[5] W. Hoff and N. Ahuja, "Surfaces from stereo: Integrating feature matching, disparity information, and contour detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989.

[6] R.D. Eastman and A.M. Waxman, "Using disparity functionals for stereo correspondence and surface reconstruction," *Comp. Vision, Graphics, and Image Proc.*, vol. 39, July 1987, pp. 73–101.

[7] J.E.W. Mayhew and J.P. Frisby, "Psychophysical and computational studies towards a theory of human stereopsis," *Artificial Intelligence*, vol. 17, 1981, pp. 349–385.

[8] S.B. Pollard, J.E.W. Mayhew, and J.P. Frisby, "Pmf: A stereo correspondence algorithm using a disparity gradient limit," *Perception*, vol. 14, pp. 449–470, 1981.

[9] G.G. Medioni and R. Nevatia, "Segment-based stereo matching," *Comp. Graphics Img. Proc.*, vol. 13, pp. 257–269, 1980.

[10] H.S. Lim and T.O. Binford, "Stereo correspondence: A hierarchical approach," *Proc. DARPA Img. Understg. Wkshp.*, Los Angeles, pp. 234–241, Feb. 1987.

[11] R.C.K. Chung and R. Nevatia, "Use of monocular groupings and occlusion analysis in an hierarchical stereo system," *Proc. IEEE Conf. Comp. Vision Patt. Recog.*, Maui, Hawaii, pp. 50–55, June 1991.

[12] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, pp. 139–154, Mar. 1985.

[13] H.H. Baker and T.O. Binford, "Depth from edge and intensity based stereo," *Proc. Seventh Int'l Joint Conf. AI*, pp. 631–636, Aug. 1981.

[14] S.T. Barnard and W.B. Thompson, "Disparity analysis of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 2, pp. 333–340, July 1980.

[15] Y.C. Kim and J.K. Aggarwal, "Positioning 3-d objects using stereo images," *IEEE J. Robotics Autom.*, vol. 3, pp. 361–373, Aug. 1987.

[16] Y.C. Kim and J.K. Aggarwal, "Determining object motion in a sequence of stereo images," *IEEE J. Robotics Autom.*, vol. 3, pp. 599–614, Dec. 1987.

[17] N. Ayache and B. Faverjon, "Efficient registration of stereo images by matching graph descriptions of edge segments," *Int'l J. Comp. Vision*, pp. 107–131, 1987.

[18] U.R. Dhond, "Stereo image interpretation in the presence of narrow occluding objects." PhD thesis, University of Texas at Austin, Department of Electrical and Computer Engineering, Austin, Tex., USA, Aug. 1992.

[19] J.J. Little and W.E. Gillett, "Direct evidence for occlusion in stereo and motion," *Img. Vision Computg.*, vol. 8, pp. 328–340, Nov. 1990.

Direct Extraction of Topographic Features for Gray Scale Character Recognition

Seong-Whan Lee and Young Joon Kim

***Abstract*—Optical character recognition(OCR) traditionally applies to binary-valued imagery although text is always scanned and stored in gray scale. However, binarization of multivalued image may remove important topological information from characters and introduce noise to character background. In order to avoid this problem, it is indispensable to develop a method which can minimize the information loss due to binarization by extracting features directly from gray scale character images.**

In this paper, we propose a new method for the direct extraction of topographic features from gray scale character images. By comparing the proposed method with Wang and Pavlidis' method, we realized that the proposed method enhanced the performance of topographic feature extraction by computing the directions of principal curvature efficiently and prevented the extraction of unnecessary features. We also show that the proposed method is very effective for gray scale skeletonization compared to Levi and Montanari's method.

***Index Terms*—Gray scale character recognition, principal curvature, principal orthogonal elements, topographic feature extraction.**

I. INTRODUCTION

It is well known that the character images of documents seen by the scanners have a wide range of gray scale for the following reasons [1]:

- Gray or different color background of documents as in magazine and business forms
- Textured background of documents as in bank checks
- Different types of ink
- Convolution distortion because of the point spread function of the scanner
- Nonuniform illumination of the scanner
- Multiplicative noise such as nonuniform paper reflection
- Additive noise due to electronics

When such a gray scale character image is binarized there is a significant information loss: Typically areas which are narrow compared to the support of the point spread function disappear, resulting in broken characters or touching characters.

Two kinds of methods have been used to overcome the disadvantages of simple binarization such as thresholding [1]. The first is to do binarization in a far more careful way than it is currently done in OCR. A representative of this kind of method is adaptive thresholding [2]. However, it requires a relatively large amount of computation, and no matter how good the binarization is, it cannot completely avoid the loss of information. The second approach is to perform recognition without binarization by using techniques related to matched filters [3], [4]. However, this approach is limited to applications such as recognition of single font characters where there is only a limited variation on the form of the characters.

A more promising approach is to perform feature extraction directly from gray scale character images [1]. By doing so, it bypasses the phase of binarization and thus avoids the loss of information caused by the operation.

In this paper, we propose a new method for extracting topographic

Manuscript received Mar. 8, 1994; revised Dec. 26, 1994.

S.-W. Lee is with the Dept. of Computer Science, Korea University, Seoul 136-701, Korea; e-mail: swlee@ailab.korea.ac.kr.

Y.J. Kim is with BIT Computer Co., Ltd., 818-11, Yeoksam-dong, Kangnam-ku, Seoul 135-080, Korea.

IEEECS Log Number P95072.

0162-8828/95$04.00 © 1995 IEEE

Authorized licensed use limited to: Korea University. Downloaded on October 25, 2008 at 01:24 from IEEE Xplore. Restrictions apply.

features directly from gray scale character images without calculating the eigenvalues and eigenvectors of the underlying image intensity surface. By comparing the proposed method with the Wang and Pavlidis' method it is realized that the proposed method enhanced the performance of topographic feature extraction by computing the directions of principal curvature efficiently and prevented the unnecessary feature extraction. We also show that the proposed method is very effective for gray scale skeletonization compared to Levi and Montanari's method [5].

II. Recent Works for Direct Extraction of Features from Gray Scale Character Images

There are two excellent recent works related to direct extraction of features from gray scale character images. The first is to use boundary features [6]. In this method, a character classifier is proposed to recognize gray scale character images by extracting structural features from character outlines. A fast local contrast based gray scale edge detector has been developed to locate character boundaries. A pixel is considered as an edge-pixel if its gray value is below a threshold and has a neighbor whose gray value is above the threshold. Edges are then thinned to one pixel width. Extracting structural features from edges is performed by convolving the edges with a set of feature templates. Extracted features are compressed to form a binary vector with 576 features and it is used as an input to a classifier.

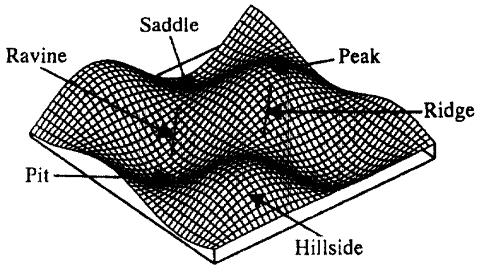

Fig. 1. An example of topographic features.

The second is to use topographic features [1], [7], [8], [9] defined by Haralick et al. [10]. Wang and Pavlidis [1] developed a technique for grouping and assembling topographically labeled pixels to form primitive features involved in solving practical pattern recognition problems. This method consists of four steps. In the first step, each pixel is classified into one of peak, pit, ridge, ravine, saddle, and hillside (Fig. 1) according to rules based upon the estimated first and second directional derivatives of the underlying image intensity surface. In the second step, basic structural information is extracted from the resulting labeled image. This is accomplished by an algorithm which creates a graph, which had been called a topographic feature graph (TFG) in a single pass through the labeled image. Each node of a TFG is connected with characteristic region formed entirely by peak and ridge points, or entirely by saddle points, or entirely by flat points. Pit, ravine, and hillside points serve as separators during the process in such a way that a connected component of a TFG corresponds to a single character, or part of a broken character, or a set of touching characters. In the third step, the TFG obtained in the second step is converted into a geometric feature graph (GFG). A GFG is a graph similar to the original TFG except that its nodes are simply points or line segments instead of connected regions of arbitrary shapes in the case of a TFG. This conversion is performed by an algorithm, which replaces each saddle or flat region by a single point and converts each peak and ridge region into one or more line segments. Some grouping and assembling of the characteristic regions of the TFG are also performed by this algorithm but the basic structure of the edge set of the TFG is preserved in the resulting GFG. The last step of the method is the recognition step. The geometric features extracted in the third step are directly fed into a hybrid structural/statistical classifier for immediate recognition.

In this method, the topographic feature extraction step requires a relatively large amount of computation, but structural information extraction step and geometric feature extraction step are simple processes which connect the results obtained from topographic feature extraction step and replace the node of TFG. Therefore, the performance of total feature extraction steps depends on the topographic feature extraction step. The topographic feature extraction step also consists of three steps as follows: the computation of the first and second derivatives, the computation of the eigenvalues and eigenvectors, and the classification of elements.

At each pixel P, if the gradient magnitude is zero, the definitions of Haralick et al. [10] are used directly to assign P a label. Otherwise, the eigenvectors are approximated by two perpendicular directions from the four natural directions defined on the image lattice, namely, the horizontal direction (H), the vertical direction (V), the right-diagonal direction (R), and the left-diagonal direction (L). The reason why the eigenvectors are approximated by two perpendicular directions is that the topographic label, namely, peak, ridge, pit, ravine, or saddle, will not be assigned precisely at the center of a pixel. For example, in Fig. 2, if we use the method of Haralick et al. [10] directly, pixel b is assigned as a hillside because the gradient magnitude is not zero. Consequently, a mis-assignment will occur.

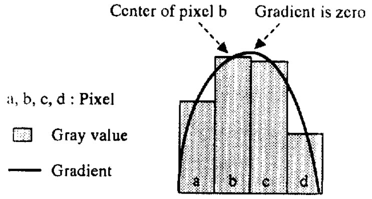

Fig. 2: An example of mis-assignment of topographic features

For ideal input images which have a uniform variation for gray value and symmetry for a brightest pixel, the gradient magnitude would be zero at a brightest pixel, but for real world images, the gradient magnitude would be rarely zero. Therefore, if the Wang and Pavlidis' method is used for topographic feature extraction, the eigenvectors must be approximated mostly by two perpendicular directions, and the calculation to get approximated directions requires much reluctant efforts. In addition, by employing such an approximation for the eigenvectors, unnecessary ridges and peaks may be extracted at the places such as bend points, starting points, end points, or corresponding hillside.

Therefore, in order to avoid these problems, it is indispensable to develop a method which can extract topographic features directly from gray scale character images without useless calculation and extraction of unnecessary topographic features.

Authorized licensed use limited to: Korea University. Downloaded on October 25, 2008 at 01:24 from IEEE Xplore. Restrictions apply.

III. A Proposed Method for Direct Extraction of Topographic Features from Gray Scale Character Images

The proposed method extracts topographic features directly from gray scale character images without calculating the eigenvalues and eigenvectors of the underlying image intensity surface and prevents the unnecessary topographic feature extraction at the places such as bend points, starting points, end points, or corresponding hillside. The outline of the proposed method is as follows:

Step 1) Calculation of gradient for adjacent pixels.

Step 2) Calculation of the derivative of gradient.

Step 3) Determination of principal curvature directions.

Step 4) Determination of principal orthogonal elements.

Step 5) Assignment of feature

Details of each step are given in the following subsections.

A. Calculation of Gradient for Adjacent Pixels

Considering $m \times n$ input image, the neighbors of a pixel P are defined by 3×3 window as shown in Fig. 3.

P_2	P_1	P_0
P_3	P	P_7
P_4	P_5	P_6

Fig. 3. 8-adjacent pixels for a pixel P

Then, the gradient m_i for adjacent pixel P_i is as follows:

$$m_i = \frac{\textit{difference of gray value}}{\textit{distance increment}} = \frac{\Delta I_i}{\Delta S} \tag{1}$$

where

$$\Delta I_i = \begin{cases} I(\mathrm{P}) - I(\mathrm{P}_i) & \text{if } i = 0, 1, 2, 3 \\ I(\mathrm{P}_i) - I(\mathrm{P}) & \text{if } i = 4, 5, 6, 7 \end{cases} \tag{2}$$

$$\Delta S = \begin{cases} 1 & \text{if } \mathrm{P}_i \text{ is adjacent pixel in horizontal or vertical directions} \\ \sqrt{2} & \text{if } \mathrm{P}_i \text{ is adjacent pixel in right-diagonal or left-diagonal directions} \end{cases} \tag{3}$$

and $I(\mathrm{P})$ represents the gray value of pixel P.

The reason why the order of calculation for i in ΔI_i is different is that the four directions shown in Fig. 4 were considered in gradient calculation.

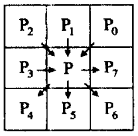

Fig. 4 Four directions in gradient calculation.

B. Calculation of the Derivative of Gradient

The derivative of gradient m'_k, $k = H, V, R, L$, in pixel P for unit distance increment is defined as follows:

$$m'_H = m_7 - m_3 \tag{4}$$

$$m'_V = m_5 - m_1 \tag{5}$$

$$m'_R = m_6 - m_2 \tag{6}$$

$$m'_L = m_4 - m_0 \tag{7}$$

The gradient and derivative of the gradient described until now can be calculated at the same time by convolving the 3×3 original image array with the Laplacian-like masks as follows:

$$m'_k = I \otimes L_k \quad k = H, V, R, L \tag{8}$$

where, $\otimes$ is the convolution operator, I is 3×3 original image array, and L_k are Laplacian-like masks defined in Fig. 5.

(a)

0	0	0
-1	2	-1
0	0	0

(b)

0	-1	0
0	2	0
0	-1	0

(c)

$-\frac{1}{\sqrt{2}}$	0	0
0	$\sqrt{2}$	0
0	0	$-\frac{1}{\sqrt{2}}$

(d)

0	0	$-\frac{1}{\sqrt{2}}$
0	$\sqrt{2}$	0
$-\frac{1}{\sqrt{2}}$	0	0

Fig. 5. Laplacian-like masks for four directions: (a) horizontal, (b) vertical, (c) right-diagonal, (d) left-diagonal.

C. Determination of Principal Curvature Directions

The principal curvature directions are limited to horizontal&vertical directions and right-diagonal& left-diagonal directions. Then, if $\|m'_H - m'_V\| \geq \|m'_R - m'_L\|$, the principal curvature directions are horizontal&vertical directions. Otherwise, the principal curvature directions are right-diagonal&left-diagonal directions.

D. Determination of Principal Orthogonal Elements

The principal orthogonal elements are defined as the four adjacent pixels in assigning features. There are two kinds of principal orthogonal elements, namely, the principal orthogonal elements of horizontal&vertical directions and right-diagonal&left-diagonal directions (Fig. 6).

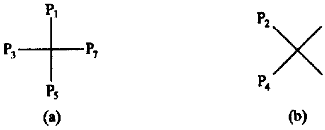

Fig. 6. The principal orthogonal elements: (a) horizontal&vertical directions, (b) right-diagonal&left-diagonal directions.

If the gradient for two adjacent pixels is monotonic increasing or decreasing in the direction whose derivative of gradient is minimum, the principal orthogonal elements are the adjacent pixels in the directions rotated from principal curvature directions by 45°. Otherwise, the principal orthogonal elements are the adjacent pixels in principal curvature directions. The reason why such conditions are applied is that the topographic features may be misassigned as ridge or peak at the places such as bend points, starting points, end points, or corresponding hillside.

E. Assignment of Feature

Let d_1 and d_2 be directions of the principal orthogonal elements. For each d_i, $i = 1, 2$, we consider the two neighbors of the pixel P, P_{i1}, and P_{i2}, in that direction. The zero crossing conditions are applied to P_{i1}, P, and P_{i2} in d_i. The rules for assigning topographic fea-

Authorized licensed use limited to: Korea University. Downloaded on October 25, 2008 at 01:24 from IEEE Xplore. Restrictions apply.

tures are as follows:

Case 1) P is not a zero crossing in either d_1 or d_2.

In this case, P is a hillside (Fig. 7a).

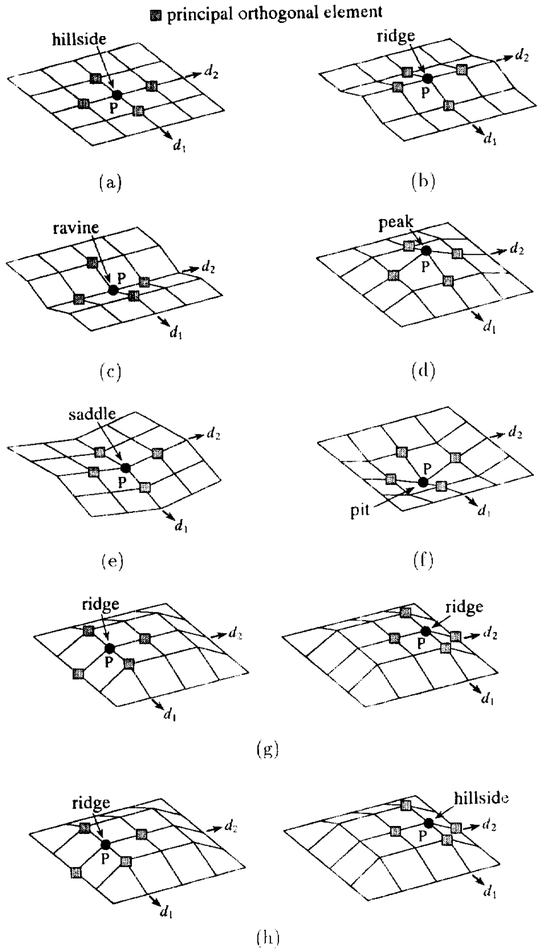

Fig. 7. Assignment of topographic features: (a) the case in which P is a hillside, (b) the case in which P is a ridge, (c) the case in which P is a ravine, (d) the case in which P is a peak, (e) the case in which P is a saddle, (f) the case in which P is a pit, (g) the case in which P is a ridge when we only want to extract the topographic features, (h) the case in which P is a ridge or a hillside when we only want to extract the skeleton.

Case 2) P is a zero crossing in exactly one of d_1 and d_2.

2.1) If zero crossing is the case that the sign of gradient is changed from positive to negative, P is a ridge (Fig. 7b).

2.2) If zero crossing is the case that the sign of gradient is changed from negative to positive, P is a ravine (Fig. 7c).

Case 3) P is a zero crossing in both of d_1 and d_2.

3.1) If the sign of gradient in d_1 is changed from positive to negative and the sign of gradient in d_2 is changed from positive to negative, P is a peak (Fig. 7d).

3.2) If the sign of gradient in d_1 is changed from positive to negative and the sign of gradient in d_2 is changed from negative to positive, P is a saddle (Fig. 7e).

3.3) If the sign of gradient in d_1 is changed from negative to positive and the sign of gradient in d_2 is changed from negative to positive, P is a pit (Fig. 7f).

Case 4) The gradient is zero in d_1, and the sign of gradient is changed from positive to zero or from zero to negative in d_2.

4.1) If we only want to extract the topographic features, P is a ridge (Fig. 7g).

4.2) If we only want to extract the skeleton from a gray scale character image, P is a ridge or a hillside (Fig. 7h); if the sign of gradient is changed from positive to zero in d_2, P is a ridge, and if the sign of gradient is changed from zero to negative in d_2, P is a hillside. This case occurs when the pixels having the same gray value are adjacent.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. Input Data

The proposed method has been implemented on a Sun SPARC-2 workstation using C language under X-window system and has been applied to 2,350 gray scale images in handwritten Hangul character database PE92 [11]. PE92 is the first database collected large-set handwritten Hangul characters in Korea. It contains 100 sets of 2,350 handwritten Hangul characters which are considered to be a complete set of characters for general and daily use. Each character image is 100 × 100 gray scale image with 256 gray levels. We show some of those images. The data which were used by Levi and Montanari [5] also have been used for gray scale skeletonization. Fig. 8 and Fig. 9 show the input data for topographic feature extraction and gray scale skeletonization, respectively.

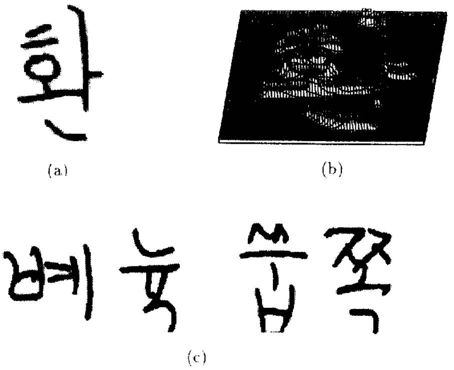

Fig. 8. Input data for topographic feature extraction: (a) gray scale image of a character, (b) 3D topographic shape of the character, (c) gray scale image of a series of characters.

B. Results of Topographic Feature Extraction

In order to compare the results of topographic feature extraction, three kinds of methods have been considered: Wang and Pavlidis' method [1], the proposed method 1 in which the conditions for determining the principal orthogonal elements were not considered, and the proposed method 2 in which these conditions were considered. In the proposed method 1, the principal orthogonal elements are the adjacent pixels in principal curvature directions.

Authorized licensed use limited to: Korea University. Downloaded on October 25, 2008 at 01:24 from IEEE Xplore. Restrictions apply.

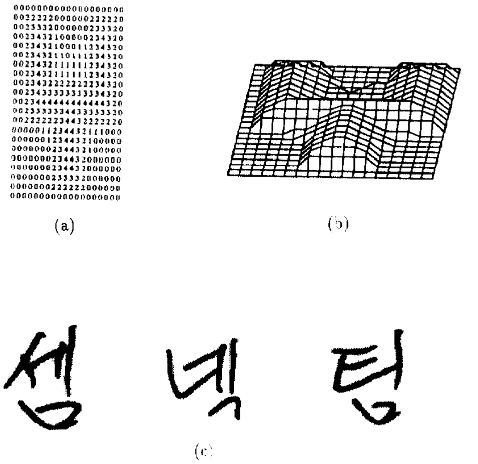

Fig. 9. Input data for gray scale skeletonization: (a) data used by Levi and Montanari [5], (b) 3D topographic shape of the data in (a), (c) gray scale image of a series of characters.

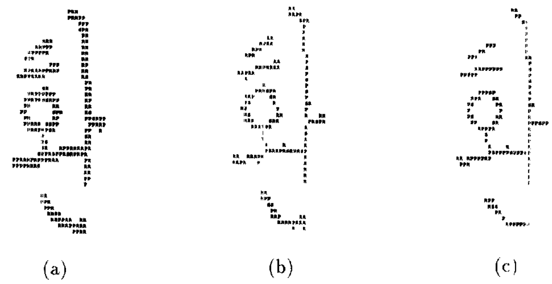

Fig. 10. The results of topographic feature extraction for the data in Fig. 8a: (a) Wang and Pavlidis' method [1], (b) proposed method 1, (c) proposed method 2.

Figs. 10 and 11 show the results of three methods for topographic feature extraction. In the results, ridge, peak, and saddle point pixels were marked by R, P, and S, respectively.

As can be seen in Figs. 10 and 11, the proposed method 2 shows superior performance in topographic feature extraction.

C. Results of Skeleton Extraction

In order to compare the results of gray scale skeletonization [12], we considered the Levi and Montanari's method [5], the Wang and Pavlidis' method [1], and the proposed method 2. Fig. 12 shows the results of three methods for gray scale skeletonization with Levi and Montanari's data [5]. In Levi and Montanari's method, the pixels which correspond to gray-weighted skeleton are marked by M, and in other methods, R and P represent the skeletal pixels. In order to extract skeleton in proposed method 2, when the sign of gradient was changed from positive to zero, we assigned a pixel to a ridge, and when the sign of gradient was changed from zero to negative, we assigned a pixel to a hillside.

The results shown in Fig. 12 reveal that the proposed method 2 is superior to other methods in the point of view of gray scale skeletonization for character recognition [13].

Fig. 13 shows the results of gray scale skeletonization for Hangul characters with the proposed method 2. In this figure, the results were thinned to one pixel width after topographic feature extraction.

D. Processing Time and Computational Complexity

In order to compare the average processing time in topographic feature extraction, we used 2,350 gray scale images in handwritten Hangul character database PE92 [11]. Each character image is

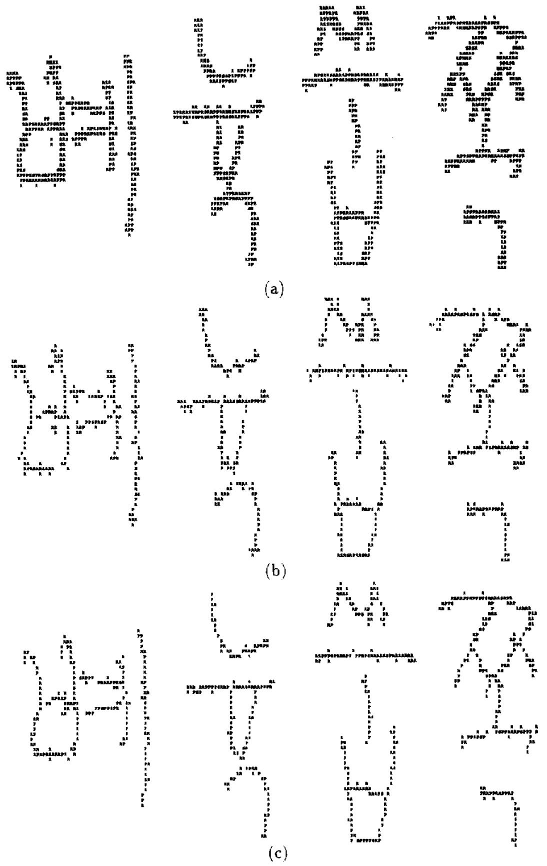

Fig. 11. The results of topographic feature extraction for the data in Fig. 8c: (a) Wang and Pavlidis' method [1], (b) proposed method 1, (c) proposed method 2.

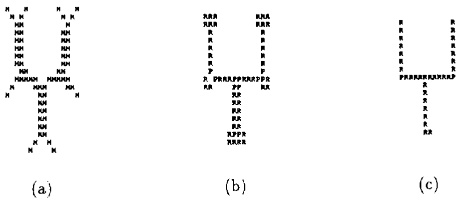

Fig. 12. Results of gray scale skeletonization for the data in Fig. 9a: (a) Levi and Montanari's method, (b) Wang and Pavlidis' method, (c) proposed method 2.

Fig. 13. Results of gray scale skeletonization for the data in Fig. 9c.

Authorized licensed use limited to: Korea University. Downloaded on October 25, 2008 at 01:24 from IEEE Xplore. Restrictions apply.

100×100 gray scale image with 256 gray levels. Table I shows the average processing time and computational complexity for Wang and Pavlidis' method [1] and the proposed method 2. As shown in Table I, in case of extracting topographic features using the method 2, the processing time was approximately five times as fast as the Wang and Pavlidis' method.

TABLE I
AVERAGE PROCESSING TIME AND COMPUTATIONAL COMPLEXITY FOR EACH METHOD

Method	Average processing time	Computational complexity
Wang and Pavlidis' method	3.82 sec	$i \times j \times (8 \times M^2 + C)$
Proposed method 2	0.76 sec	$i \times j \times (4 \times N^2 + C)$

i : the width of input image
j : the height of input image
M : neighborhood size for calculating the first and second partial derivatives
N : neighborhood size for convolution with Laplacian-like masks
C : the number of comparative operations

V. CONCLUDING REMARKS

In this paper, we proposed a new method for extracting topographic features directly from a gray scale character image without calculating eigenvalues and eigenvectors of the underlying image intensity surface.

For real world character images, the gradient magnitude would be rarely zero at the center of pixel. Therefore, if the Wang and Pavlidis' method is used for topographic feature extraction, the eigenvectors must be approximated mostly by two perpendicular directions, and the calculation to get approximated directions requires much reluctant efforts. In addition to that, by employing such an approximation for the eigenvectors, unnecessary ridges and peaks may be extracted at the places such as bend points, starting points, end points, or corresponding hillside.

In case of extracting topographic features using the proposed method, the processing time was approximately five times as fast as the Wang and Pavlidis' method because only simple comparative operations were used, and unnecessary topographic features need not be extracted at bend points, starting points, and end points by taking the local information of gray scale character image into account in determining principal orthogonal elements. In addition to that, experimental results with Levi and Montanari's data revealed that the proposed method was also very effective for gray scale skeletonization for character recognition.

ACKNOWLEDGMENTS

The authors wish to thank the anonymous reviewers for their helpful comments in improving the earlier draft of this paper. This research was supported by the 1992 Directed Basic Research Fund of Korea Science and Engineering Foundation.

REFERENCES

[1] L. Wang and T. Pavlidis, "Direct gray scale extraction of features for character recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1,053-1,067, Oct. 1993.

[2] T. Pavlidis and Wolberg, "An algorithm for the segmentation of bilevel images," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 570-575, Miami, June 1986.

[3] H. Murakami and B.V.K.V. Kumar, "Efficient calculation of primary images from a set of images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 4, pp. 511-515, 1982.

[4] C.M. Leung, "A practical basis set for Chinese character recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 532-537, San Francisco, June 1985.

[5] G. Levi and U. Montanari, "A gray-weighted skeleton," *Information and Control*, vol. 17, pp. 62-91, 1970.

[6] S.W. Lam, A.C. Girardin, and S.N. Srihari, "Gray scale character recognition using boundary features," *Proc. SPIE Conf. Machine Vision Applications in Character Recognition and Industrial Inspection*, vol. 1,661, pp. 98-105, San Jose, 1992.

[7] T. Pavlidis, "Recognition of printed text under realistic conditions," *Proc. Second IPTP Conf.*, pp. 65-76, Tokyo, Jan. 1992.

[8] T. Pavlidis, L. Wang, J. Zhou, W.J. Sakoda, and J. Rocha, "Recognition of poorly printed text by direct extraction of features from gray scale," *Proc. SPIE Conf. Machine Vision Applications in Character Recognition and Industrial Inspection*, vol. 1,661, pp. 1-9, San Jose, Feb. 1992.

[9] L. Wang and T. Pavlidis, "A geometric approach to machine-printed character recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 665-668, Champaign, Ill., June 1992.

[10] R.M. Haralick, L.T. Watson, and T.J. Laffey, "The topographic primal sketch," *Int'l J. Robotics Research*, vol. 2, pp. 50-72, 1983.

[11] D.H. Kim, Y.S. Hwang, S.T. Park, E.J. Kim, S.H. Paek, and S.Y. Bang, "Handwritten Korean character image database PE92," *Proc. Second Int'l Conf. Document Analysis and Recognition*, pp. 470-473, Tsukuba Science City, Japan, Oct. 1993.

[12] L. Lam, S.-W. Lee, and C.Y. Suen, "Thinning methodologies—a comprehensive survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 869-885, Sept. 1992.

[13] S.-W. Lee, L. Lam, and C.Y. Suen, "A systematic evaluation of skeletonization algorithms," *Pattern Recognition and Artificial Intelligence*, vol. 7, no. 5, pp. 1,203-1,225, 1993.

Efficient Color Histogram Indexing for Quadratic Form Distance Functions

James Hafner, Harpreet S. Sawhney, Will Equitz,
Myron Flickner, and Wayne Niblack

***Abstract*—In image retrieval based on color, the weighted distance between color histograms of two images, represented as a quadratic form, may be defined as a match measure. However, this distance measure is computationally expensive (naively $O(N^2)$ and at best $O(N)$ in the number N of histogram bins) and it operates on high dimensional features ($O(N)$). We propose the use of low-dimensional, simple to compute distance measures between the color distributions, and show that these are lower bounds on the histogram distance measure. Results on color histogram matching in large image databases show that prefiltering with the simpler distance measures leads to significantly less time complexity because the quadratic histogram distance is now computed on a smaller set of images. The low-dimensional distance measure can also be used for indexing into the database.**

***Index Terms*—Color histogram matching, image querying, image databases, efficient multidimensionalal feature matching, histogram indexing.**

I. INTRODUCTION

In the *query by image content* (or QBIC) project, we are developing a system for efficient indexing and retrieval of images from a large database based on their content defined in terms of shapes, colors, textures, and user sketches [16]. Other efforts towards similar goals are presented in [2], [7], [8], [11], [12], [19].

Manuscript received May 5, 1993; revised Mar. 6, 1995.

J. Hafner, H.S. Sawhney, M. Flickner, and W. Niblack are with the Machine Vision Group, IBM Almaden Research Center, K54, 650 Harry Road, San Jose, CA 95120; e-mail sawhney@almaden.ibm.com.

W. Equitz is with Storm Technology, 1861 Landings Dr., Mountain View, CA 94043. This work was performed while this author was at the IBM Almaden Research Center.

IEEECS Log Number P95069.

0162-8828/95$04.00 © 1995 IEEE

Authorized licensed use limited to: Korea University. Downloaded on October 25, 2008 at 01:24 from IEEE Xplore. Restrictions apply.