Image Understanding with Connected Component Histograms

Yang Wang
Computer Science Dept.
Southwest Missouri State University, Springfield, MO 65804

Prabir Bhattacharya
Dept. of Computer Science and Engineering
University of Nebraska-Lincoln, Lincoln, NE 68588

Abstract

In this paper, we propose a method for image understanding by applying a theory of parameter-dependent connected components developed by us in a previous work. We may study various properties of an image at the connected component level. Using the information obtained from various component histograms, we can understand the structure of images in either micro-view or macro-view.

1. Introduction

Although the human brain can quite successfully deal with image understanding, it is still unclear how to instruct a computer to perform that task with comparative success. Any given image has certain properties or features of its own. Are we aware of all of them? If not, to what level can we understand the image by using a subset of its properties? Researchers have studied the properties of images from various directions, such as the distribution of the gray values of pixels [1, pp. 92–94], [2, vol.1, pp. 231–237], connected components [2, vol.2, pp. 241–244], [3] and transformations [2, vol.1, pp. 13–29], [4]. We believe that image understanding may be thought of as a process which is composed of three levels – pixel, component and content understanding. At the pixel understanding level, we concentrate on the properties of the pixels in the image. We may understand the structure of an image in very detail but may have no sense of the content of the image. At the content understanding level, we concentrate on the properties of some groups of pixels. Usually, each group of pixels represents a meaningful object. Content understanding is a complex intellectual process and this is a *high level understanding*.

In practical problems, it is not enough to understand an image only at the pixel level. We often need to understand the content of an image. This raises the question: how to group the pixels in an image such that each group of pixels represents an object of the image under consideration? In this paper, we shall refer to any *connected subset* [5] of pixels of an image as a *connected component* of the image. For any given grouping method, the pixels of an image may be grouped into a set of connected components. Thus, we may understand the structure of the image in terms of its connected components besides its pixels. However, for an arbitrary grouping method, a connected component of an image may not necessarily represent a meaningful object in the image. So, we may regard understand images by its connected components as the component understanding level or *intermediate level* of image understanding.

There are many ways to group the pixels in an image. A traditional definition for a connected component of an image is a maximal connected subset of the pixels in the image which have the same gray value [3]. In a binary image, the objects are usually represented by either black pixels or white pixels. The connected components of the image can be easily obtained by the technique of thresholding [2, vol. 2, pp. 61–66]. In this case, the properties of the connected components, such as size, location, are corresponding to the properties of the objects of the images. In a multi-gray-value image, an object usually contains pixels which have different gray values. If we limit each connected component to have only one gray value, an object might be "cut" into several pieces and each piece belongs to a different connected component. We may obtain the connected components of a multi-gray-value image by using multi-thresholding [2, vol. 2, pp. 66–68] or some other modified thresholding technique [2, vol. 2, pp. 68–71], [6, 7, 8, 9, 10]. In this case, the properties of connected components usually do not correspond to the properties of the objects in the images. However, in practical situations, for a comprehensive image understanding we often would like to know the relationship among the objects, rather than among the "parts" of the objects. So, we feel that it would not be convenient in practice to con-

straint each connected component to have the same gray value. Instead, it may be more useful to relax this condition to allow the pixels in the same component to have different gray values, so that each object of an image may correspond to a connected component.

Recently, we have introduced [5] the concept of (ϵ, δ)-components of gray images that takes into account both the gray values of the pixels and the differences of the gray values of the neighboring pixels, where ϵ, δ are two parameters. Each (ϵ, δ)-component may contains the pixels which have different gray values. We have discussed [5] some properties of (ϵ, δ)-components which may help us to analyze and understand the structure of an image in a higher level. In [5], we described an algorithm to find the (ϵ, δ)-components for a given image and the values of the parameters. The experimental results gave in [11, 12] have show that for some appropriate parameter values, an (ϵ, δ)-component may represent an object of an image reasonably well. So, the properties of the (ϵ, δ)-components describes the properties of the corresponding objects of the image. Thus, we may understand the content of an image through the properties of its (ϵ, δ)-components. In this paper, we shall discuss how to analyze and understand the structure of an image in terms of its (ϵ, δ)-components.

2. Preliminaries

To make this paper self-contained, we review briefly some definitions from our previous work [5].

A gray image Σ is represented by a set of points each of which has a certain *gray value* representing the intensity of brightness of the point. We shall use $\sigma(p)$ to denote the gray value of the point p. Although, theoretically, $\sigma(p)$ could be any number, we shall take it, for convenience, to be a non-negative integer. Two points in a gray image Σ are called *adjacent* if they share either a vertex or an edge. Our treatment does not depend on the grid system chosen to represent an image and the way in which the points share vertices or edges. Instead of considering whether the points are 4-, 6- or 8-neighbors (see [3] or [13] for definitions), we shall simply consider here only whether two points are adjacent or not.

A *path* between two points p_0 and p_n in a gray image Σ is a sequence of points $p_0, p_1, \ldots, p_n$ such that $p_i \in \Sigma$ and p_i and p_{i-1} are adjacent for all $1 \leq i \leq n$. Given non-negative integers ϵ and δ, we say that two distinct points p, $q \in \Sigma$ are (ϵ, δ)-*connected* if there exists a path $p = p_0, p_1, \ldots, p_n = q$, such that the maximal variation of the gray values of the points on the path is less than or equal to ϵ, and the maximal variation of the gray values of any two adjacent points along the path is less than or equal to δ. Such a path will be called an (ϵ, δ)-*connected path* between p and q. A subset of Σ is called an (ϵ, δ)-*connected set* if each pair of points of the subset is (ϵ, δ)-connected. The definition of (ϵ, δ)-connectedness gives us a convenient tool to study the variation of gray values in an image. By varying the parameters ϵ and δ, we may investigate the diverse distribution of gray values. Given a point $p \in \Sigma$, any maximal (ϵ, δ)-connected set containing p is called a *related* (ϵ, δ)-*connected component* (in short, (ϵ, δ)-*RCC*) of p, and is denoted by C_p^Σ; the point p is called the *seed point* of its (ϵ, δ)-RCCs. (By a *maximal* (ϵ, δ)-connected set, we mean an (ϵ, δ)-connected set S such that there exists no other (ϵ, δ)-connected set which contains S properly, *i.e.*, for any point $p' \notin C_p^\Sigma$, there is at least one point $q \in C_p^\Sigma$, such that p', q are not (ϵ, δ)-connected.) For a given image, an (ϵ, δ)-RCC obtained by the algorithm discussed in [5] is called an (ϵ, δ)-component of the given image. The *spectrum* of the gray values of a set of pixels C is defined to be $max_C - min_C$, where min_C, max_C denote the minimum and the maximum of the gray values in C respectively.

3. Image Understanding by Connected Components

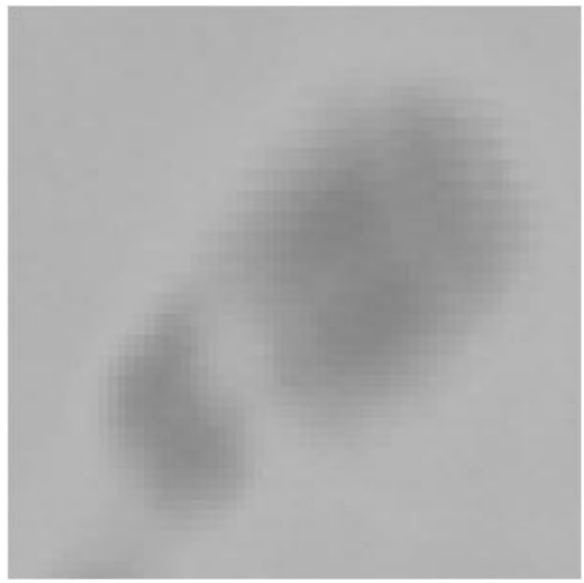
Any digitized image has certain inherent properties which are somehow shown as a scene that may not always represent a significant meaning to humans. The concept of the (ϵ, δ)-components is one of the inherent properties of an image. The algorithm described in [5] provides us a method to find the (ϵ, δ)-components for a given image and the values of the parameters.

Histogram has been used as a primary tool to analyze the structure of images in the pixel level understanding. In the conventional rectangular grid system, any pixel has only two distinguishable properties with respect to other pixels: its gray value and its location. From the gray-value-histogram of pixels, we could get only distribution of the gray values of pixels in an image and the location of each individual pixel would not help us significantly to understand the structure of an image. However, a connected component is often composed of a group of pixels. It posses much richer distinguishable properties, such as size, shape, location, maximal gray value, minimal gray value, with respect to other connected components than a single pixel does. So, we may expect that the various property histograms of (ϵ, δ)-components could provide us much more information about the structure of an image than the gray-value histogram of pixels. Besides the individual properties, the (ϵ, δ)-components of an image with different values of ϵ and δ have also statistical properties. Thus, we may study the structure of the image through the individual properties and the statistical properties of its (ϵ, δ)-components. At the component understanding level, if we concentrate on the individual properties of each com-

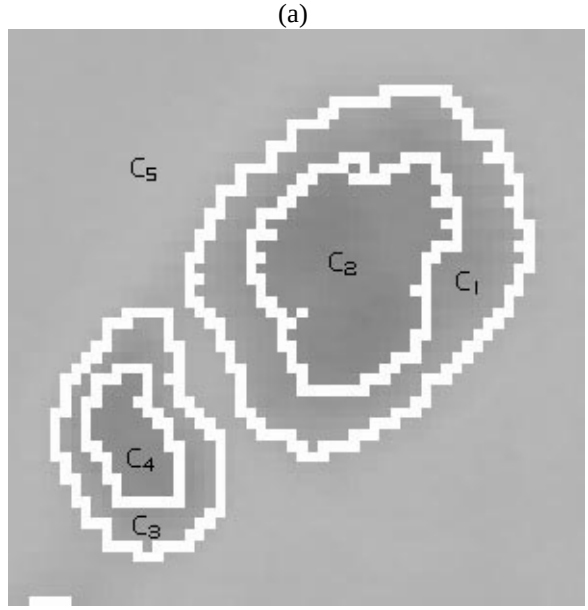
ponent, we understand the image in a *micro-view*. If we consider also the statistical properties of the components, we understand the image in a *macro-view*.

3.1. Image understanding in a micro-view

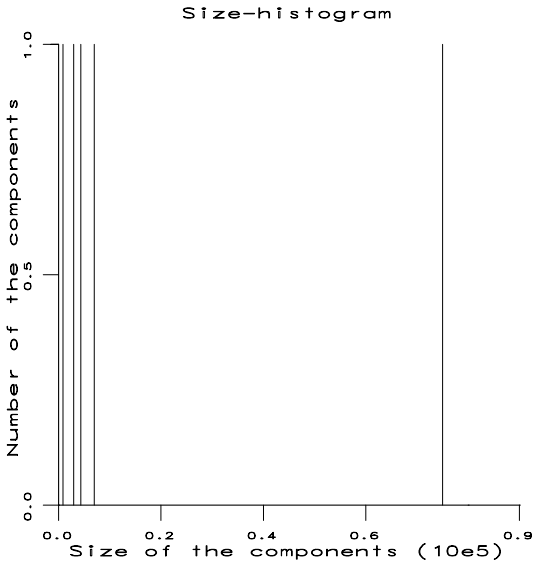
For a given ϵ and δ, we know that the corresponding (ϵ, δ)-components of an image have some common properties. From the definition of the (ϵ, δ)-component [5], it follows that the spectrum of the gray values of each component is less than or equal to ϵ. Although, within an (ϵ, δ)-component, the difference of the gray values of any two neighboring pixels may not be always less than or equal to δ, and it can be shown [5] that there is always a path between the two pixels such that the difference of the gray values between any two neighbor pixels on the path is less than or equal to δ. Since all the pixels within an (ϵ, δ)-component are known, we can easily find the size and maximal, minimal, or average gray values of each (ϵ, δ)-component. Although the exact size of a component may not be important (since it depends on the scale of the image), the relative size of the component to that of the other components is very helpful to understand the image in many cases. Furthermore, since we can trace down the boundary of each (ϵ, δ)-component, so, we are able to describe the shape of each (ϵ, δ)-component and locate its position. Thereafter, we are able to calculate other desirable geometric features of each component. If we can describe the shape of a component efficiently, it will help us to understand the content of the image. For example, we may check if an (ϵ, δ)-component is of some simple geometrical shape, such as a rectangle, a circle or a triangle. However, so far, there is no convenient formula to represent an arbitrary shape. So, in most practical problems we have to use the entire (or somehow encoded) boundary to describe the shape of a component. Figure 1 shows a simplified example. Figure 1-(a) is an original gray image. Figure 1-(b) contains its five (ϵ, δ)-components when $\epsilon = 20$ and $\delta = 10$. The component C_1 has the size 6947 pixels and the average gray value 162. The component C_2 has the size 4,335 pixels and the average gray value 138. Similarly, the components C_3, C_4 and C_5 have the sizes 2947, 851, 75020 and the average gray value 160, 142 and 181 respectively. From these information we understand that the image is composed of five connected components in each of which the variation of the gray values of pixels is no more than 20 and the difference of the gray value of a neighboring pixel is no more than 10. Also, we understand that C_1, C_2 are two smaller and darker components, C_5 is a bigger and brighter component. Therefore, with the positions of the connected components, we could have a good understanding of the structure of the image.



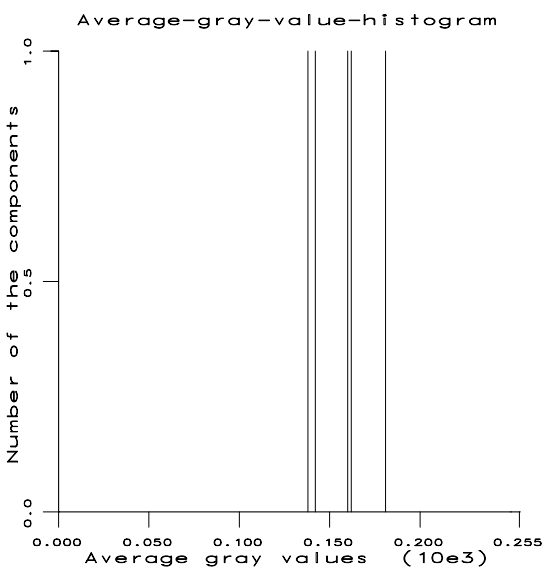
(a)



(b)



(c)



(d)

Figure 1: *(a) A gray image. (b) The* (ϵ, δ)*-components with* $\epsilon = 20$, $\delta = 10$. *(c) The corresponding size-histogram of (a). (d) The corresponding average-gray-value-histogram of (a).*

3.2. Image understanding in a macro-view

To understand an image in a macro-view, we concentrate on the statistical properties of its (ϵ, δ)-components. The most obvious statistic is the number of (ϵ, δ)-components. We now consider various histograms of the (ϵ, δ)-components – instead of the pixels – to analyze the distributions of the different properties of the (ϵ, δ)-components in an image. We shall refer to such a histogram a *property histogram* of (ϵ, δ)-components. Corresponding to each property of the (ϵ, δ)-components, we can create a property histogram of the (ϵ, δ)-components, where the X-axis represents the values of the property and the Y-axis represents the number of the (ϵ, δ)-components. Thus, a point (x, y) in a property histogram of (ϵ, δ)-components represents the information that there are y number of (ϵ, δ)-components in the image whose corresponding property has the value x. For example, we may create the size-histogram, max-gray-value-histogram, average-gray-value-histogram of (ϵ, δ)-components. Figure 1-(c) shows the size-histogram of the (ϵ, δ)-components of Figure 1-(b); Figure 1-(d) shows the average-gray-value-histogram of the (ϵ, δ)components of Figure 1-(b). If the shapes of the components are limited only to those of simple geometrical shapes (such as a rectangle, a disk), we may easily create a shape-histogram of components also. Where, the values along the X-axis may be a code of the different shape names. From a shape-histogram, we could easily find the number of rectangles, triangles in the image (This would be specially helpful to analyze, e. g., engineering drawings). From the various property histograms of (ϵ, δ)-components of an image, we could analyze the distributions of the (ϵ, δ)-components for corresponding properties. Thus, we could understand the image not only at the pixel-level but also at a higher level – the (ϵ, δ)-*component level*.

Note that the (ϵ, δ)-components of an image depend on the values of ϵ and δ. Changing the value of ϵ or δ, a property histogram of (ϵ, δ)-components may also be changed correspondingly. Comparing the different property histograms of an image for the different values of ϵ and δ, we may find the variations of the (ϵ, δ)-components with the values of ϵ and δ, and it could help us further to understand the structure of an image. Let Σ be an image, max_Σ and min_Σ be the maximal and the minimal gray value of the image respectively, then, the spectrum of the image $spectrum_\Sigma = max_\Sigma - min_\Sigma$. When $spectrum_\Sigma \leq \delta \leq \epsilon$, the entire image would be one (ϵ, δ)-component [5] – such an (ϵ, δ)-component does not help us to understand the image. So, we need to consider here only the situation when $0 \leq \delta \leq \epsilon \leq spectrum_\Sigma$ [5]. If a large amount of change in the ϵ-, δ-values causes a small change in a property histogram, the (ϵ, δ)-components are comparatively stable on the property within the corresponding ranges of the ϵ-, δ-values. If a small change in the ϵ-, δ-values causes

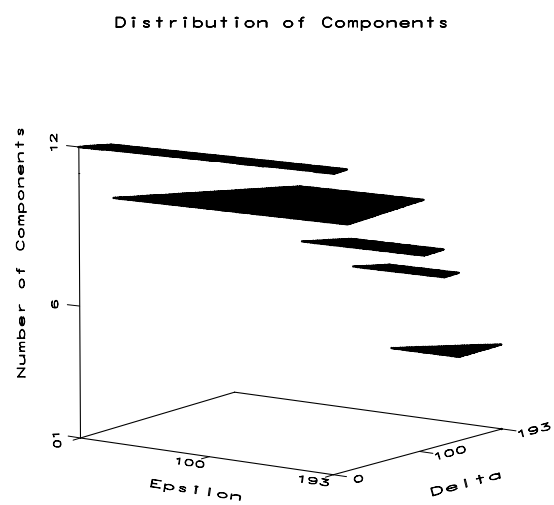
a big change in a property histogram, it implies that these ϵ-, δ-values change the (ϵ, δ)-components significantly by means of the corresponding property, and our experimental results show that these values are often good guides for segmentation.

Let $CN_{\epsilon,\delta}$ represent the number of the (ϵ, δ)-components of an image when the values of the parameters are ϵ and δ respectively, and
$max_{CN} = max_{0 \leq \delta \leq \epsilon < spectrum_\Sigma}\{CN_{\epsilon,\delta}\}$
$min_{CN} = min_{0 \leq \delta \leq \epsilon < spectrum_\Sigma}\{CN_{\epsilon,\delta} \mid CN_{\epsilon,\delta} > 1\}$
If max_{CN} is small for an image, we understand that the sizes of many homogeneous patches of the image are big. Otherwise, if max_{CN} is big for an image, then, the sizes of the most homogeneous patches of the image are small. For example, from a given image whose spectrum is 193, we obtain its $CN_{\epsilon,\delta}$ with different values of ϵ and δ as shown in Figure 2-(b). Since $max_{CN} = 12$ is small, we know that the image must be composed of several big homogeneous patches, even before we actually see this image. Figure 2-(a) shows this original image which matches with our expectation. Figure 3-(b) shows $CN_{\epsilon,\delta}$ for another image whose spectrum is 71. In this case, $max_{CN} = 168$ is big, so, the image contains many small homogeneous patches. Figure 3-(a) is the original image. From the distribution of the (ϵ, δ)-components with the different values of ϵ and δ shown in Figure 2-(b) or Figure 3-(b), we see that the number of (ϵ, δ)-components changes a lot in some ranges of the values of ϵ and δ, but a few in some other ranges of the values of ϵ and δ. Intuitively, we should pay more attention to those values of ϵ and δ which cause a big change of the number of (ϵ, δ)-components. Comparing Figure 2-(b) with Figure 3-(b), we know that the variation of gray values of pixels in Figure 2-(a) is more abrupt than that in Figure 3-(a). In a similar manner, we may analyze other properties of the image from its corresponding property histogram of the (ϵ, δ)-components.

4. Conclusion

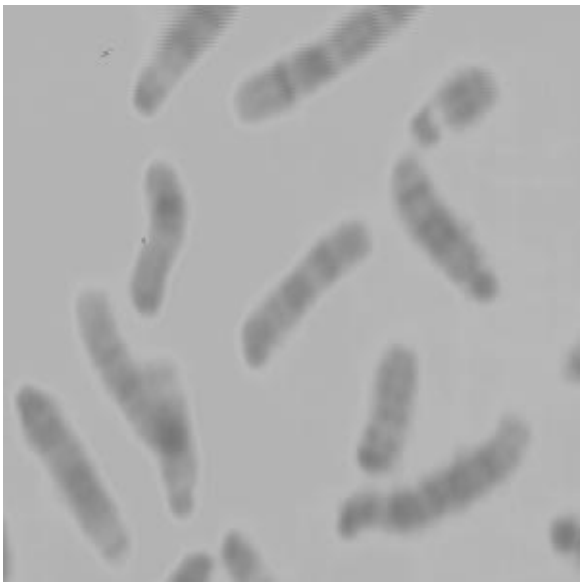
In this paper, using various component histograms, we have discussed a method to understand the structure of an image at an intermediate level through the (ϵ, δ)-components, the inter-relationships of the (ϵ, δ)-components, and the relationships of the (ϵ, δ)-components with different values of ϵ and δ. Combining certain pre-knowledge and information obtained from various (ϵ, δ)-component histograms, we could find the appropriate values of ϵ and δ, such that the objects in an image could be represented by its corresponding (ϵ, δ)-components reasonably well. Object extraction and segmentation may then be done by locating these (ϵ, δ)-components. Thus, using (ϵ, δ)-components as a tool, we may understand an image beginning from the intermediate level to the higher level.

(a)

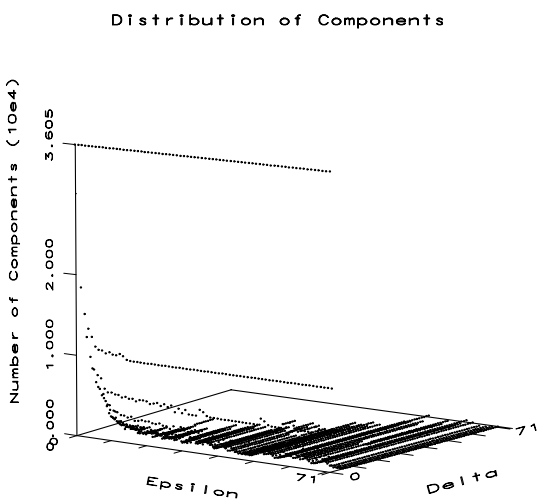


(b)

Figure 2: *(a) A gray image. (b) The $CN_{\epsilon,\delta}$s of the image in (a).*



(a)



(b)

Figure 3: *(a) A chromosome image. (b) The distribution of the* (ϵ, δ)*-components of the image in (a).*

Support of a 1998 Summer Faculty Fellowship by Southwest Missouri State University is gratefully acknowledged.

5. References

1. B. K. P. Horn. *Robot Vision*. The MIT Press, Cambridge, MA, 1986.
2. A. Rosenfeld and A. C. Kak. *Digital Picture Processing*. Academic Pr., New York, 1982.
3. A. Rosenfeld. On Connectivity Properties of Grayscale Pictures. *Pattern Recognition*, **16**:47–50, 1983.
4. F. Matus and J. Flusser. Image Representations via a Finite Radon Transform. *IEEE Trans. Patt. Anal. Mach. Intell.*, **15**:996–1006, 1993.
5. Y. Wang and P. Bhattacharya. On Parameter–Dependent Connected Components of Gray Images. *Pattern Recognition*, **29**:1359–1368, 1996.
6. F. Deravi and S. K. Pal. Grey Level Thresholding Using Second-order Statistics. *Pattern Recognition Letters*, **1**:417–422, 1983.
7. J. C. Olivo. Automatic Threshold Selection Using the Wavelet Transform. *Computer Vision, Graphics, and Image Processing*, **56**:205–218, 1994.
8. N. R. Pal and S. K. Pal. Entropic Thresholding. *Signal Processing*, **16**:97–108, 1989.
9. S. C. Sahasrabudhe and K. S. Das Gupta. A Valley–Seeking Threshold Selection Technique. *Computer Vision and Image Processing*, **93**:55–65, 1993.
10. P. K. Sahoo, Soltani, A. K. C. Wong, and Y. C. Chen. A Survey of Thresholding Techniques. *Computer Vision, Graphics, and Image Processing*, **41**:233–260, 1988.
11. Y. Wang and P. Bhattacharya. A Theory of Parameter–Dependent Connected Components of Gray Images and Segmentation. In *Proc. IEEE Internat. Conf. on Image Processing*, **3**:69–72, Washington DC., Oct. 1995.
12. Y. Wang and P. Bhattacharya. Image Analysis and Segmentation Using Gray Connected Components. In *Proc. IEEE Internat. Conf. on Systems, Man and Cybernetics*, **1**:444–449, Beijing, China, Oct. 1996.
13. T. Y. Kong and A. Rosenfeld. Digital Topology: Introduction and Survey. *Computer Vision, Graphics, and Image Process.*, **48**:357–393, 1989.