

# Application of binary mathematical morphology to separate overlapped objects

Bhabatosh Chanda

*Nodal Centre for Knowledge Based Computing, Indian Statistical Institute, 203 B.T. Road, Calcutta 700 035, India*

## *Abstract*

Chanda, B., Application of binary mathematical morphology to separate overlapped objects, Pattern Recognition Letters 13 (1992) 639-645.

A simple algorithm has been proposed to separate the images of overlapped objects from binary image. The algorithm is designed using Mathematical Morphological operations and, hence, can be implemented on parallel machines.

*Keywords.* Mathematical morphology, object recognition, partial occlusion, object separation.

## 1. Introduction

Identifying objects and estimating their pose is one of the major objectives in computer vision research (Ballard and Brown (1982), Fu et al. (1987)). When one such vision system is put to operate in a real environment to recognize, pick up and place different objects, it is expected to work reliably and fast. As a consequence, recognition and pose estimation algorithms are usually developed using simple features, as well as based on the assumption that the objects are well separated. This assumption may not be always true in real life, and an object may lie on another. Some work

has been done to recognize this type of overlapped objects (Ayache and Faugeras (1986), Boles and Cain (1982), Turney et al. (1985)). These algorithms are computationally expensive. Simpler recognition algorithms would work if images of these overlapped objects are separated. The present work addresses to this problem and tries to give a Mathematical Morphological algorithm to separate the images of overlapped objects.

Mathematical Morphology is becoming more and more a popular tool for dealing with shapes of objects in an image (Serra (1982), Haralick et al. (1987)). The main difference between a Mathematical Morphological technique and a traditional image processing technique is that the former treats images as an ensemble of sets rather than as a multi-dimensional signal. The language of Mathematical Morphology is that of Set Theory; and the operations are defined in terms of the interaction between an object and a structuring element.

*Correspondence to:* B. Chanda, Nodal Centre for Knowledge Based Computing, Indian Statistical Institute, 203 B.T. Road, Calcutta 700 035, India.

This report is organized as follows. Section 2 gives some definitions of Mathematical Morphology. The problem is defined more formally in Section 3. Section 4 presents a possible solution step by step in terms of Mathematical Morphological operations, and hence leads to a complete algorithm. Experimental results and conclusions are cited in Sections 5 and 6, respectively.

## 2. Preliminaries

Before describing the proposed algorithm, let us present some basic Mathematical Morphological operations. This will help us understanding the algorithms described in later sections.

A digital image is defined by a finite valued function over a discrete domain  $\mathbb{Z}^2$ . Let us assume the digital domain  $D \subset \mathbb{Z}^2$  is a rectangular array of size  $M \times N$ , i.e.,

$$D = \{(r, c) \mid r = 0, 1, 2, \dots, M-1; \\ c = 0, 1, 2, \dots, N-1\}$$

obtained by sampling of step size  $h$  along two orthogonal directions. Here we consider only binary images. A pixel  $(r, c)$  in  $D$  is defined as a foreground pixel or an object pixel if its value is 1, and as a background pixel if its value is 0. Therefore, a digital object  $A$  is a subset of pixels of  $D$  that have value 1 and are connected to each other. Binary Mathematical Morphology is a well known tool for manipulating shapes represented as a set of pixels. Preliminary operations are as follows. Since sets of pixels having value 1, i.e. digital objects, are of interest, the same notation may be used for both the digital object itself, and the binary image containing that digital object.

*Translation.* Let  $A$  be a subset of  $\mathbb{Z}^2$  and  $t$  be a point of  $\mathbb{Z}^2$ . We denote translation of  $A$  by the point  $t$  by  $A_t$  and define it by

$$A_t = \{p \in \mathbb{Z}^2 \mid p = a + t \text{ for some } a \in A\}.$$

*Dilation.* The dilation of a set  $A \subset \mathbb{Z}^2$  by another set  $B \subset \mathbb{Z}^2$  is denoted by  $A \oplus B$  and is defined by

$$A \oplus B = \{p \in \mathbb{Z}^2 \mid p = a + b \text{ for some } a \in A \\ \text{and } b \in B\}.$$

The set  $B$  is called a structuring element.

*Erosion.* The erosion of a set  $A \subset \mathbb{Z}^2$  by another set  $B \subset \mathbb{Z}^2$  is denoted by  $A \ominus B$  and is defined by

$$A \ominus B = \{p \in \mathbb{Z}^2 \mid p + b \in A \text{ for every } b \in B\}.$$

*Opening.* The opening of a set  $A \subset \mathbb{Z}^2$  by another set  $B \subset \mathbb{Z}^2$  is denoted by  $A \circ B$  and is defined by

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

*Closing.* The closing of a set  $A \subset \mathbb{Z}^2$  by another set  $B \subset \mathbb{Z}^2$  is denoted by  $A \bullet B$  and is defined by

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

*Conditional Dilation.* Let  $D$  be a structuring element. The conditional dilation of an image  $J$  by  $D$  with respect to an image  $I$  is denoted by  $J \oplus|_I D$  and is defined in an iterative manner. Let  $J_0 = J$  and  $J_n = (J_{n-1} \oplus D) \cap I$ , then

$$J \oplus|_I D = J_m$$

where  $m$  is the smallest index satisfying some constraint, say,  $J_m = J_{m-1}$ .

*Morphologic Pattern Spectrum.* There are several ways to define the morphologic pattern spectrum  $S_i(A, K)$  of an object  $A$  with respect to a set of structuring elements  $K = \{K_1, K_2, K_3, \dots\}$ . We may define  $S_i(A, K)$  as

$$S_i(A, K) = \#(A \circ K_i) \quad \text{for } i = 1, 2, 3, \dots$$

where  $\#(X)$  denotes the area of the object  $X$  or, more precisely, the number of pixels in  $X$ .

## 3. Problem definition

The present algorithm is devised to work with binary images of overlapped objects. Images captured may be gray level images. In that case, the image should be thresholded to get a binary image. A threshold is selected from the gray level histogram based on the assumption that there are only two types of regions: object and background. Thresholding of an image may be preceded by removal of noise by gray scale Mathematical Morphological opening and closing operations with

zero height structuring element. A second and direct method of getting a binary image is capturing the image using backlighting arrangement such that the background becomes bright and the object black. We implemented the second method to get the binary image. Like many other vision problems we start from the input image and reach the goal image through some processing. Processing techniques as well as values of relevant parameters are chosen using the domain knowledge. Based on this strategy the present problem may be formally represented as follows.

*Input image.* The input is a binary image which contains the silhouettes of industrial objects. The objects are elongated, like screw-driver, spanner, hammer etc. and each of their silhouettes has at least one long straight strip. Let us call the width of this long straight strip the 'principal width'. In the image an object may be overlapped by at most one other object somewhere within this strip. An image frame contains only one connected component of 1-pixels. That means during processing the image, we encounter only two types of situations: (i) the image contains a single object, and (ii) the image contains two objects overlapping one another. Since we are working only with the silhouettes of the objects, which object is lying on which does not matter in the subsequent processing.

Suppose, in the binary image, object<sub>1</sub> is represented by set  $A$  of pixels of value 1 and object<sub>2</sub> is represented by set  $B$  of pixels of value 1. Each of these  $A$ ,  $B$ , etc. has an elongated part (represented by a relatively thin and long straight strip), and some non-elongated parts. Now after a certain translation and rotation the sets  $A$  and  $B$  are transformed to  $A'$  and  $B'$ , respectively. Hence, the silhouette of the overlapped objects in the binary image is represented by

$$I = A' \cup B'.$$

*Knowledge.* The principal widths of the set of objects whose images are presented to the vision system are known [from the prototypes or models of the objects]. Suppose these widths are  $w_1$ ,  $w_2$ ,  $w_3$ , and so on, and let  $w$  be the minimum among them. Secondly, the lengths of the straight strips are also known. Suppose these lengths are  $l_1$ ,  $l_2$ ,

$l_3$ , and so on, and let  $l$  be the minimum among them. Thirdly, the angle between the axes, i.e., the lines parallel to the said long strips of the overlapped objects is not less than  $t^\circ$ .

*Goal.* The generation of the images of the isolated objects from the observed images of the overlapped objects, i.e., the generation of the images containing  $A'$  and  $B'$  individually starting from the binary image  $I$ .

#### 4. Proposed algorithm

In this section, we briefly describe the proposed algorithm alongwith the motivation, explanation and expected output of each step.

##### Step 1

First of all we like to determine the number of objects in the given image, and if the number of objects is more than one, then their orientation.

Generate  $S_i(I, K)$  where  $K_i$  ( $i = 1, 2, \dots, 180/d$ ) is a straight line segment shorter than  $l$  and much longer than  $w$ , and making an angle  $(i-1)d^\circ$  with the  $x$ -axis.

The interval  $d^\circ$  is chosen to be less than or equal to  $\arcsin(w/l)$ . This guarantees that for at least one value of  $i$ ,  $K_i$  must lie within the thin straight strip of  $A'$  (say) for some amount of translation. As a result, if  $I$  is opened with this  $K_i$  then the straight strip of  $A'$  will result; and it may be said that the straight strip (as well as  $A'$ ) makes an angle approximately equal to  $(i-1)d^\circ$  with the  $x$ -axis. Non-elongated parts of  $A'$  are absent in the result of opening since the length of the structuring element is usually greater than the diameter of the non-elongated parts. Now if the orientation of the structuring element is varied in either direction (i.e., the value of  $i$  is increased or decreased), it is expected that the structuring element does not fit within  $A'$  (although at some places the structuring element may fit in practice). Therefore,  $\#(I \circ K_i)$  falls sharply on both sides of the said value of  $i$ , and we see a sharp peak. If the width of  $A$  is such that for more than one value of  $i$ ,  $K_i$  matches the straight strip we get a dome type peak. (However,

even in this situation a particular  $i$  may show the best match.)

The same is true for  $B'$  also. Now if  $t$  is sufficiently greater than  $d$ , then two well separated peaks are formed corresponding to the straight strips of  $A'$  and  $B'$ .

If  $S_i(I, K)$  contains only one peak then the image contains one object only, and the image can be passed for subsequent processing. If  $S_i(I, K)$  [as shown in Figure 1, in the ideal case] contains two peaks for  $i=r$  and  $s$  (say), then it can be said that there are two objects, one of which is overlapped by the other, and their orientations are roughly  $(r-1)d^\circ$  and  $(s-1)d^\circ$ , respectively.

*Step 2*

Once we know that the image contains multiple objects as well as their orientations, we extract a thin straight strip of each object. Suppose the orientation of  $A'$  is  $(r-1)d^\circ$ ; then  $R$  as defined below represents the thin straight strip of  $A'$ :

$$R = I \circ K_r.$$

*Step 3*

Now  $R$  is conditionally dilated by  $D$ , a small disk structuring element, with respect to the input image

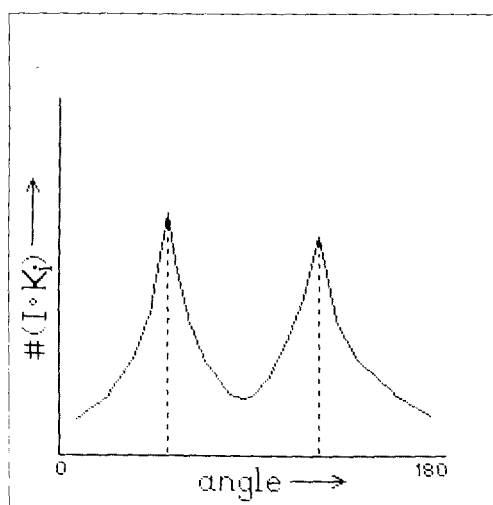


Figure 1. An ideal pattern spectrum:  $\#(I \circ K_i)$  versus angle that generates different structuring elements  $K_i$ .

until  $(R_{m+1} - R_m)$  contains two connected components only; where  $R_0 = R$  and  $R_i = (R_{i-1} \oplus D) \cap I$ .

Here the thin strip  $R$  of  $A'$  is grown to append non-elongated parts of  $A'$  to itself. If  $R$  is dilated by a disk structuring element  $D$  then it grows in all possible directions, but the intersection with  $I$  retains only those points which belong to  $I$ . As a result, the non-elongated parts of  $A'$  are grown and a part of the straight strip of  $B'$  is also augmented to  $R$  at the place of overlap. If we count the number of connected components of  $(R_m - R_{m-1})$  at this phase the count will be more than two, because the appended regions are at both ends of the straight strip portion of  $B'$  as well as at some other non-elongated portions of  $A'$ . After some iterations the non-elongated parts of  $A'$  are completely grown. If we continue the process,  $R$  will grow only along the straight strip of  $B'$ , and the number of connected components of  $(R_m - R_{m-1})$  will be two. We use this criterion to stop the iteration of conditional dilation. Finally,  $R_m$  becomes the object  $A'$  along with a part of the thin strip of  $B'$ .

Steps 2 and 3 are repeated with  $K_s$  also, and  $S_n$  is obtained. In fact,  $R_m$  and  $S_n$  are the images of the isolated objects with the exception that blister type regions are grown at the position where they overlapped.

*Step 4*

Next we obtain the common part of  $R_m$  and  $S_n$  which is denoted by  $X$ , i.e.,

$$X = R_m \cap S_n.$$

$X$  may be described in two different ways. In the first case, it may be said that  $X$  is the union of a portion of the straight strip of  $A'$  and some unwanted portion that has come from  $B'$ . And in the second case, it is the union of a portion of the straight strip of  $B'$  and unwanted portions of  $A'$ . In either of the cases, we are interested to delete the unwanted portions.

*Step 5*

Suppose  $X$  is opened again with a line structuring element of orientation  $(r-1)d^\circ$  and of length

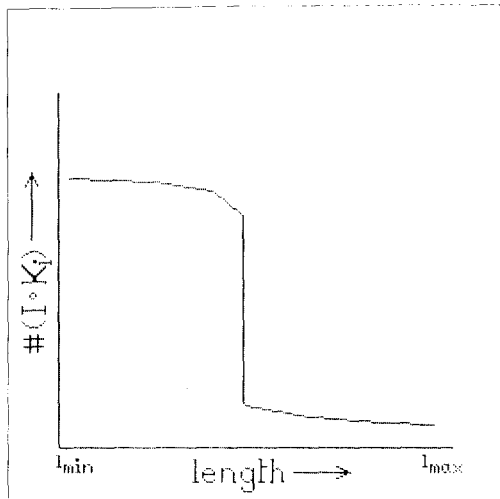


Figure 2. An ideal pattern spectrum:  $\#(I \circ K_i)$  versus length that generates different structuring elements  $K_i$ .

varying from the minimum possible value of  $l_{\min}$  to the maximum possible value of  $l_{\max}$ . As long as the length of the structuring element is less than  $w_B \cdot \sin(s-1)d^\circ$ , where  $w_B$  is the principal width of  $B$ ,  $X$  opened by the structuring element gives back a set which is ideally the same as  $X$  (however, in practice, the result of opening is not exactly the set  $X$  but is very close to it). But as the length of the structuring element crosses this limit, unwanted portions are deleted and the result of opening retains a portion of the straight strip of  $A'$  only since its length is much higher than the said limit. If we plot  $\#(X \circ K'_i)$  versus the length of  $K'_i$  then, in the ideal case, we should get a pattern spectrum as shown in Figure 2. Let  $K'_r$  be the structuring element of length greater than the value corresponding to the point where the curve in Figure 2 shows the sharp fall.  $K'_s$  is another structuring element obtained in a similar way.

*Step 6*

As described in Step 5, if  $X$  is opened by structuring elements  $K'_r$  and  $K'_s$ , respectively, the unwanted portions of  $B'$  and  $A'$  would be deleted. Suppose

$$X_1 = X \circ K'_r \quad \text{and} \quad X_2 = X \circ K'_s.$$

Hence,  $X_1$  is  $X$  without unwanted portions of

$B'$ ; or in other words,  $X_1$  is a portion of the straight strip of  $A'$  free from any significant noise due to the presence of  $B'$ . Similarly,  $X_2$  is  $X$  without unwanted portions of  $A'$ ; or in other words,  $X_2$  is a portion of the straight strip of  $B'$  free from any significant noise due to the presence of  $A'$ .

*Step 7*

From both  $R_m$  and  $S_n$  the common portion  $X$ , which contains a desired portion of the straight strip as well as some unwanted portions, is deleted, and then only the desired portion of the straight strip is added to get the desired objects. For example, since both  $R_m$  and  $X$  contains the same unwanted portions of  $B'$  and  $R_m$  is closer to the desired object  $A'$ ,  $X$  is deleted from  $R_m$ . As a result,  $R_m$  becomes free from the unwanted portion of  $B'$ ; but at the same time it loses a portion of the straight strip because  $X$  also contains that portion of  $A'$ . Now the union of the resultant image and  $X_1$ , which contains only the desired portion of the straight strip of  $A'$ , gives the desired image  $R'$  (say) corresponding to  $A'$ . Similarly, we obtain the image  $S'$  corresponding to  $B'$ . That means, if  $R'$  and  $S'$  are defined as

$$R' = (R_m - X) \cup X_1 \quad \text{and}$$

$$S' = (S_n - X) \cup X_2,$$

respectively, then the images containing  $R'$  or  $S'$  are the goal images.

**5. Experimental results**

The algorithm proposed in the previous section is implemented on several images. The result due to only one image is shown here. The program is written in C language and is executed on MicroVax II in Ultrix 32 environment.

As mentioned earlier, before applying the algorithm we have to set values of some assumed parameters. The values are usually set based on the knowledge of the object models. The values used in the experiment using the present set of images are:

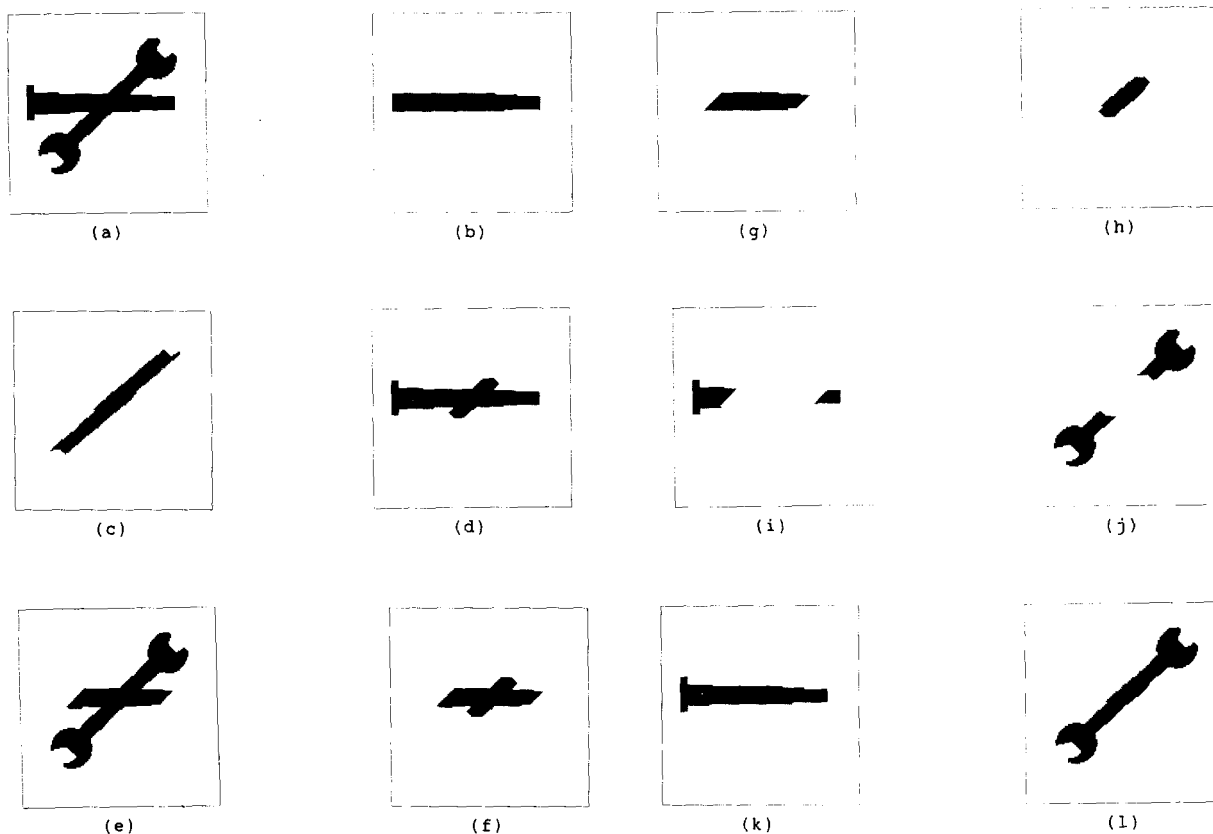


Figure 3. (a) Input binary image  $I$  of overlapped objects, (b)  $I \circ K_r$ , (c)  $I \circ K_s$ , (d)  $R_m$ , (e)  $S_n$ , (f)  $X$ , (g)  $X_1$ , (h)  $X_2$ , (i)  $R_m - X$ , (j)  $S_n - X$ , (k) goal image  $R'$ , and (l) goal image  $S'$ .  $K_r$ ,  $K_s$ ,  $R_m$ ,  $S_n$ ,  $X$ ,  $X_1$ ,  $X_2$  are described in the text.

$$d = 5, \quad t = 20, \quad l = 40,$$

$$\text{radius of disk } D = 2,$$

$$l_{\min} = \min\{w_1, w_2, w_3, \dots\} / 2,$$

$$l_{\max} = \max\{w_1, w_2, w_3, \dots\} * 2.$$

The experimental result is shown in Figure 3 in a step-by-step manner. The orientations of  $A'$  and  $B'$  found for this figure are  $0^\circ$  and  $40^\circ$ .

**6. Conclusions**

In this report a fast parallel algorithm has been proposed to separate the images of overlapped objects from their binary image. The algorithm is very useful in object recognition applications. The main advantage is that the recognition of in-

dividual objects and overlapped objects can be done by the same module if the input image is pre-processed by the proposed algorithm. However, the algorithm is designed only for simple overlapping cases. For example, if more than two objects overlap one another, or if they overlap at non-elongated portions, then the proposed algorithm cannot produce the expected result. For the former case the present algorithm may be modified to look for more than two peaks and two process accordingly. However, the second situation is not that simple, we are still looking for a solution.

**Acknowledgement**

The author likes to thank Mr. Srijeeb Majumder and Ms. Mily Gupta for their extensive help in programming and also for technical discussions.

**References**

- Ayache, N. and O.D. Faugeras (1986). HYPER: a new approach for the recognition and positioning of two dimensional objects. *IEEE Trans. Pattern Anal. Machine Intell.* 8, 44-54.
- Ballard, D.H. and C.M. Brown (1982). *Computer Vision*. Prentice-Hall, Englewood Cliffs, NJ.
- Boles, R.C. and R.A. Cain (1982). Recognizing and locating partially visible objects: the local-feature-focus method. *Int. J. Robotics Res.* 1, 57-82.
- Fu, K.S., R.C. Gonzalez and C.S.G. Lee (1987). *Robotics*. McGraw-Hill, Singapore.
- Haralick, R.M., S.R. Sternberg and X. Zhuang (1987). Image analysis using mathematical morphology. *IEEE Trans. Pattern Anal. Machine Intell.* 9, 532-550.
- Serra, J. (1982). *Image Analysis and Mathematical Morphology*. Academic Press, London.
- Turney, J.L., T.N. Mudge and R.A. Volz (1985). Recognizing partially occluded parts. *IEEE Trans. Pattern Anal. Machine Intell.* 7, 410-421.