Segmentation of Bilevel Images Using Mathematical Morphology

Jin–Chang Cheng and Hon–Son Don
Department of Electrical Engineering
State University of New York
Stony Brook, NY 11794

Abstract

This paper presents the results of a study on the use of morphological skeleton transformation to segment grayscale images into bilevel images. When bilevel image is digitized, the result is a grayscale image due to the point spread function of digitizer, non-uniform illumination and noise. Our method can recover the original bilevel image from the grayscale image. The theoretical basis of the algorithm is the physical structure of the skeleton set. A connectivity property of the grayscale skeleton transformation is used to separate and remove the background terrain. The object pixels can then be obtained by applying a global threshold. Experimental results are given.

Keywords: segmentation, mathematical morphology, thresholding

1. Introduction

Assume that the input gray scale image is intended to have only two levels, such as the printed texts and line drawings. The most commonly used method in extracting object region from background is "thresholding". Thresholding classifies the pixels of a given image into two groups, that is, object and background. If the object is clearly distinguishable from the background, the gray scale histogram will be bimodal and the threshold for segmentation can be easily chosen as the bottom of the valley. However, gray scale histograms are not always bimodal. Several methods have been proposed to solve this problem so that the valley seeking technique can still be applied. Some other publications on automatic selection of thresholding values include iterative method [9], illumination-independent contrast measure [11], moment-preserving method [13], and methods based on entropy of the histogram [3, 8], on local estimation of uniform error analysis [1] and on other image statistics. Some aforementioned algorithms use simple global threshold while others use (multiple) local thresholds. However, simple global thresholding of such images produces poor results due to the nonuniformity of the light distribution across the images. Local thresholding techniques need multiple passes over the data. This causes fast data throughput difficult. Recently, Pavlidis and Wolberg [6] proposed a method based on a model of distortion of the images and tried to invert the distortion process. Another iterative method was proposed by Pérez and Gonzalez [7]. They used Taylor series expansion method on an illumination-reflectance model. Like many iterative algorithms, the result of their method depends on the selection of initial values.

Mathematical morphology was first introduced by Matheron[5] and Serra[10]. It is characterized by its robustness, speed, accuracy and flexibility. It can extracts information about the geometrical structure of an object by transforming it with another smaller and simpler object, called structuring element. Many applications of mathematical morphology have been proposed in literature, such as noise removal, feature detection (edges, holes, corners, ..., etc), thickenings and thinnings. Skeletonization by using binary mathematical morphological operations was introduced by Serra[10].

Our major topic in this paper, is the segmentation of bilevel image. It is done by first extracting the skeleton of the given gray scale image by the gray scale morphological operations and then using the connectivity property of the skeleton to segment the given image into bilevel image. For many kinds of image, like machine parts, printed texts and line drawings, the object pixels and the noises can only be inscribed by spheres of small size, while the pixels in the relatively smooth background can be inscribed by spheres of much larger size. The centers of the inscribing spheres (i.e., skeleton points) of these different kinds of regions usually are quite distance apart. Therefore they can be easily separated. After properly separating spheres of small size from large ones, we retain only the foreground (object or noise) pixels and remove the background pixels. A global thresholding can then be applied to extract the object pixels.

In our experiments, the algorithm is compared against global thresholding scheme, a local thresholding scheme reported by White and Rohrer[14] as well as a segmentation algorithm reported by Pavlidis and Wolberg[6].

In Section 2, the basic morphological set transformations are summarized. In Section 3, we discuss the morphological skeleton representation of grayscale images. In Section 4, we propose the algorithms for bilevel image segmentation. Experimental results and conclusions are given in Section 5 and Section 6, respectively.

This work was supported by the National Science Foundation under Grant IRI-8710856 and U.S. Army Research Office under Contract DAAL 0388K0033.

2. Basic Morphological Set Operations

In the following, we summarize some basic morphological transformations. The details of these formula can be found in [2].

Let E^N denote the set of N-tuple integers, A, B be subsets of E^N, x be a vector in E^N.

Definition 1 : The binary dilation $\oplus$ of A by B is defined by

$$A \oplus B = \{ c \in E^N \mid c=a+b, \text{ for some } a \in A \text{ and } b \in B \}$$

Definition 2: The binary erosion $\ominus$ of A by B is defined by

$$A \ominus B = \{ x \in E^N \mid x+b \in A, \text{ for every } b \in B \}$$

Definition 3: The binary opening $\circ$ of A by B is defined by

$$A \circ B = (A \ominus B) \oplus B$$

Definition 4: The binary closing $\bullet$ of A by B is defined by

$$A \bullet B = (A \oplus B) \ominus B$$

Let f and g be functions defined on E^N, x and y are vectors in E^N.

Definition 5: Let $A \subseteq E^{N+1}$ and $F=\{x \in E^N \mid$ for some $y \in E$, $(x,y) \in A\}$. The top (or top surface) of A is defined by

$$T[A](x) = \max\{ y \mid (x,y) \in A \}.$$

Definition 6: The umbra of f, denoted by U[f] is defined by

$$U[f] = \{(x,y) \mid y \le f(x), x \in E^N, y \in E \}.$$

Definition 7: The grayscale dilation of f by B is defined by

$$f \oplus_g B = T[U[f] \oplus U[B]]$$

Proposition 1:

$$(f \oplus_g B)(x) = \max_{b \in B} \{ f(x-b)+B(b) \}. \quad (1)$$

Definition 8: The grayscale erosion of f by B is defined by

$$f \ominus_g B = T[U[f] \ominus U[B]]$$

Proposition 2:

$$(f \ominus_g B)(x) = \min_{b \in B} \{ f(x+b)-B(b) \}. \quad (2)$$

Definition 9: The grayscale opening of f by B is defined by

$$f \circ_g B = (f \ominus_g B) \oplus_g B. \quad (3)$$

Definition 10: The grayscale closing of f by B is defined by

$$f \bullet_g B = (f \oplus_g B) \ominus_g B. \quad (4)$$

Proposition 3: Umbra Homomorphism Theorem

$$U[f \oplus_g B] = U[f] \oplus U[B] \quad (5)$$

$$U[f \bullet_g B] = U[f] \bullet U[B] \quad (6)$$

Other formulas can be found in [2, 5, 10].

3. Morphological Skeleton Transformation of Gray Scale Images

The skeleton S(f) of a discrete N-dimensional binary image f is defined as the set of the centers of the maximal N-dimensional spheres, inscribable inside the image f. An N-dimensional sphere is *maximal* if it is not properly contained in any other spheres which are totally included in f. Hence, a maximal sphere must touch the boundary of f at least at two different points. The morphological skeleton transformation of a discrete binary image has been studied in [5,10]. For gray scale images, the maximal sphere can be similarly defined and the gray scale skeleton transformation can be derived by using the concept of umbra. Let B be a N-dimensional gray scale structuring element with semi-sphere shape of flat bottom. We define the skeleton S(f) of a discrete N-dimensional gray scale image f as the set of the centers of B which the umbra U[B] of B is the maximal umbra inscribable inside the umbra U[f] of f. An umbra of a semi-sphere is maximal if it is not properly contained in any other umbra of semi-sphere totally included in U[f]. Therefore, the gray scale skeleton can be derived as follows:

$$S_n(f) = (U[f] \ominus U[nB]) - (U[f] \ominus U[nB]) \circ U[B] \quad (7)$$

and

$$S(f) = \bigcup_{n=0}^{M} S_n(f)$$

where $S_n(f)$ denotes the n^{th} skeleton subset of U[f], B is the N-dimensional gray scale structuring element and nB is defined as the gray scale dilation of B with itself n times:

$$nB = B \oplus_g B \oplus_g \cdots \oplus_g B \quad (n \text{ times}).$$

By $S_0(f) = U[f] - U[f] \circ_g U[B]$, we can see that the 0^{th} skeleton subset of U[f] consists of those points in U[f] and they can not be inscribed or touched by any unit-size maximal sphere B. Similarly, $S_n(f)$ consists of points which are the translated centers of size-n sphere nB when the translation of U[nB] is inscribable inside U[f] but that of U[(n+1)B] is not. To find the skeleton by using gray scale erosions and openings, we need the following property.

Proposition 4 :

$$S_n(f) = U[f \ominus_g nB] - U[(f \ominus_g nB) \circ_g B] \quad (8)$$

proof: By (6), $U[f] \ominus U[nB] = U[f \ominus_g nB]$. Let $h = f \ominus_g nB$, then by (3)(5), $(U[f] \ominus U[nB]) \circ U[B] = U[h] \circ U[B] = (U[h] \ominus U[B]) \oplus U[B] = U[h \ominus_g B] \oplus U[B] = U[(h \ominus_g B) \oplus_g B] = U[h \circ_g B] = U[(f \ominus_g nB) \circ_g B]$. Therefore, $S_n(f) = U[f \ominus_g nB] - U[(f \ominus_g nB) \circ_g B]$.

The morphological skeleton transformation is invertible. The discrete gray scale image f can be exactly reconstructed as the finite union of its M+1 skeleton subsets dilated by the structuring element B of proper size .

$$f = T[\bigcup_{n=0}^{M} S_n(f) \oplus U[nB]] \quad (9)$$

The structuring element B in (8) can be considered as an unit-distance probe. Skeleton point in $S_n(f)$ has the minimum distance of n to the boundary of image f. The distance measure can be interpreted by using city-block, chess-board, rectangular-solid distance, Euclidean distance or any other distance measures. Note that the first three distances satisfy the equation $nB = B \oplus_g B \oplus_g \ldots \oplus_g B$ (n times). The Euclidean distance does not satisfy this equation. Therefore the Euclidean distance is not suitable to be used as the

unit-size structuring element for the iterative skeleton transformation (8). Although some structuring elements do not look like spheres in shape, we still adapt the term "maximal sphere" to the structuring elements not spherical in shape but are maximal.

4. Segmentation of Bilevel Image

In this section, we present our algorithm for bilevel image segmentation. The basic idea of the segmentation algorithm as well as its computational complexity are discussed in Section 4.1. In Section 4.2, the theoretical analysis of our algorithm are given. The selection of the structuring element and the robustness of the algorithm are discussed in Section 4.3. Finally, a comparison with the Top Hat Transformation[10] is given in Section 4.4.

4.1. Basic Approach

Given an ideal bilevel image f_i whose gray values at pixel (x,y) are $f_i(x,y)=a$ for object pixels, and $f_i(x,y)=0$ for background pixels. Let its blurred image f be denoted by

$$f(x,y) = f_i(x,y)+noise(x,y)+offset(x,y),$$

where noise(x,y) and offset(x,y) are the added noise and the background offset caused by non-uniform lighting distribution, respectively. Our aim is to find a base surface b(x,y) which is the estimation of the offset, i.e., b(x,y)=offset (x,y). Then the bilevel image f_i can be reconstructed by thresholding f(x,y)-b(x,y).

To find the base surface b for a given gray scale image f, we begin with the skeletonization of the input image f by using the gray scale skeleton transformation (8). Recall that the skeleton point is the center of the maximal inscribed spheres. We will call the skeleton point which inscribes the object (background) pixels as *object (background) center*. The skeleton of a connected discrete image may be disconnected. We will explain in the following that the disconnection occurs in the intersection of object pixels and background pixels. Within object pixels, the radius change is continuous and the distance between two neighboring skeleton points are less than or equal to two in noise-free case (slightly larger than two in noisy case). The magic number two is due to the possible increment of the size by one between two consecutive maximal spheres and the one pixel downshift of the larger sphere. Since the sizes of the spheres which inscribe the background pixels are relatively much larger than those of the spheres which inscribe the object pixels. Therefore, in the intersection, the distance between the object center and its successive background center will be much greater than two. This big jump of distance between two successive centers causes the disconnection of the skeleton. By detecting this distance change, we can easily locate those background centers.

Two skeleton points are *connected* if the distance between them is less than or equal to two. Otherwise, we say that they are *disconnected*. For disconnected skeleton points, the skeleton associated with the larger radius is a background center. Any skeleton point connected to a background center is a background center and any skeleton point connected to an object center is an object center.

After the background centers are found, the base surface b can then be constructed from those background centers by using the inverse grayscale skeleton transformation (9). The base surface b is the top of the union of those umbras which are the translation of the umbra of nB by Q, for every background center Q in $S_n(f)$. Here, we list our segmentation algorithm as follows:

Algorithm:

Input: a grayscale image f.

Output: a segmented bilevel image.

Steps:

0. Initialize the base surface b(x,y)=0.
1. Using (8), find the skeleton subsets $S_n(f)$, for $0 \le n \le M$.
2. Construct b by the top of the union of those umbras which are the translation of the umbra of nB by Q, for every background center Q in $S_n(f)$:
$$b = T[\bigcup_Q (U[nB])_Q]$$
3. Remove the background offset by subtraction: g(x,y)=f(x,y)-b(x,y).
4. Thresholding the background-removed image g(x,y).

The value of M in step 1 can be the size of the minimum square mask of M by M pixels which can not be completely contained inside any object in the image. The computation time is dominated by the skeletonization process in step 1. Assume we choose the structuring element B as an 1-D unit-size rectangle (see Fig. 1a) for 1-D line scan process and unit-size rectangular solid for 2-D process. Then, by (2), the grayscale erosion(dilation) of f by B needs 2 parallel min(max) operations in 1-D or 8 parallel min(max) operations in 2-D by using a parallel computer. Since the grayscale erosion(dilation) of f by nB can be done by successively eroding f by B n times, $0 \le n \le M$, step 1 needs $\frac{1}{2}M(M+1)$ grayscale erosions by B plus M+1 grayscale erosions by B and M dilations by B for the opening operations in (8), which amounts to a total of (M^2+5M+2) grayscale erosion/dilation operations. Therefore, step 1 totally takes (M^2+5M+2) and $4(M^2+5M+2)$ parallel min/max operations on a parallel computer for 1-D and 2-D fashion, respectively, for the selected structuring element B. Note that the 2-D fashion takes 4 times longer than 1-D line scan fashion.

4.2. Theoretical Analysis

For the simplicity of description, we assume the image is a one-dimensional image which can be the row- (or column-) version of a two-dimensional image. All the proper-

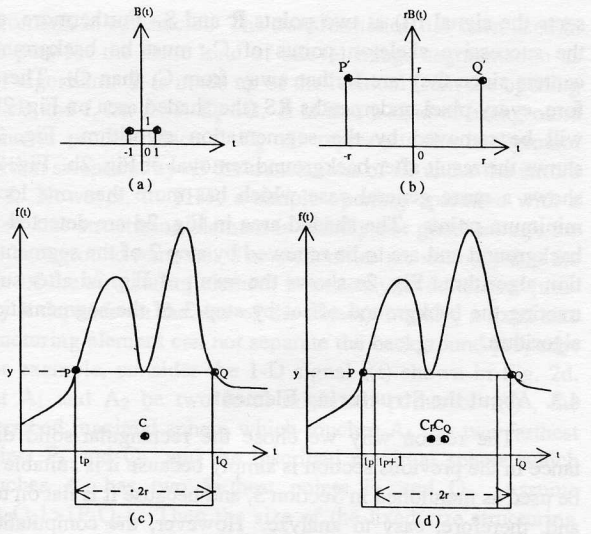

Fig. 1: (a) 1-D rectangular unit-size structuring element, (b) 1-D rectangular size-r structuring element, (c) A maximal inscribed size-r sphere with center at C, touchs the local minimum point and two points P and Q on the discrete signal f, when $t_P - t_Q$ is even, (d) Two neighboring maximal inscribed spheres with centers at C_P and C_Q, respectively, when $t_P - t_Q$ is odd.

ties derived below can be naturally and easily extended to two-dimensional image. We also assume the pixels on the image boundary belong to the background. In the continuous image case, a maximal sphere inscribed in f must touch the top of f at least at two points. These two points which have maximum distance among those touched points are called *two farthest points*. While in discrete image case, due to the nature of discrete skeleton transformation (8), the discrete skeleton may be of width up to two. The two farthest points may be touched by one maximal sphere or by two neighboring maximal spheres of the same size. In either case, without loss of generality, we still say that these two farthest points are *touched by a maximal sphere*. First, we give the following observation:

Proposition 5 : Assume the structuring element B is a 1-D unit-size rectangle (see Fig. 1a) , that is, $B(t)=1$ for $|t|\le 1$, and $B(t)$ is undefined for $|t|>1$. Let $P=(t_P, f(t_P))$ and $Q=(t_Q, f(t_Q))$ be two farthest points which are touched by a maximal sphere rB of radius r, $t_P \le t_Q$, then

(i) $f(t_P)=f(t_Q)$,

(ii) $r=\frac{1}{2}\lfloor t_Q - t_P \rfloor$, and $t_Q - t_P = 2r$, if $t_Q - t_P$ is even; $t_Q - t_P = 2r+1$, if $t_Q - t_P$ is odd,

(iii) if and only if $t_Q = t_P$ or $t_Q = t_P+1$ then $r=0$,

(iv) If $t_Q - t_P$ is even, then the inscribed maximal sphere is centered at $(t_P+r, f(t_P)-r)$, otherwise P and Q are touched by two neighboring maximal spheres which are centered at $(t_P+r+1, f(t_P)-r)$ and $(t_P+r, f(t_P)-r)$, respectively.

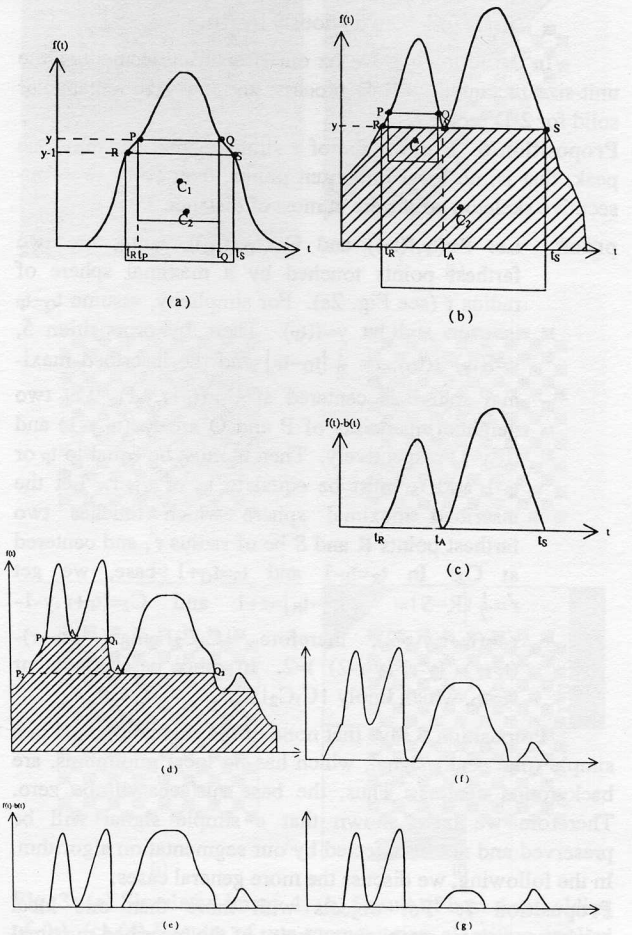

Fig. 2: (a) 1-D representation of bilevel one-peak signal, marked with two consecutive and connected skeleton points C_1 and C_2, (b) A two-peak signal, marked with two consecutive but disconnected skeleton points C_1 and C_2, respectively. Pixels in shaded area are background pixels, (c) Result of the signal in Fig. 2b after the removal of the shaded area (background). Note that all signals are plotted with continuous line, they actually represent discrete signals.

proof: Since B is an unit size rectangular and $rB = B \oplus_g B \oplus_g \cdots \oplus_g B$ (r times), therefore rB is a size-r rectangular, i.e., $rB(t)=r$ for $|t|\le r$ and $rB(t)$ is undefined for $|t|>r$ (see Fig. 1b). When $t_Q - t_P$ is even (see Fig. 1c), umbra of rB slides around and fit within umbra of f, the two corner points P′ and Q′ in rB will match f at two farthest points $P=(t_P, f(t_P))$ and $Q=(t_Q, f(t_Q))$, respectively. Since $rB(-r)=rB(r)=r$, we get $t_Q - t_P = |P-Q| = |P'-Q'| = 2r$, and the skeleton point is at $(\frac{1}{2}|P-Q|, f(t_P)-r) = (t_P+r, f(t_P)-r)$. When $t_Q - t_P$ is odd (see Fig. 1d), P will be touched by the maximal sphere rB of centered at $(t_P+r, f(t_P)-r)$ and matched P while Q will be touched by the maximal sphere rB of centered at

$(t_P+r+1,\ f(t_P)-r)$ and matched Q'. So, $t_Q-t_P=2r+1$. (iii) is followed obviously by (ii).

In the following, we fix our structuring element as the unit-size rectangle for 1-D process, and unit-size rectangular solid for 2-D process.

Proposition 6: If the shape of a simple signal has only one peak, that is, no local minimum points, then every two consecutive skeleton points are at most of distance 2.

proof: Let $P=(t_P,f(t_P))$ and $Q=(t_Q,f(t_Q))$, $t_P<t_Q$, are two farthest points touched by a maximal sphere of radius r (see Fig. 2a). For simplicity, assume t_Q-t_P is even and let $y=f(t_P)$. Then, by proposition 5, $y=f(t_P)=f(t_Q)$, $r=\frac{1}{2}\lfloor t_Q-t_P\rfloor$ and the inscribed maximal sphere is centered at $C_1=(t_P+r,y-r)$. Let two neighboring points of P and Q are $R=(t_R,y-1)$ and $(t_S,y-1)$ respectively. Then t_R must be equal to t_P or t_P-1 and t_S must be equal to t_Q or t_Q+1. Let the inscribed maximal sphere which touches two farthest points R and S be of radius r', and centered at C_2. In $t_R=t_P-1$ and $t_S=t_Q+1$ case, we get $r'=\frac{1}{2}|R-S|=\frac{1}{2}\lfloor t_S-t_R\rfloor=r+1$ and $C_2=(t_P+r',y-1-r')=(t_P+r,y-r-2)$, therefore $|C_1C_2|=\max((t_P+r)-(t_P+r),(y-r)-(y-r-2))=2$. In other cases, $t_R=t_P$ or $t_S=t_Q$, which imply $|C_1C_2|\le 2$.

Proposition 6 says that none of the skeleton points in a simple (one peak) signal, which has no local minimums, are background centers. Thus, the base surface will be zero. Therefore we have shown that a simple signal will be preserved and not be affected by our segmentation algorithm. In the following, we discuss the more general cases.

Proposition 7: For objects with more than one local minimum points, any skeleton point of the inscribed maximal sphere which touches one of the local minimum points must be a background center.

proof: Let $A=(t_A,f(t_A))$ be an local minimum point of f, $y=f(t_A)$, and the inscribed maximal sphere which touches A is of radius r, centered at C_2, and touches f at two farthest points $R=(t_R,f(t_R))$ and $S=(t_S,f(t_S))$, $t_R+1<t_A<t_S-1$, respectively (see Fig. 2b). By proposition 5, we have $f(t_A)=f(t_R)=f(t_S)=y$, $r=\frac{1}{2}\lfloor t_R-t_S\rfloor$ and the center $C_2=(t_R+r,f(t_A)-r)$. Assume points $P=(t_P,f(t_P))$ and $Q=(t_Q,f(t_Q))$ are two respective neighbors of points R and A, and the inscribed maximal sphere which touches points P and Q is of radius r' and centered at C_1, then we have $f(t_P)=f(t_Q)=f(t_A)+1$, $t_P=t_R$ or $t_P=t_R+1$, $t_Q<t_A$, $r'=\frac{1}{2}\lfloor t_Q-t_P\rfloor$ and center $C_1=(t_P+r',f(t_P)-r')$. Therefore, $|C_1C_2|=\max(|t_R+r-t_P-r'|,|f(t_A)-r-f(t_P)+r'|)=r-r'+1$. Since $r=\frac{1}{2}\lfloor t_R-t_S\rfloor$, $r'=\frac{1}{2}\lfloor t_Q-t_P\rfloor$, we immediately get $|C_1C_2|>2$, and therefore C_2 is a background center.

By proposition 7, every skeleton points C_2 touches one of the local minimum points must be a background center. The translation of rB to C_2 is nothing but the inscribed maximal sphere with center at C_2. Its top surface makes a horizontal baseline RS which passes the minimum point and intersects the signal f(t) at two points R and S. Furthermore, all the successive skeleton points of C_2 must be background centers since they are farther away from C_1 than C_2. Therefore, every pixel underneaths RS (the shaded area on Fig. 2b) will be removed by the segmentation algorithm. Fig. 2c shows the result after background removal of Fig. 2b. Fig. 2d shows a more general case which has more than one local minimum points. The shaded area in Fig. 2d are detected as background and are to be removed by step 2 of the segmentation algorithm. Fig. 2e shows the result of Fig. 2d after subtracting the background offset by step 3 of the segmentation algorithm.

4.3. About the Structuring Element

The reason why we chose the rectangular-solid distance in the previous section is simply because it is suitable to be used as mentioned in Section 3, and because it is flat on top and, therefore, easy to analyze. However, the computation time as well as the segmentation results will be the same by using any one of the three distance measures.

By using these kind of structuring elements, the connectivity of the skeleton may be biased and the spatial localization may be inaccurate. However, our goal is the segmentation of bilevel images, not the representation of the image structures by using the skeleton. These two problems will not affect the results of segmentation, since the algorithm will transform back to the original signal shape by (9) after the background are removed. The shape of the original signal above the background terrain will be preserved no matter which one of the structuring elements is used.

Another nice property of this algorithm is that it uses a fixed unit-sized structuring element for all kinds of input images. This property makes the automation possible. In addition, most morphological hardware architecture implementations are limited to a fix-sized structuring element. However, the choices (of the shape, size and/or orientation) of the structuring element are varying in most applications using mathematical morphology, like feature extraction, edge detection, etc, depending heavily on the image to be processed. For example, the choice of the structuring element to extract features depends on the size of the interested features and the requirement of how accurate the boundary localization is. If the size of the structuring element is too small, noises as well as unnecessary details may be extracted. If it is too large, some small features may not be extracted. Since our segmentation algorithm uses the structuring element only as a distance measure, the choice of structuring element is relatively not critical and is independent of the size and the orientation of the images. Any small size structuring element should be appropriate. No a priori knowledge of the size of the object and the orientation of the light source(s) is necessary. Therefore, this segmentation algorithm is robust.

4.4. Comparison with Top Hat Transformation

Top Hat Transformation (THT) [10] was originally proposed by Meyer and was applied to the extraction of

chromatin in cell nuclei. The transformation has been used in the past for the same kind of bilevel image segmentation as our algorithm. It is made up of the residual from the opening by rB and f, i.e. f - f $\circ_g$ rB. It is also a kind of background removal process since f $\circ_g$ rB represents the background portion swept out by all the translations of rB under the umbra of f. However, in THT, a simple opening operation with a fixed size structuring element is usually not good enough for the background removal. The difficulties are not only that the size of the structuring element is not easy to determine (usually it depends on the size of the object), but also a fixed size structuring element can not separate the background properly. For example, consider the 1-D signal f(t) shown in Fig. 2d. Let A_1 and A_2 be two local minimum points of f(t) , the inscribed maximal sphere which touches A_1 has two farthest points P_1 and Q_1, and the inscribed maximal sphere which touches A_2 has two farthest points P_2 and Q_2. Assume $|A_2Q_2| > |P_1Q_1|$. Then the size of the fixed size structuring element is either $\geq$ or $< \frac{1}{2}|P_1Q_1|$. In the first case, if the size is $\geq \frac{1}{2}|P_1Q_1|$, then the two peaks in-between P_1A_1 and A_1Q_1 can not be separated because the structuring element can not touch the minimum point A_1 (see Fig. 2f). On the other hand, if the size of the structuring element is less than $\frac{1}{2}|P_1Q_1|$, then the base under the peak in-between A_2Q_2 will touch the small structuring element too much, and will thus be removed too much. As a result, the peak in-between A_2Q_2 will be narrowed and distorted (see Fig. 2g). Therefore the background can not be removed completely or the signal may be altered in either case by using a fixed size structuring element.

On the contrary, our segmentation algorithm uses a fixed, unit size structuring element, which will preserve the original shapes of all peaks (see Fig. 2e).

5. Experimental Results

In this section, we present some experimental results to show the performance of our segmentation algorithm. We used a simulated checkerboard image under two nonuniform illuminations and two real images for the experiments. For the purpose of comparison, the images in all experiments were segmented by 1) our algorithm, 2) a global threshold selected interactively to yield "optimum" visual results, 3) the White-Rohrer algorithm [14], and 4) the Pavlidis-Wolberg algorithm [6]. The White-Rohrer algorithm has been used successfully in character extraction applications.

The simulated image is a checkerboard of size 256 × 256 pixels. The light regions have gray scale intensity 40 and the dark regions have gray scale intensity 0. To simulate the nonuniform illumination effect, illumination patterns were added to this perfect checkerboard. Fig. 3a shows the checkerboard image superimposed an illumination pattern which changes linearly in the x and y directions plus a jump in the middle of the pattern, i.e.,

$$\text{offset}(x,y) = k_1 \cdot x + k_2 \cdot y + k_3 \cdot u$$

where k_i's are constants and u=0, if $x \leq x_0$ and $y \leq y_0$; u=2, if $x > x_0$ and $y > y_0$; otherwise u=1. The jump simulates the situation that some part of the image are under shadow. Fig. 3b shows the result of thresholding the image in Fig. 3a. The difficulty of thresholding the image is obvious. The image shown in Fig. 3a is easily segmented by our segmentation algorithm. We applied it in both 2-D fashion and 1-D row scan fashion, respectively, and got the same good results (Fig. 3c and Fig. 3d). Note that the nonuniform illumination effect is completely removed.

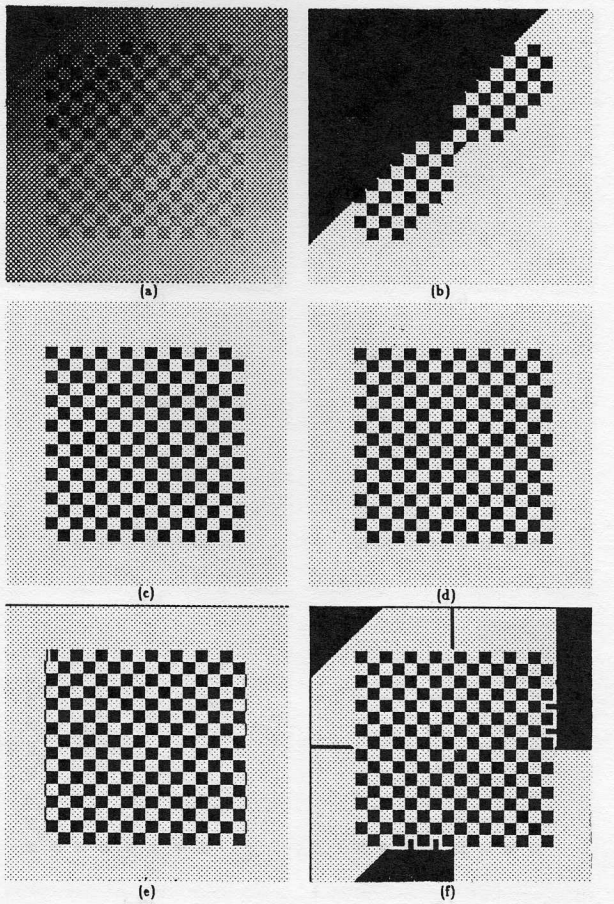

Fig. 3: (a) Simulated image 1, (b) Result of best thresholding of (a), (c)-(d) Results of our segmentation algorithm applied to (a), by using 2-D and 1-D scan fashion, respectively. (e) Results of White-Rohrer[14] algorithm, (f) Results of Pavlidis-Wolberg[6] algorithm.

Fig. 4a shows the checkerboard image added to a Gaussian illumination pattern, defined by

$$\text{offset}(x,y) \propto \exp\left[\frac{-[(x-x_0)^2+(y-y_0)^2]}{\sigma^2}\right]$$

with σ=62. Fig. 4b shows the result of thresholding the image in Fig. 4a. Using our segmentation algorithm in both 2-D fashion and 1-D row scan fashion, respectively, we got the same good results (Fig. 4c and Fig. 4d).

Fig. 5a shows a text image of size 192 × 512 pixels. The data were obtained by a vidicon camera with a nonuniform lighting condition. Thresholding result of the text image is also shown in Fig. 5b. Fig. 5c and Fig. 5d show the results

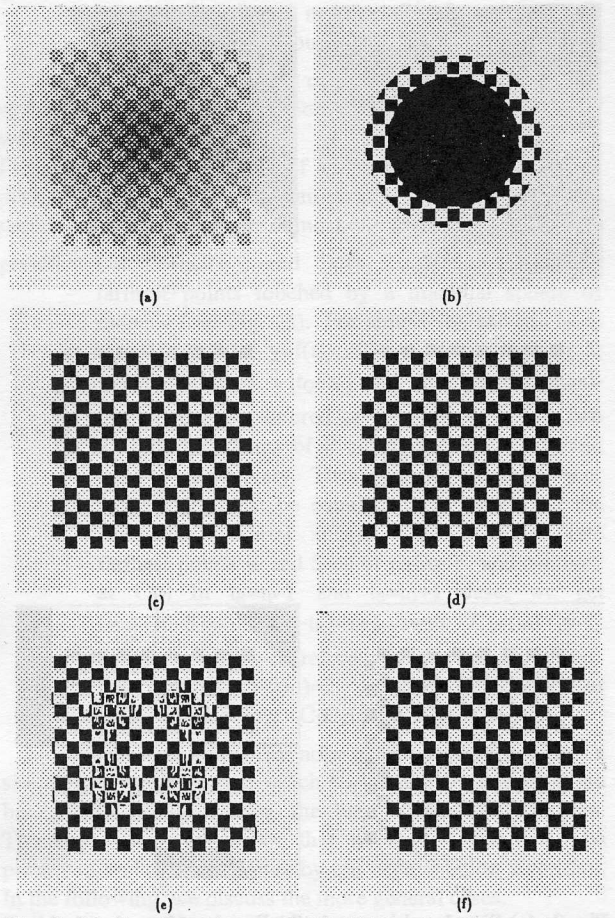

Fig. 4: (a) Simulated image 2, (b)-(f) Results of algorithms as in Fig. 3.

after segmentation by our algorithm using 2-D fashion and 1-D scan fashion, respectively. Fig. 5e and Fig. 5f show the segmented results by the White-Rohrer algorithm and the Pavlidis-Wolberg algorithm, respectively.

Experimental results using tool images are shown in Fig. 6 and Fig. 7. Fig. 6a and 7a show two tool images of size 256 × 256 pixels. Fig. 6b and 7b show the results obtained by global thresholding. Fig. 6c and 7c show the results after segmentation by using our algorithm in 2-D scan fashion and Fig. 6d and 7d show the results in 1-D scan fashion. Fig. 6e and 7e show the results by the White-Rohrer algorithm. Fig. 6f and 7f show the results by the Pavlidis-Wolberg algorithm.

6. Discussion and Conclusions

We have proposed a bilevel image segmentation algorithm by using gray scale morphological operations. The morphological skeleton transformations on binary images are extended to the gray scale images, which forms the basis of our segmentation algorithm. The segmentation algorithm skeletonizes the input gray scale image first, then check for discontinuity of skeleton set. The background due to the nonuniform illumination effect is then removed according to the discontinuity information. Finally, a global thresholding

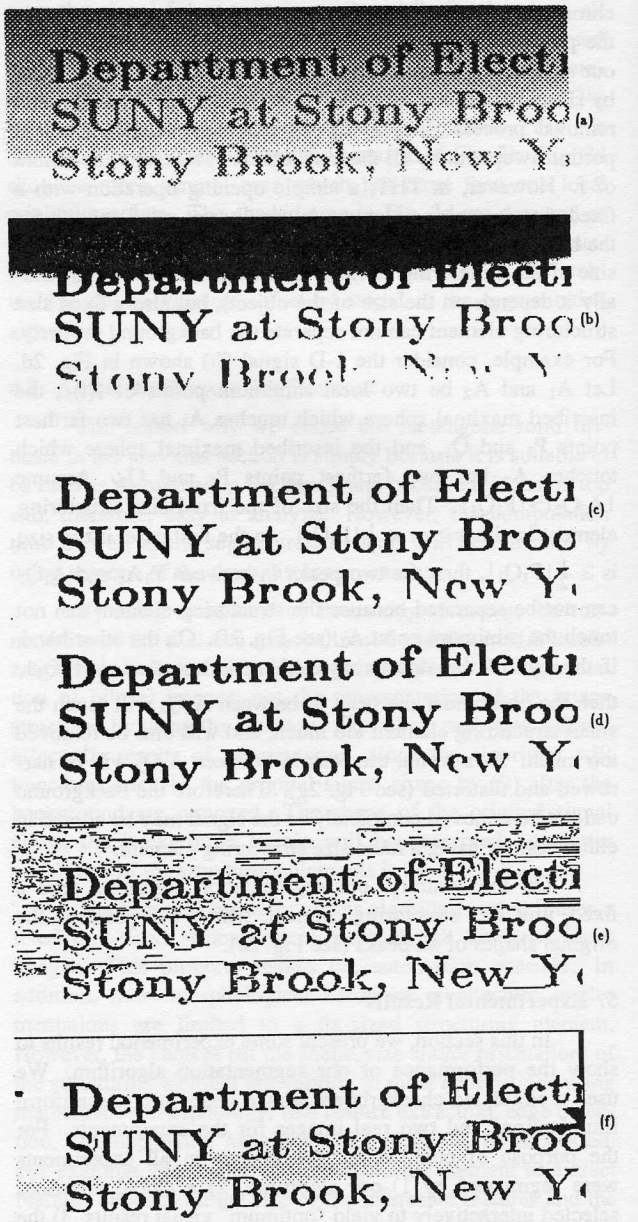

Fig. 5: (a) Text image, 192 × 512 pixels, (b)-(f) Results of algorithms as in Fig. 3.

is applied to get the proper segmentation result. The algorithm is also shown to be better than the best manually selected global threshold under different illumination conditions. It is also better than a local threshold algorithm proposed by White and Rohrer and another algorithm proposed by Pavlidis and Wolberg. A visual comparison from the experimental results leaves the impression that this segmentation algorithm running in 1-D scan fashion can give as good results as that running in 2-D fashion. However, the computation time is four times faster by using 1-D scan fashion.

References

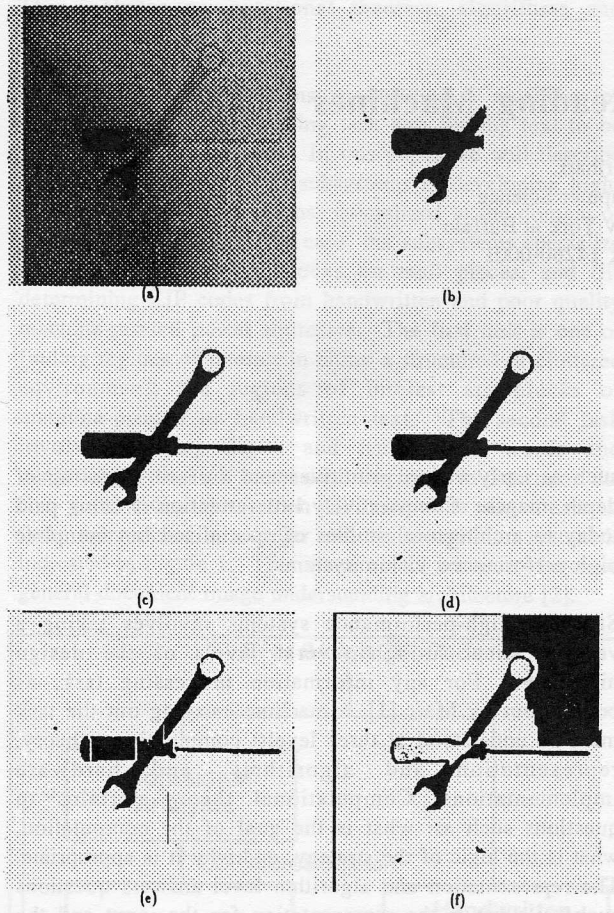

Fig. 6: (a) Tool image 1, 256 × 256 pixels, (b)-(f) Results of algorithms as in Fig. 3.

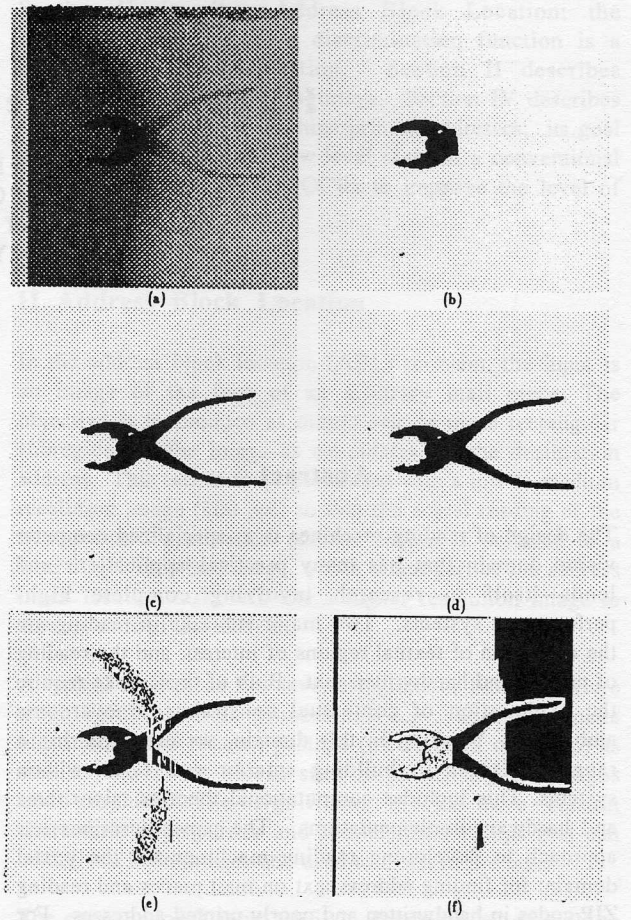

Fig. 7: (a) Tool image 2, 256 × 256 pixels, (b)-(f) Results of algorithms as in Fig. 3.

[1] S. M. Dunn, D. Harwood, and L. S. Davis, "Local estimation of the uniform error threshold", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, no.6, November 1984.

[2] R.M. Haralick, S.R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology", *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, no.4, July 1987.

[3] 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, no.3, pp.273-285, 1985.

[5] G. Matheron, *"Random sets and integral geometry"*, New York, Wiley, 1975.

[6] T. Pavlidis and G. Wolberg "An algorithm for the segmentation of bilevel images", *Proc. CVPR-86,* Miami Beach, Florida, pp.570-575, June, 1986

[7] A. Pérez and R. C. Gonzalez, "An iterative thresholding algorithm for image segmentation", *IEEE Trans. Pattern Anal. Mach. Intell.* vol. PAMI-9, no.6, pp.742-751, 1987.

[8] T. Pun, "Entropic thresholding: A new approach", *Computer Graphics Image Processing,* vol.16, no.3, pp.210-239, 1981.

[9] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method", *IEEE Trans. Systems Man Cybernet,* vol. SMC-8, no.8, pp.630-632, 1978.

[10] J. Serra, *"Image analysis and mathematical morphology"*, Academic Press, New York, 1982.

[11] A. Shio, "An Automatic Thresholding Algorithm Based on an Illumination-Independent Contrast Measure", *Proc. CVPR-89,* pp.632-637, June 1989.

[12] S. R. Sternberg, "Biomedical image processing", *Computer,* vol.16, no.1, pp.22-34, January 1983.

[13] W. H. Tsai, "Moment-preserving thresholding: A new approach", *Computer Vision, Graphics and Image Processing,* vol.29, no.3, pp.377-393, 1985.

[14] J. M. White and G. D. Rohrer, "Image thresholding for optical character recognition and other applications requiring character image extraction", *IBM J. Res. Develop.*, vol.27, no.4, pp.400-411, 1983.