Skeletonization of Grayscale Character Images Using Pixel Superiority Index

Kyung-Won Kang, Jang Won Suh and Jin H. Kim
Department of Computer Science & Center for AI Research, KAIST
373-1 Gusong-dong, Yusong-ku, Taejon, 305-701, Korea
E-mail: {kwkang, suh, jkim}@ai.kaist.ac.kr

In this paper, we present pixel superiority index as a tool for designing a skeletonization algorithm which utilizes topographic features efficiently. We clarify a relationship between pixel superiority index and topographic features. Then, using the relationship, we transform a problem of skeletonization into a problem of skeleton growing. Preliminary experiments show that the proposed algorithm generates a comparatively good quality of skeleton even when stroke width is thick. Applying the proposed algorithm to a stroke-based structural Hangul recognizer, 17.5% of error reduction is obtained compared to a conventional thinning algorithm with dynamic thresholding.

1 Introduction

Stroke analysis approach is considered as the most feasible solution for the recognition of handwritten oriental scripts such as Chinese Hanzi, Korean Hangul and Japanese Kana. It is because not only that stroke is, especially in those scripts, the key structural feature for character identity but also that the stroke representation of character patterns permits more intuitive design of recognition algorithms exploiting resemblance to human conception of those patterns [5]. Stroke is often obtained connecting skeletal pixels. The objective of skeletonization, particularly for the handwritten characters, should be recovering the movement track of pen tip. Ignoring this objective, skeletal pixels are computed by thinning binary images obtained by simple threshold logic. As a consequence, such binarization approach reveals several intrinsic defects such that adjacent strokes are often merged together due to the ignorance of the shallow valley between two strokes(figure 1). Although it may be overcome by intelligently selecting the threshold value, or applying the threshold logic locally, it is, by no means, a more difficult problem.

To overcome such a problem, direct skeletonization from grayscale images is attempted to utilize topological features such as peak, ridge, valley, ravine, saddle, etc. Those topological features are extracted directly from grayscale images without the usual step of binary thresolding [1,2]. As the pioneer of such attempt, Haralick proposed a mathmatical description of topographic primal sketch of a grayscale digital image [1]. In this method, each pixel is classified into one of peak, pit, ridge, ravine, saddle, flat or hillside and skeletal pixels

(a) (b) (c)

Figure 1: Shape distortion by binarization. (a) Grayscale image of a Hangul character, (b) Binary image of (a), and (c) Thinned image of (b)

can be defined as a set of peak, ridge and saddle points. Topographic feature classification was more developed by Pavlidis [2]. He developed techniques for grouping and assembling topographically labeled pixels to form primitive features. Though Pavlidis' method is more efficient than Haralick's, it is also very time-consuming to determine the principal direction. To reduce time complexity, Lee used a derivative of gradient in determining the principal direction by assuming that only four directions are possible [3]. Similar to Lee's method, Suh proposed a digital version of approximation by 4-directional scanning, which is much more efficient than the previous methods because it does not calculate the principal direction [4].

However, these approaches have a common problem in extracting skeleton from topographic features when stroke width is thick. In that case, a simple collection of feature points doesn't satisfy a necessary condition of skeleton, or a property of one-pixel width. One of the solutions is to apply a conventional binary thinning algorithm on the set of skeletal pixels [7]. This straightforward approach brings shape distortion, an intrinsic problem of conventional thinning approaches. The other solution is to use a concept of a primary line and a gap filling process [4]. But this method has a difficulty in extracting primary lines, because primary lines are closely related to thickness of skeletal pixels. Another solution is to extract strokes directly from topographic features bypassing skeletonization. Pavlidis grouped topographic features and represented a character with a graph [2]. Although graph-representation of a character is requisite in stroke analysis, the graph may disregard some information like curvature extremity by grouping topographic features, which are important for stroke analysis.

This paper presents a new grayscale skeletonization algorithm by pixel su-

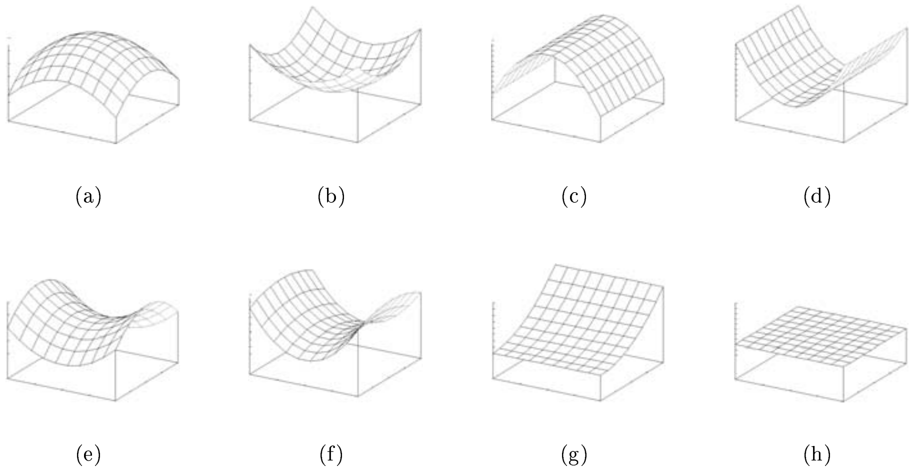


Figure 2: Characteristics of topographic features. (a) Peak, (b) Pit, (c) Ridge, (d) Ravine, (e) Ridge saddle, (f) Ravine saddle, (g) Hillside, and (h) Flat

periority index. Pixel superiority index indicates the topographic significance of a pixel compared to its neighbors and approximates topographic features digitally. The proposed algorithm is very efficient in the aspect of time complexity. Moreover, it guarantees a comparatively good quality of skeleton even when stroke width is thick.

The rest of this paper is organized as follows. In section 2, we summarize previous works related to topographic feature classification. In section 3, the proposed algorithm is explained. Experimental results are analyzed in section 4 and a conclusion is given in section 5.

2 Related Works

2.1 Topographic feature classification

Haralick proposed a direct feature extraction method from grayscale images called the topographic primal sketch [1]. Topographic features such as peak, pit, ridge, ravine, hillside and flat are mathematically described by first-order and second-order partial derivatives of an image surface. Characteristics of topographic features are shown in figure 2.

Feature classification of a pixel consists of three steps; surface fitting, calculation of a Hessian matrix and its eigen-vectors, and feature labeling. For

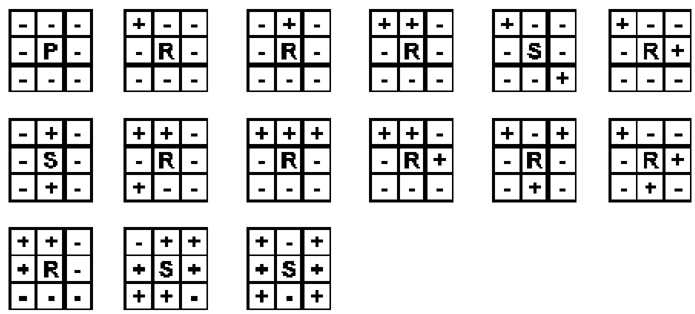


Figure 3: Local intensity configurations of skeletal pixels in no consideration of rotation and symmetry ('+' means that its gray-value is greater than gray-value of the center pixel and '-' means the opposite. 'P', 'R' and 'S' represent peak, ridge and saddle, respectively)

each pixel, a local image surface is fitted in the form of $z = f(x, y)$ by least square fitting or spline curves. Then, the Hessian matrix at the pixel is calculated using second-order partial derivatives of $f(x, y)$. Eigen-vectors of the Hessian matrix represent principal directions, which are the directions in which the normal curvature of the local surface attains an extreme value. Finally, the pixel is topographically labeled by classification rules. For details, refer to the original paper.

2.2 Topographic feature extraction by 4-directional scanning

One of the authors proposed an efficient topographic feature classification method that uses only local intensity configuration of neighborhood [4]. Without calculation of principal directions, each pixel is topographically classified according to geometric features along 4 principal directions. Here, only 4 principal directions, horizontal, vertical, right-diagonal and left-diagonal directions, are assumed to be sufficient to define topographic features. A geometric feature is one of local maxima, local minima, ascending and descending. Then, each pixel is topographically classified by assigned 4 geometric features as follows.

Peak: four local maxima

Ridge: only one local maxima and no local minima,
two adjacent local maxima and no local minima or
three local maxima and no local minima

Saddle: two adjacent local maxima and two local minima,
three local maxima and one local minima or
one local maxima and three local minima

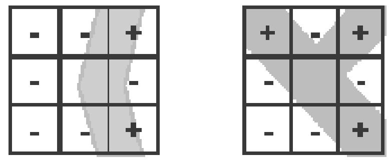


Figure 4: Cases of skeletal pixels which are classified as non-skeleton by 4-directional scanning (Meanings of '+' and '-' are the same as those in figure 3 and skeleton is colored with gray)

Pit, ravine, flat and hillside can be defined in the similar way from Haralick's mathematical descriptions. Figure 3 shows all local intensity configurations of skeletal pixels.

3 Skeletonization Based on Pixel Superiority Index

The main disadvantages of skeletonization based on topographic features are in time complexity and skeleton extraction. In the aspect of efficiency, digital approximation made a successful cost reduction in time complexity [3,4]. But approximation usually involves an error. When we assume only 4 principal directions, small difference in digital sampling can make an error in topographic feature classification(figure 4). On the other hand, there were several attempts to extract skeleton of one-pixel width from the set of skeletal pixels. However, these approaches do not fully utilize the topographic features of the original images. So important information for stroke analysis can be missed. In this paper, we present pixel superiority index as a tool for designing a skeletonization algorithm which utilizes a topography efficiently.

Pixel superiority index(PSI) of a pixel is defined as the number of *inferior* pixels in its neighborhood. Here, *inferior* means having a smaller or equal gray-value. We use 8-neighbors as neighborhood and PSI has a value between 0 and 8. In the extreme case, if PSI of a pixel is 8, the pixel is topographically labeled as peak. From intuition and evidence in figure 3 and 4, two important facts are revealed. The first one is that if a pixel has a *inferior* skeletal pixel in its neighborhood, it is also a skeletal pixel. The second one is that all skeletal pixels have PSIs larger than 4. The only exceptional cases are ravine saddle points, which are known to be meaningless as skeleton [8].

From these observations, we view skeletonization as a sequence of skeleton expansion starting from initial skeleton. Initial skeleton is a set of peak points, namely, pixels of $PSI = 8$. And the initial skeleton is expanded using pixels of $PSI = 7$. This expansion is repeated with smaller PSI and terminates when $PSI = 4$. During expansion, spurious branches are removed and one-pixel

(a) (b) (c)

Figure 5: Pixel superiority index. (a) Grayscale image of a Hangul character, (b) A set of pixels of $PSI \geq 5$, and (c) A set of pixels of $PSI \geq 6$

width property is preserved in skeleton. That is, skeleton is generated in the opposite way of conventional thinning.

However, this approach has a few defects in practice. It is too costly to detect spurious branches during expansion. And expansion is usually accompanied by a search process, which is time-consuming. To avoid the problem initial skeleton is generated from a set of pixels of larger than some fixed PSI and skeleton expansion is repeated only on lower levels. Experimentally we confirmed that when PSI is equal to or larger than 6, there are almost no spurious branches in the set of pixels(figure 5). So we set the relevant PSI value as 6. In that case, the proposed skeletonization algorithm consists of only two stages; initial skeleton generation and skeleton expansion.

Initial skeleton is generated from S, a set of pixels of $PSI \geq 6$ by thinning pixels in topographic order. Generally, a pixel of the larger PSI is topographically more important. So we define topographic order as the order of PSI. That is, initial skeleton is a set of pixels which are left after removing boundary pixels from S in the order of PSI of pixels. Boundary pixels are removed if the removal doesn't change 8-connectivity of S. Pixel removal tests are well explained in a survey paper [5].

As a next stage, initial skeleton is expanded by simple connected component analysis. Because initial skeleton is a subset of skeletal pixels, connectivity in skeletal pixels may not preserve. Each connected component in initial skeleton is examined whether it is connected to another component through pixels of $PSI = 5$ by a labeling technique(figure 6). Pixels of $PSI = 5$ are labeled as chessboard distance from initial skeleton. If there exists another component connected to it, then two nearest pixels of each components are linked following the labeling order. At this time, spurious branches are detected and removed

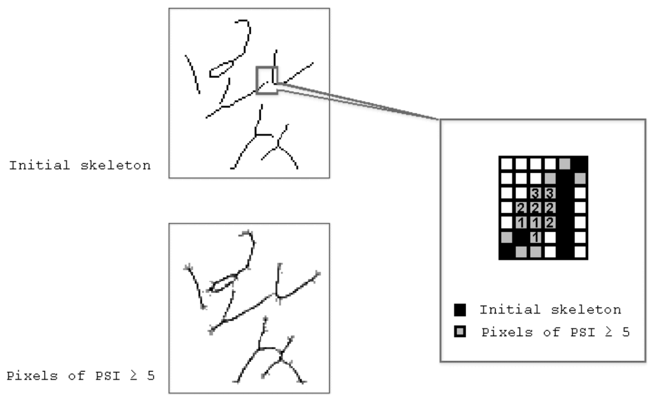


Figure 6: Strokes of initial skeleton that are connected through skeletal pixels are linked by simple connected component analysis.

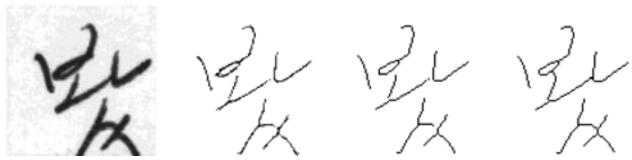
taking into account the length of a branch and the gray-values of pixels in the branch [6].

4 Experimental Results

Real images obtained from scanner have many low-level hills, which can be classified as skeletal pixels. Low-level thresholding was applied to remove such noise. Then average smoothing operator with 3 by 3 mask was used to smooth the surface of character images. Average smoothing has similar effects with less computation compared to Gaussian smoothing because only integer operations are used. Figure 7 shows the overall process of the proposed algorithm.

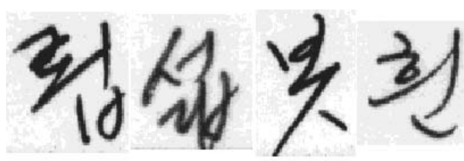
Examples in figure 8 show some Hangul characters and their skeletons generated by our algorithm, which are compared to those of Zhang and Suen's thinning algorithm [10]. Zhang and Suen's algorithm was applied after thresholding grayscale images by dynamic thresholding [9]. Shape distortions induced by binarization were recovered by using the proposed algorithm.

To evaluate the performance of the algorithm, we measured the correct classification rate of a stroke-based structural Hangul recognizer which had been developed at KAIST [11]. For training, 52,000 characters of 520 classes were used. For testing, 13,000 characters of 520 classes were used. In training, Zhang and Suen's thinning algorithm after dynamic thresholding was applied to extract strokes from grayscale images. The comparison result of our algorithm with Zhang and Suen's thinning algorithm is shown in table 1. The



(a) (b) (c) (d)

Figure 7: The overall skeletonization process. (a) Original grayscale image, (b) Pixels of $PSI \geq 6$, (c) Initial skeleton, and (e) Skeleton



(a)

(b)

(c)

Figure 8: Skeletonization results. (a) Grayscale images, (b) Skeletons generated by the proposed method, (c) Skeletons generated by Zhang and Suen's thinning algorithm

	Recognition rate for test data
The proposed algorithm	85.5%
Dynamic thresholding + Zhang and Suen's Thinning	82.4%

Table 1: The correct classification rate of a stroke-based structural Hangul recognizer

error was reduced by 17.5% using the proposed algorithm. We expect more improvements to be made if the proposed algorithm is used in training the recognizer.

5 Conclusion

We proposed a novel grayscale skeletonization algorithm based on pixel superiority index(PSI). Using the relationship between PSI and topography of an image surface, we modified the problem of skeletonization from boundary pixel removal to skeleton growing. The preliminary experimental results showed that the performance of the algorithm is better than a conventional binary thinning algorithm with dynamic thresholding. Applying the proposed algorithm to a stroke-based structural Hangul recognizer, 17.5% of error reduction was obtained.

Ongoing works include extraction of information on stroke interaction in grayscale images. The proposed algorithm doesn't utilize topographic features such as saddle points to reduce time complexity; additional mask operations are needed to extract particular topographic features in our method. For higher performance of a structural recognizer, such information is indispensable. One possibility of utilizing saddle points is to find local curvature-maximum points along skeletal pixels, or ridge lines.

References

1. Layne T. Watson, Thomas J. Laffey, and Robert M. Haralick, Topographic Classification of Digital Image Intensity Surfaces Using Generalized Splines and the Discrete Cosine Transformation, *Computer Vision, Graphics, and Image Processing*, Vol. 29, pp. 143-167, 1985
2. Li Wang and Theo Pavlidis, Direct Gray-Scale Extraction of Features for Character Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1053-1066, 1993
3. S.W. Lee and Y.J. Kim, Direct Extraction of Topographic Features for

Gray Scale Character Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 7, pp. 724-729, 1995
4. Jang Won Suh and Jin H. Kim, Stroke Extraction from Gray-Scale Character Image, in *Proc. IWFHR-V*, pp. 365-368, 1996
5. Louisa Lam, Seong-Whan Lee and Ching Y. Suen, Thinning Methodologies – A Comprehensive Survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 9, pp. 869-885, 1992
6. S. Ablameyko, C. Arcelli and G. Ramella, Removing Noninformative Subsets During Vectorization of Gray-Level Line Patterns, *SPIE*, Vol. 2727, pp. 1514-1521, 1996
7. J.-I. Toriwaki and T. Fukumura, Extraction of Structural Information from Gray Pictures, *Computer Vision, Graphics and Image Processing*, Vol. 7, pp. 30-51, 1978
8. Ø.D. Trier, T. Taxt, and A.K. Jain, Recognition of Digits in Hydrographic Maps: Binary Versus Topographic Analysis, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 19, No. 4, pp. 399-404, 1997
9. J.N. Kapur, P.K. Sahoo, and A.K.C. Wong, A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram, *Computer Vision, Graphics and Image Processing*, Vol. 29, pp. 273-285, 1985
10. T. Y. Zhang and C. Y. Suen, A Fast Parallel Algorithm for Thinning Digital Patterns, *Communications of ACM*, Vol. 27, No. 3, pp. 236-239, 1984
11. Ho-Yon Kim and Jin H. Kim, Handwritten Hangul Recognition Using Extended Hierarchical Random Graph, in *Proc. IWFHR-VI*, pp. 577-586, 1998