Thus, by maximizing $-\alpha\Gamma_1$ (i.e., minimizing $\alpha\Gamma_1$), the local slant angles are to be stabilized. However, at ligatures and blanks, it is still easy to change the angles. This is because such parts contain few black pixels and therefore the effect of Γ_1 becomes weak. The value of Γ_2 equals to zero if two consecutive correction lines do not touch at their ends. Otherwise, the value of Γ_2 equals to the number of black pixels on the first (if $p_i = p_{i-1}$) or last (if $q_i = q_{i-1}$) quarter part of the ith correction line. The function Γ_2 also has a stabilizing effect as well as Γ_1. In addition, Γ_2 has the effect to prevent the concentration of the correction lines around black pixels, which causes unnatural slant-corrected images.

2.3.3. Algorithm Implementation and Computational Complexity. Now one can implement the DP algorithm of Fig.2 on digital computers by substituting θ_i by p_i and q_i. The loop of θ_i at Step 3 of Fig.2 is modified as the loop of possible combinations of p_i and q_i (which satisfy the constraints (5) and (7)). The DP recursion at Step 4 becomes as follows:

$$g_i(p_i, q_i) := \max_{p_{i-1}, q_{i-1}} [g_{i-1}(p_{i-1}, q_{i-1}) + f(p_i, q_i \mid p_{i-1}, q_{i-1})]. \tag{9}$$

The maximum selection in this DP recursion is performed with respect to the possible combinations of p_{i-1} and q_{i-1} which satisfy the constraints (5)–(7).

The resulting DP algorithm requires $O(MNW)$ computations, since the number of all possible combinations of p_i and q_i is $O(W)$ and the calculation of $g_i(\theta_i)$ (i.e., DP recursion (9)) requires $O(N)$ computations. This computational complexity is comparative to that of the conventional uniform slant correction techniques.

2.3.4. Slant-Corrected Image. Slant-corrected images, denoted as $\boldsymbol{B} = \{b(i, j)\}$, are obtained by *mapping* the pixels on the correction line between $(p_i, 1)$ and (q_i, N) of $\boldsymbol{A}$ onto ith row of $\boldsymbol{B}$, for all positions of i. This fact indicates that the nonuniform slant correction is closely related with elastic image matching, such as piecewise linear two-dimensional warping[10].

3. Experimental Results

In this section, the effectiveness of the present nonuniform slant correction technique is both qualitatively and quantitatively evaluated through experimental results. The

Figure 5. Examples of slant correction. For each word, the preprocessed image (top), the uniform slant correction result (middle), and the nonuniform slant correction result (bottom) are shown. The nonuniform slant correction results are similar (a), superior (b), and inferior (c), to the uniform slant correction results.

Figure 6. Artificially slanted words. (a) Original image. (b) Slanted by constant angle. (c) Slanted by sinusoidally changing angle.

parameters W, λ, ϵ, α, and β were fixed at 63, 4, 25, 1, and 2, respectively. A computer with Pentium III (500MHz) required 600 ms for the nonuniform slant correction of 256×64 word images.

In the following two experiments, the present technique was compared to the uniform slant correction technique developed by incorporating the additional constraint, $\theta_i = \theta_{i-1}$, into the present technique. While the results of this uniform slant correction technique somewhat differ from those of each conventional technique, the results will reflect the general characteristics of the conventional techniques.

3.1. Qualitative Evaluation

For qualitative evaluation of the present technique, several cursive word images included in the CEDAR CDROM-1 (city and state words) were subjected to a slant correction test. Each word image was normalized to be 64 pixels in height, binarized, and then padded with W white pixels at both left and right sides.

Fig.5 shows preprocessed word images and their slant-corrected images. The slant-corrected images of Fig.5(a) and (b) show that the present technique can correct the slants of component characters preserving their inherent shapes. Especially in (b), where obvious nonuniform slant appears in each word image, it is shown that the present technique provides near-perfect correction while the uniform slant correction (i.e., conventional) technique fails.

As shown in Fig.5(c), the present technique sometimes over-corrects the slants of several alphabets, such as 'X', so as to force their inherently slanted strokes to be perpendicular. This over-correction is still an open problem of the present technique. Most uniform slant correction techniques relax this problem by taking a average slant. This suggests that the use of wider context on the evaluation of the local slant angles will be one of promising remedies. A more radical remedy is indicated in Section 4 as a future work.

3.2. Quantitative Evaluation

For quantitative evaluation, an experiment to measure the mean square error between given and estimated slant angle was conducted by using artificially slanted machine-

Table 1. Mean square error (multiplied by 10^3) between given and estimated slant angle for the images in Fig.6

	constant slant					sinusoidally changing slant				
image no.	1	2	3	4	5	1	2	3	4	5
uniform correction	0	0	0	0	0	427	537	286	185	186
nonuniform correction	0.188	0.007	0.337	8.47	6.66	18.3	169	41.5	110	152

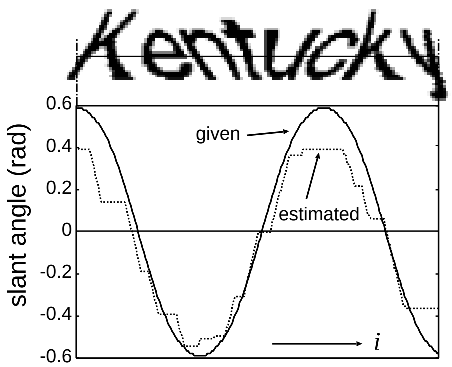

Figure 7. Given and estimated slant angles for a nonuniformly slanted image of "Kentucky".

printed word images. Fig.6(a), (b), and (c) show original machine-printed word images, slanted images with a constant angle, and slanted images with a sinusoidally changing angle, respectively.

Table 1 shows the mean square error between given and estimated slant angle, i.e., $\sum_{i=1}^{M}(\overline{\theta}_i - \psi_i)^2/M$ where ψ_i is the given slant angle at position i. This evaluation shows that (i) the present technique provides almost similar estimation to the uniform slant correction when a constant slant appears, and (ii) the present technique provides further accurate estimation than the uniform slant correction technique when nonuniform slant appears. Fig.7 shows the estimated slant angle obtained by the present technique as a function of position i with the given slant angle for the case of word "Kentucky". It is shown that the estimated angle follows the given slant angle appropriately.

4. Conclusion and Future Work

A nonuniform slant correction technique for handwritten word recognition was presented. The nonuniform slant correction problem was formalized as a constrained optimization problem where local slant angles were variables to be optimized and then solved efficiently by a dynamic programming (DP)-based algorithm. Experimental results indicates that the present technique possesses advantageous characteristics of over conventional uniform slant correction techniques. The results also show that the over-correction is the remaining problem.

The present technique can be embedded into word recognition systems based on segmentation-by-recognition, such as Kim and Govindaraju[2]. In [2], a slant correction technique is performed as preprocessing and then segmentation-by-recognition is performed in an optimization framework (based on 2-level DP). The embedment of the present technique into this framework will realize a slant-correction-and-segmentation-by-recognition where segmentation boundaries are represented by the correction lines and optimized with the help of a character recognizer. This scheme will relaxe the over-correction problem since character shape information can be used in the estimation of local slant angles.

Acknowledgement: The authors thank the Center of Excellence for Document Analysis and Recognition (CEDAR), the State University of New York, for providing us CEDAR CD-ROM 1. The authors also thank Toshiyuki Ishida (Fujitsu Chugoku Systems Ltd.) for his valuable comments.

References

[1] R.M. Bozinovic and S.N. Srihari, "Off-Line Cursive Script Word Recognition," IEEE Trans. PAMI, Vol. 11, No. 1, pp. 68–83, Jan. 1989.

[2] G. Kim and V. Govindaraju, "A lexicon driven approach to handwritten word recognition for real-time applications," IEEE Trans. PAMI, Vol. 19, No. 4, pp. 366-379, April 1997.

[3] D. Guillevic and C.Y. Suen, "Cursive Script Recognition: A Sentence Level Recognition Scheme," Proc. 4th IWFHR, pp. 216–223, Dec. 1994.

[4] E. Kavallieratou, N. Fakotakis and G. Kokkinakis, "A slant removal algorithm," Patt. Recog., Vol. 33, No. 7, pp. 1261–1262, Jul. 2000.

[5] G. Nicchiotti and C. Scagliola, "Generalised projections: a tool for cursive handwriting normalisation," Proc. 5th ICDAR, pp. 729-732, Sep. 1997.

[6] F. Kimura, M. Shridhar, and Z. Chen, "Improvements of a lexicon directed algorithm for recognition of unconstrained handwritten words," Proc. 2nd ICDAR., pp. 18–22, Oct. 1993.

[7] L. Simoncini and Zs.M. Kovács-V, "A System for Reading USA Census'90 Hand-Written Fields," Proc. 3th ICDAR, Vol. II, pp. 86–91, Aug. 1995.

[8] Y. Ding, F. Kimura, Y. Miyake and M. Shridhar, "Accuracy Improvement of Slant Estimation for Handwritten Words," Proc. 15th ICPR., Vol. 4, pp. 527–530, Sep. 2000.

[9] A. D. S.Britto JR., R. Sabourin, E. Lethelier, F. Bortolozzi, and C. Y. Suen , "Improvement in handwritten numeral string recognition by slant normalization and contextual information," Proc. 7th IWFHR., pp. 323–332, Sep. 2000.

[10] S. Uchida and H. Sakoe, "Piecewise linear two-dimensional warping," Proc. 15th ICPR., Vol. 3, pp. 538–541, Sep. 2000.