Chapter 4

Analysis of Morphological Document Image Segmentation
ANALYSIS OF MORPHOLOGICAL DOCUMENT IMAGE SEGMENTATION

Mathematical morphology is closely related to integral geometry. It quantifies many aspects of the geometrical structure of images in a way that agrees with human intuition and perception. Mathematical morphology has been widely used for many applications of image processing and analysis [1-5]. Morphological processing can be employed for many purposes including pre-processing, edge detection, segmentation, and object recognition. Morphological expressions are defined as a combination of image operations, the simplest of which is the operation of erosion and dilation. The morphological approach is based on the analysis of an image in terms of some predetermined geometric shape templates known as structuring elements. The manner in which the structuring elements can be embedded into the original shape by using a specific sequence of operators leads eventually to shape classification or discrimination. The morphological operators are the filters that encode the original shape for providing features needed for shape discrimination. However the structuring elements design is difficult due to its computational intractability.

Part of this chapter is published in and/ or communicated to
4.1 Introduction Document Image Segmentation

One of the most important operations in Computer Vision and Image Processing is image segmentation. Since last 30 years, hundreds of segmentation algorithms are proposed; still it is a highly important area of research in image processing and computer Vision. The aim of image segmentation is the domain-independent partition of the image into a set of regions, which are visually distinct and uniform with respect to some property, such as gray level, texture or color. The problem of segmentation has been, an important research field and many segmentation methods have been proposed in the image segmentation literature. Image segmentation techniques may be classified into number of groups depending on the approach of the algorithm. These include feature thresholding, contour based techniques, region based techniques, clustering, template matching, edge or boundary based, Morphological Watershed, hybrid etc [1-5]. Each of these approaches has its own merits and demerits in terms of applicability, suitability, performance, computational cost and so on.

In particular, many of the existing techniques for image description and recognition, image visualization and object based image compression highly depend on the segmentation results. The general segmentation problem involves the partitioning of a given image into a number of homogeneous segments (Spatially connected groups of pixels), such that the union of any two neighboring segments yields a heterogeneous segment. Alternatively, segmentation can be considered as a pixel labeling process in the sense that all pixels that belong to the same homogeneous region are assigned the same label. There are several ways to define homogeneity of a region based on the particular objective of the segmentation process. However, independently of the homogeneity criteria, the noise corrupting almost all acquired images is likely to prohibit the generation of error-free image partitions.

Image segmentation is a fundamental part of the 'low level' aspects of computer vision and has many practical applications such as in medical imaging, industrial automation and satellite imagery. Traditional methods for image segmentation have approached the problem either from localization in class space using region information, or from localization in position, using edge or boundary information. Segmentation algorithms for monochrome images generally are based on one of two basic properties of gray-level values: discontinuity and similarity.

In the first category, the approach is to partition an image based on abrupt changes in gray level. The principal areas of interest within this category are detection
of isolated points and detection of lines and edges in an image. The principal approaches in the first category are based on edge detection, and boundary detection. Basically, the idea underlying most edge-detection techniques is the computation of a local derivative operator. The first derivative of the gray level profile is positive at the leading edge of a transition, negative at the trailing edge, and zero in areas of constant gray level. Hence the magnitude of the first derivative can be used to detect the presence of an edge in an image.

For Image segmentation many segmentation methods have been proposed. A typical Digital Image Segmentation system is as shown in fig. 4.1. An input image is submitted to preprocessing phase, where image enhancement operation is carried out by frequency and spatial domain methods. The result of preprocessing phase is submitted to segmentation phase, where actual segmentation operation is made and output image is in segmented form. The actual segmentation is stopped when the object of interest is isolated.

![Fig. 4.1 Digital Image Segmentation System](image)

In image analysis [6], segmentation is the partition of a digital image into multiple regions (sets of pixels), according to some criterion. The goal of segmentation is typically to locate certain objects of interest, which may be depicted in the image. Segmentation could therefore be seen as a computer vision problem. The image can be acquired from some problem domain by using various kind of image acquisition devices passed to preprocessing unit. The key function of preprocessing and Image enhancement phase is to improve the image in the way that increases the chances of success for the other process.

Image segmentation is often used as an initial transformation for general image analysis and understanding. Therefore image segmentation has a large number of
applications in varying fields. The aim of an automatic image segmentation system is to mimic the human visual system in order to provide a meaningful image sub division. There is currently substantial and growing interest in the field of document image understanding where several research groups are developing and designing systems to automatically process and extract relevant information from document such as engineering drawings, maps, magazines, newspapers, forms and mail envelopes. A thresholding technique must be able to segment a digitized image into different objects with similar properties. There are general technique for image segmentation and limit the scope of the problem to document images. The new approaches goes beyond the segmenting only one bright object from an image to an approach that recursively segment the brightest object from image with recursion, living the darkest object in the given digitized image.

**Segmentation Techniques**

In segmentation phase the image (such as multi-resolution, multi-spectral) is divided into constituent parts as shown in fig. 4.2.

![Segmentation Diagram](image)

**Fig. 4.2 Typical Image segmentation**

The image segmentation techniques can be broadly grouped into the following categories:

i) Thresholding or Histogram based techniques

ii) Edge or Boundary based method

iii) Region based method

iv) Morphology based

v) Hybrid Techniques

Following fig. 4.3 shows the classification of image segmentation techniques.
4.2 Morphology Based Techniques

Morphology is biological term refers to study of form and structure, in imaging, the term is not used so generically [7]. Morphological operators often take a binary image and a structuring element as input and combine them using a set operator (intersection, union, inclusion, complement) [1,3]. They process objects in the input image based on characteristics of its shape, which are encoded in the structuring element. The mathematical details are explained in Mathematical morphology. Usually, the structuring element is sized $3 \times 3$ and has its origin at the center pixel. It is shifted over the image and at each pixel of the image its elements are compared with the set of the underlying pixels. If the two sets of elements match the condition defined by the set operator (e.g. if set of pixels in the structuring element is a subset of the underlying image pixels), the pixel underneath the origin of the structuring element is set to a pre-defined value (0 or 1 for binary images). A morphological operator is therefore defined by its structuring element and the applied set operator.

For the basic morphological operators the structuring element contains only foreground pixels (i.e. ones) and 'don't care's'. These operators, which are all a combination of erosion and dilation, are often used to select or suppress features of a certain shape, e.g. removing noise from images or selecting objects with a particular direction. The more sophisticated operators take zeros as well as ones and 'don't care's' in the structuring element. The most general operator is the hit and miss; in fact, all the other morphological operators can be deduced from it. Its variations are often used to simplify the representation of objects in a (binary) image while preserving their
structure, e.g. producing a skeleton of an object using skeletonization and tidying the result up using thinning.

Mathematical Morphology refers to a branch of nonlinear image processing and analysis that concentrates on the geometric structure within an image, it is mathematical in the sense that the analysis is based on set theory, topology, lattice, random functions, etc. [7,8]. As well as mathematical morphology is considered as a powerful tool to extract information from images [8]. Where as erosion and dilation are considered the primary morphological operations and the operations of opening and closing are secondary operations and are implemented using erosion and dilation operations [1]. Morphological operators can also be applied to gray scale images to reduce noise or to brighten the image. However, for many applications, other methods like a more general spatial filter produces better results. In this research work, Mathematical morphology is used for segmentation of text document image because morphological operations have following salient features [9]:

1. Morphological operations provide for the systematic alteration of the geometric content of an image while maintaining the stability of the important geometric characteristics.
2. There exists a well-developed morphological algebra that can be employed for representation and optimization.
3. It is possible to express digital algorithms in terms of a very small class of primitive morphological operations.
4. There exist rigorous representations theorems by means of which one can obtain the expression of morphological filters in terms of the primitive morphological operations.

In this research work mathematical morphology based document image segmentation technique is proposed which provides promising results.

Morphology means structure. The field of mathematical morphology contributes a wide range of operators to image processing, all based around a few simple mathematical concepts from set theory [6]. The operators are particularly useful for the analysis of images and common usages include edge detection, noise removal, image enhancement and image segmentation. The two most basic operations in mathematical morphology are erosion and dilation [1,3,10]. Both of these operators take two pieces of data as input: an image to be eroded or dilated, and a structuring element (also known as a kernel). The two pieces of input data are each treated as
representing sets of coordinates in a way that is slightly different for binary and grayscale images. For a binary image, white pixels are normally taken to represent foreground regions, while black pixels denote background. (It should be noted that in some implementations this convention is reversed, and so it is very important to set up input images with the correct polarity for the implementation being used). Then the set of coordinates corresponding to that image is simply the set of two-dimensional Euclidean coordinates of all the foreground pixels in the image, with an origin normally taken in one of the corners so that all coordinates have positive elements. For a grayscale image, the intensity value is taken to represent height above a base plane, so that the grayscale image represents a surface in three-dimensional Euclidean space. Then the set of coordinates associated with this image surface is simply the set of three-dimensional Euclidean coordinates of all the points within this surface and also all points below the surface, down to the base plane.

The structuring element is already just a set of point coordinates (although it is often represented as a binary image). It differs from the input image coordinate set in that it is normally much smaller, and its coordinate origin is often not in a corner, so that some coordinate elements will have negative values. In many implementations of morphological operators, the structuring element is assumed to be a particular shape (e.g. a $3 \times 3$ square) and so is hardwired into the algorithm. Binary morphology can be seen as a special case of gray level morphology in which the input image has only two gray levels at values 0 and 1. Erosion and dilation work by translating the structuring element to various points in the input image, and examining the intersection between the translated kernel coordinates and the input image coordinates. For instance, in the case of erosion, the output coordinate set consists of just those points to which the origin of the structuring element can be translated, while the element still remains entirely 'within' the input image.

Virtually all other mathematical morphology operators can be defined in terms of combinations of erosion and dilation along with set operators such as intersection and union. Some of the more important are opening, closing and skeletonization.

### 4.3 Morphological Operations

Morphology technique is used to operate on images via different operations [1,3,11,12]. Some of them are described in this section.

#### 4.3.1 Erosion and Dilation
Erosion and Dilation constitute the basis for more complex morphological operators and can be defined as follows. Let \( A: z^2 \rightarrow z \) be a gray-scale image and \( B: z^2 \rightarrow z \) a structuring element. The gray scale erosion of \( A \) by \( B \) denoted by \((A \Theta B)\), is defined as [1]

\[(A \Theta B)(x, y) = \min\{A(x + s, y + t) - B(s, t)\} \]

------- (1)

Where \((x + s, y + t) \in D_A\), \((s, t) \in D_B\) and \(D\) represents the discrete domain of the images.

The gray scale dilation of \( A \) by \( B \), denoted by \((A \oplus B)\), is defined as

\[(A \oplus B)(x, y) = \max\{A(x - s, y - t) + B(s, t)\} \]

------- (2)

Where \((x - s, y - t) \in D_A\), \((s, t) \in D_B\).

4.3.2 Closing and Opening

By iteratively applying erosion and dilation, one can eliminate image details, smaller than the structuring element, without affecting its global geometric features. Visually, closing operation smoothes the contours by filling in narrow gulfs and eliminating small holes. The closing operation is simply a dilation operation followed by an erosion operation. The gray scale closing of \( A \) by \( B \), denoted by \((A \bullet B)\), is defined as,

\[(A \bullet B) = (A \oplus B) \Theta B \]

----- (3)

On the other hand, opening smoothes contours by suppressing small islands. The opening operation is simply an erosion operation followed by a dilation operation. The gray scale opening of \( A \) by \( B \), denoted by \((A \circ B)\), is defined as,

\[(A \circ B) = (A \Theta B) \oplus B \]

----- (4)

4.3.3 Top-hat and Bottom-hat Transform

When we subtract the opened image from the original image is known as Top-hat Transform \((T_h)\) is defined as,

\[T_h = I - (I \circ B) \]

----- (5)

When we subtract the closed image from the original image \((I)\) is known as Bottom-hat Transform \((B_h)\) is defined as,
\[ B_h = I - (I \bullet B) \] ------ (6)

Where, \( I \) is the image, \( B \) is the Structuring element.

4.4. Structuring Element

The theoretical foundation of binary mathematical morphology is set theory. In binary images, those points in the set are called the ‘foreground’ and those in the complement set are called the ‘background’. Besides dealing with the usual set-theoretic operations of union and intersection, morphology depends heavily on the translation operation. For convenience, ‘\( \cup \)’ denotes the set-union, ‘\( \cap \)’ denotes set-intersection and ‘+’ inside the set notation refers to vector addition in the following equations. We need two general definitions that are used extensively to extend morphological operations [1].

Structuring Element: Before continuing, we describe what a structuring element is in morphological operations. A structuring element is a small image that is overlapped on input image to compute a certain definition. The basic operations of binary and also gray-scale images depend on what structuring elements are used. In this section, only those for binary morphology are considered. Figure 4.4 contains some examples that are used commonly for binary images.

![Fig. 4.4 Examples of 3x3 Structuring Element](image)

In these cases, all origins of structuring elements are located on the center. Depending on the shape of structuring element, the origin could be on a different location. The pixels marked with ‘1’ are the points that should be considered during any binary morphological operations.
Structuring element is a small grid representing pixels, which are either set (1), not set (0), or “don’t care”. It is applied to images to change the structure of the image content. The structuring element (SE) is generally of square dimension of size 3x3, 5x5 and sometimes greater depending upon the application. For the 3x3 structuring element cases, nine logical values are defined and labeled as in fig. 4.5.

\[
\begin{array}{ccc}
X3 & X2 & X1 \\
X4 & X & X0 \\
X5 & X6 & X7 \\
\end{array}
\]

Fig. 4.5 A 3X3 structuring element

Every pixel in the input image is evaluated with its eight neighborhoods to produce a resulting output pixel value. Structuring elements are also known as morphological mask, which play an important role in morphological operation like opening and closing. The structuring element consists of a pattern specified as the coordinates of a number of discrete points relative to some origin. Normally Cartesian coordinates are used and so a convenient way of representing the element is as a small image on a rectangular grid. An important point to note is that although a rectangular grid is used to represent the structuring element, not every point in that grid is part of the structuring element in general. Hence the elements shown in Fig. 4.4 contain some blanks. In many cases, these blanks are represented as zeros.

When a morphological operation is carried out, the origin of the structuring element is typically translated to each pixel position in the image in turn, and then the points within the translated structuring element are compared with the underlying image pixel values. The details of this comparison and the effect of the outcome depend on which morphological operator is being used. If we use both opening and closing operations in a program then structuring element or morphological mask remains same for the both operation [13].

4.4.1 Basic Structuring Elements

Structuring element (SE) type is specified by its shape. Depending on shape, structuring element can take additional parameters. There are two types of structuring element [1,3]

i) **Flat** Structuring Element.

ii) **Non-flat** Structuring Element.
Flat structuring element have maximum two parameters, this vary as per requirement, whereas, non-flat structuring element have one additional parameter i.e. height, that means non-flat structuring element may be used for 3D images. For example:

![Diagram of a flat structuring element](image)

**Fig. 4.6** Disk shape flat structuring element with radius $R=3$

**Flat Structuring Elements are,**

i) Arbitrary  
ii) Pair  
iii) Diamond  
iv) Disk  
v) Periodic line  
vi) Rectangle  
vii) Line  
viii) Square  
ix) Octagon

**Non-flat Structuring Elements are,**

i) Arbitrary  
ii) Ball

**4.4.1.1 Flat Structuring Elements**

i) Arbitrary  

Arbitrary creates a flat structuring element where NHOOD specifies the neighborhood. NHOOD is a matrix containing 1's and 0's; the location of the 1's defines the neighborhood for the morphological operation. The center (or origin) of NHOOD is its center element, given by floor

$$((\text{size (NHOOD)} + 1)/2).$$
ii) Diamond

Diamond creates a flat, diamond-shaped structuring element, where R specifies the distance from the structuring element origin to the points of the diamond. R must be a nonnegative integer scalar.

iii) Disk

Disk creates a flat, disk-shaped structuring element, where R specifies the radius. R must be a nonnegative integer. N must be 0, 4, 6, or 8. When N is greater than 0, the disk-shaped structuring element is approximated by a sequence of N periodic-line structuring elements. When N equals 0, no approximation is used, and the structuring element members consist of all pixels whose centers are no greater than R away from the origin. If N is not specified, the default value is 4.
iv) Line

Line creates a flat, linear structuring element, where LEN specifies the length, and DEG specifies the angle (in degrees) of the line, as measured in a counterclockwise direction from the horizontal axis. LEN is approximately the distance between the centers of the structuring element members at opposite ends of the line.

![Line Structuring Element Diagram]

v) Octagon

Octagon creates a flat, octagonal structuring element, where R specifies the distance from the structuring element origin to the sides of the octagon, as measured along the horizontal and vertical axes. R must be a nonnegative multiple of 3.
vi) Pair

Pair creates a flat structuring element containing two members. One member is located at the origin. The second member's location is specified by the vector \(\text{OFFSET}\). \(\text{OFFSET}\) must be a two-element vector of integers.

![Fig. 4.11 Octagon shaped Structuring element](image)

![Fig. 4.12 Pair shaped Structuring element](image)

vii) Periodic line

Periodic line creates a flat structuring element containing \(2\times P+1\) members. \(V\) is a two-element vector containing integer-valued row and column offsets. One structuring element member is located at the origin. The other members are located at \(1\times V, -1\times V, 2\times V, -2\times V, ..., P\times V, -P\times V\).

![Fig. 4.13 Periodic Line shaped Structuring element](image)

viii) Rectangle

Rectangle creates a flat, rectangle-shaped structuring element, where \(MN\) specifies the size. \(MN\) must be a two-element vector of nonnegative integers. The first element of \(MN\) is the number of rows in the structuring element neighborhood; the second element is the number of columns.
ix) Square

Square creates a square structuring element whose width is $W$ pixels. $W$ must be a nonnegative integer scalar.

ii) Ball

Ball creates a non-flat, ball-shaped structuring element (actually an ellipsoid) whose radius in the X-Y plane is $R$ and whose height is $H$. Note that $R$ must be a nonnegative integer, $H$ must be a real scalar, and $N$ must be an even nonnegative integer. When $N$ is greater than 0, the ball-shaped structuring element is approximated by a sequence of $N$ non-flat, line-shaped structuring elements. When $N$ equals 0, no approximation is used, and the structuring element members consist of all pixels whose centers are no greater than $R$ away from the origin. The corresponding height values are determined from the formula of the ellipsoid specified by $R$ and $H$. If $N$ is not specified, the default value is 8.
In this work, all types of flat and non-flat structuring elements are applied on document image to find out exact shape and size of structuring element suitable for document image segmentation.

4.4.2 Reflection of Structuring Element about origin

The structuring element consists of a pattern specified as the coordinates of a number of discrete points relative to some origin. Normally Cartesian coordinates are used and so a convenient way of representing the element is as a small image on a rectangular grid. Figure 4.16 shows a structuring element of various sizes. In each case a ring around that point marks the origin. The origin does not have to be in the center of the structuring element, but often it is. As it is evident from the figure, structuring elements that fit into a 3×3 grid with its origin at the center is the most commonly seen type. An important point to note is that although a rectangular grid is used to represent the structuring element, not every point in that grid is part of the structuring element in general. Hence the elements shown in figure 4.16 contain some blanks. In many texts, these blanks are represented as zeros. When a morphological operation is carried out, the origin of the structuring element is typically translated to each pixel position in the image in turn, and then the points within the translated structuring element are compared with the underlying image pixel values. The details of this comparison and the effect of the outcome depend on the morphological operator is used.

Fig. 4.16 The examples of Morphological Mask/Structuring Element [14].

Morphological operations run much faster when the structuring element uses approximations (N > 0) than in case of (N = 0). However, structuring elements that do not use approximations (N = 0) are not suitable for computing granulometries. For all shapes but for 'arbitrary', structuring elements are constructed using a family of techniques collectively known as structuring element decomposition. The principle is that dilation by some large structuring elements can be computed faster by dilation with a sequence of smaller structuring elements. For example, dilation by an 11-by-11
square structuring element can be accomplished by dilating first with a 1-by-11 structuring element and then with an 11-by-1 structuring element. Structuring element decompositions used for the 'disk' and 'ball' shapes are approximations; all other decompositions are exact.

4.5 Properties of Morphology Operators

We briefly describe some properties of morphology operators as follows:

**Translation**

Let $A$ and $B$ be subsets of $\mathbb{Z}^2$. The *translation* of $A$ by $x$ is denoted $A_x$ and is defined as

$$A_x = \{v : v = a + x, \text{for } a \in A\}.$$

**Reflection**

The *reflection* of $B$, denoted $\hat{B}$, is defined as

$$\hat{B} = \{v : v = -b, \text{for } b \in B\}.$$

**Difference**

Set of points that belongs to $A$ but not to $B$

$$A - B = \{w | w \in A, w \notin B\}$$

$$= A \cap B^c$$

**Thinning**

Thins set $A$. The first two equations give the basic definition of thinning. The last two equations denote thinning by a sequence of structuring elements.

$$A \odot B = A - (A \star B)$$

$$= A \cap (A \star B)^c$$

$$A \odot \{B\} = ((...(A \odot B^1) \odot B^2)...) \odot B^n)$$

$$\{B\} = \{B^1, B^2, B^3,...B^n\}$$

**Thickening**

Thickens set $A$.

$$A \cdot B = A \cup (A \star B)$$

$$A \cdot \{B\} = ((...(A \cdot B^1) \cdot B^2)...) \cdot B^n)$$

**Pruning**
X₄ is the result of pruning set A. The number of times that the first equation is applied to obtain X₄ must be specified. Structuring elements V are used for the first two equations.

\[ X_1 = A \oslash \{B\} \]
\[ X_2 = \bigcup_{k=1}^{8} (X_1 \ast B^k) \]
\[ X_3 = (X_2 \oslash H) \cap A \]
\[ X_4 = X_1 \cup X_3 \]

4.6 Morphology for Binary Images

4.6.1 Binary Dilation

Binary Dilation With A and B as sets in \(Z^2\), the dilation of A by B (usually A is an image and B is the structuring element), denoted by \(A \oslash B\), is defined as

\[ A \oslash B = \{Z \in Z^2 | Z = a + b \text{ for some } a \in A \text{ and } b \in B\} \]

It can be shown that dilation is equivalent to a union of translation of the original image with respect to the structuring element:

\[ A \oslash B = \bigcup_{b \in B} (A)_b \]

Dilation is found by placing the center of the template over each of the foreground pixels of the original image and then taking the union of all the resulting copies of the structuring element, produced by using the translation. From Figure 4.17, it is clear how dilation modifies the original image with respect to the shape of the structuring element. Dilation generally has an effect of expanding an image; so consequently, small holes inside foreground can be filled. In another sense, dilation can be a morphological operation on a binary image defined as:
$A \oplus B = \{A | (\hat{B})_z \cap A \neq \Phi\}$

This equation is based on obtaining the reflection of B about its origin and shifting this reflection by $z$. The dilation of A by B is the set of all displacements, $z$, such that B and A overlap by at least one element. Based on this interpretation, the equation above may be written as

$A \oplus B = \{Z | [(\hat{B})_z \cap A] \subseteq A\}$

Figure 4.18 illustrates the dilation operation using a binary image. The original image is dilated with an 11x11 ‘disk’ type structuring element.

**Fig. 4.18 Binary Dilation Example**

### 4.6.2 Binary Erosion

Erosion of a binary image $A$ by structuring element $B$, denoted by $A \ominus B$, is defined as

$A \ominus B = \{Z | Z + b \in A, \forall b \in B\}$

Whereas dilation can be represented as a union of translates, erosion can be represented as an intersection of the negative translates. So, the given definition of erosion above can be redefined as

$A \ominus B = \bigcap_{b \in B} A_{-b}$

Where ‘$-b$’ is the scalar multiple of the vector $b$ by $-1$. 
The erosion of the original image by the structuring element can be described intuitively by template translation as seen in the dilation process. Erosion shrinks the original image and eliminates small enough peaks (Note: the terms ‘expand’ for dilation and ‘shrink’ for erosion refer to the effects on the foreground). Figures 4.19 and 4.20 clearly illustrate these effects. The original image is eroded with 7x7 disk-shape structuring element.

\[ A \circ B = (A \Theta B) \oplus B \]

From the definition, the original image A is first eroded and then dilated by the same structuring element B. In terms of set theory, this opening process can also be defined as

\[ A \circ B = \bigcup \{(B)_x | (B)_x \subseteq A\} \]

The whole procedure of opening can be interpreted as “rolling the structuring element about the inside boundary of the image” as illustrated in fig. 4.21.
The effects of the opening process on the original image are smoothing, reducing noise from quantization or the sensor and pruning extraneous structures. These effects result from the fact that the structuring element cannot fit into the regions. Therefore, it can be said that the result of the opening process heavily depends on the shape of structuring elements. Figure 4.22 presents an example of the opening process. The vertices of the triangle foreground have been cut out because the image is “opened” with ‘square’ type structuring element, whereas those of the square are preserved.

4.6.4 Binary closing

Closing of a binary image $A$ by a structuring element $B$, denoted by $A \ast B$, is defined as

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

In the closing operation, dilation and erosion are applied successively in that order. Note that this order is reversed for the opening process. In another aspect, the closing process on a binary image can be defined as:
Z is an element of AB if and only if \((B + y) \cap A \neq \emptyset\) for any translate \((B + y)\) containing \(Z\).

The closing operation can be described as in Figure 4.23 as “rolling the structuring element on the outer boundary of the image.”

![Illustration of Binary Closing Process](image)

Fig. 4.23 Illustration of Binary Closing Process [1]

The closing process has the effect of filling small holes in the original image, smoothing as the opening process does, and filling up the bay in the foreground. Sometimes, it is said that the closing has an effect of clustering each spatial point to be connected. Figure 4.24 shows how the closing operation works.

![Binary Closing Example](image)

Fig. 4.24 Binary Closing Example

4.7 Morphology for Gray-Scale Images

In this section, the binary morphology is extended to the gray-scale case. The key issue is to use the ‘Maximum’ and ‘Minimum’ functions to define gray-scale morphological operators. Using these concepts, gray-scale morphology can be easily extended from binary morphology with the same concept. The differences between binary and gray-scale morphology results from the definitions of dilation and erosion because other operations basically depend on these. Except for these definitions, gray-scale morphology is fairly similar to the binary case. Hence, in this section, definitions for gray-scale dilation and erosion as well as some of examples for gray-scale morphological operations are given. [1,3,5].
4.7.1 Gray-scale Dilation and Erosion

Before discussing basic gray-scale morphological operations, it should be noted that the structuring elements of the gray-scale morphological operations could have the same domains as those in binary morphology. However, a gray-scale structuring element has certain values ('b's) instead of having only position value '1' or '0' showing its domain. A grayscale image can be considered as a three-dimensional set where the first two elements are the x and y coordinates of a pixel and the third element is grayscale value. It can be also applied to the gray-scale structuring element. With this concept, gray-scale dilation can be defined as follows:

**Gray-scale Dilation**

Gray-scale dilation of \( f \) by \( b \), denoted by \( f \oplus b \), is defined as

\[
(f \oplus b)(s,t) = \max \{ f(s-x,t-y) + b(x,y) | (s-x,t-y) \in D_f ; (s,y) \in D_b \}
\]

Where \( D_f \) and \( D_b \) are the domains of \( f \) and \( b \), respectively. Similarly, gray-scale erosion can be defined as an extension of binary erosion.

**Gray-scale Erosion**

Gray-scale erosion, denoted by \( f \ominus b \), is defined as

\[
(f \ominus b)(s,t) = \min \{ f(s+x,t+y) - b(x,y) | (s+x,t+y) \in D_f ; (s,y) \in D_b \}
\]

Where \( D_f \) and \( D_b \) are the domains of each image or function.

Specific concepts and operation procedures are already explained in binary morphology. Gray-scale dilation and erosion are duals with respect to function completion and reflection. That is, the relation between these can be expressed as

\[
(f \ominus b)^c(s,t) = (f^c \oplus \hat{b})(s,t)
\]

Where \( f^c = -f(x,y) \) and \( \hat{b} = b(-x,-y) \)

The minimum operator will interrogate a neighborhood with a certain domain and select the smallest pixel value to become the output value. This has the effect of causing the bright areas of an image to shrink or erode. Similarly, grayscale dilation is performed by using the maximum operator to select the greatest value in a neighborhood.

4.7.2 Gray-scale Opening and Closing

Gray-scale opening and closing are defined below in a similar manner as the binary case. The only difference is, when the operations are carried out, these opening...
and closing operations use gray-scale dilation and erosion described in the previous section. As binary morphological operations do, gray-scale opening is anti-extensive and gray-scale closing is extensive. Both operations make an original image smooth along to the nature of minimum and maximum functions. Also, both operations have ‘increasing’, ‘idempotent’ properties.

4.8 Hit or Miss Transformation

This transform is used to look for particular patterns of foreground and background pixels. It is very simple object recognition technique. All other morphological operations can be derived from it.
Input: 1. Binary Image
   2. Structuring Element, containing 0s and 1s.
Figure 4.25 shows a Hit-and-miss Structuring Element containing 0’s, 1’s, don’t care’s and usually a “1” at the origin.

![Fig. 4.25 Structuring element for Hit and Miss Transform](image)

4.9 Applications of Gray Scale Morphology

Although there are some applications requiring basic gray-scale morphological operations, most applications of morphology are developed for binary images. A list of binary morphological applications is as mentioned below:

- Boundary extraction
- Region filling
- Extraction of connected components
- Convex Hull
- Thinning
- Thickening
- Skeletons
- Pruning

All these applications are carried out by applying a series of basic operations with different types of structuring elements. For gray-scale morphological case, it can be expanded to

- Smoothing
- Morphological gradient
- Top-hat, Bottom-hat Transformations
- Textual segmentation
- Granulometry

The work in this thesis focuses on how to use morphological structuring element and how to improve that for document image segmentation.

**Smoothing**

One of the best solution to achieve smoothing is to perform a morphological opening operation followed by the closing operation. The net result of these two operations is to remove or attenuate both bright and dark artifacts and noise.

**Morphological Gradient**

The removal of small dark and bright artifacts, dilation and erosion often is used to compute the morphological gradient of an image, denoted by \( g \); for linear filters the gradient filter yields a vector representation with a magnitude and direction.

\[
g = (f \oplus b) - (f - b)
\]

**Top-hat Transformation**

The morphological top-hat transform of an image, denoted \( h \), is defined as

\[
h = f - (f \circ b)
\]

Where \( f \) is the input image and \( b \) is the structuring element function. This transform, which owns its original name to the use of a cylindrical or parallelepiped structuring element function with a flat top is useful for enhancing detail in the presence of shading.

**Textural Segmentation**

The segmentation problem can be described as the task of partitioning an image into homogeneous regions. For textured images one of the main conceptual difficulties is the definition of a homogeneity measure in mathematical terms. The general
approach to unsupervised texture segmentation is based on four cascaded design decisions, concerning the questions of image representation, texture homogeneity, objective functions and optimization procedures. Texture segmentation is formulated as a pair wise data clustering problem with a sparse neighborhood structure. The pair wise dissimilarities of texture blocks are computed using a multi scale image representation, which are tuned to spatial frequencies at different scales and orientations [15,16]. Texture segmentation involves subdividing an image into differently textured regions. Mostly textural segmentation is based on filter bank model, where the filter is Gabor filter derived from Gabor Elementary function. The goal is to transform texture differences into detectable filter output discontinuities at texture boundaries, by locating these boundaries; one can segment the image into differently textured regions where distinct discontinuities occur, however, only if the Gabor filter parameters are suitably chosen. [17].

Granulometry

Granulometry is a field that deals with determining the size distribution of particles in an image. An opening operation with structuring element used of increasing size is performed on the original image. The difference between the original image and its opening is computed after each pass when a different structuring element is completed. At the end of the process, these differences are normalized and then used to construct a histogram of particle size distribution. This approach is based on the idea that opening operation of a particular size have the most effect on regions of the input image that contain particles of similar size. Thus, a measure of the relative number of such particles is obtained by computing the difference between the input and output images.

In this research work, morphological segmentation technique is proposed which works on the structure, shape and size of document images. In the next chapter we present our proposed segmentation model based on morphology. The morphological approach presented here works well in all situations like high contrast, low contrast, uneven background, heavily noisy background, etc as well as in this approach no need to take threshold value to separate foreground and background in pre or post processing. In reported work, mathematical morphology worked for pre & post processing only. But as mentioned, it works well but no one proposed such technique fully based on mathematical morphology.
References:


