Chapter from the book Vision Systems: Segmentation and Pattern Recognition
Downloaded from: http://www.intechopen.com/books/vision_systems_segmentation_and_pattern_recognition

Interested in publishing with IntechOpen?
Contact us at book.department@intechopen.com
1. Introduction

Robotics advances have generated an increasing interest in new research projects and developments. Nowadays this science has several new applications characterized by working in non-structured dynamic environments. As a result, the research on this emerging area is growing, and, specially, vision algorithms are constantly being improved. In many cases, navigation needs real-time answers. As robots often work in dynamic environments, it would be desirable that the system takes a decision and applies it before external conditions change.

Much work has been done on solving the problem of planning shortest paths between different locations within an environment (also known as a workspace) scattered with obstacles. For these solutions, the obstacles are usually considered as solid objects, and a collision-free path (of possibly shortest distance) must be found to navigate around them. However, not all path planning applications can be modelled as such a problem.

On the other hand, Mathematical Morphology (MM) is a useful tool in image analysis, commonly used to extract components of the image, like contours, skeletons and convex forms. Although there are some approaches that take into account topographical maps in order for a robot to navigate through a workspace, few approaches actually deal with Mathematical Morphology operations.

In this chapter we will focus on some of the research that we have completed in this field in the last few years. This way, two different robotic MM-based applications are discussed:

- Path planning, which is strongly influenced by the precision of the acquisition process. Thus, it can be modified both by the quality of the information obtained from the environment, and the attributes of the system and the environment in which it works.

- Image segmentation, which is an essential part of any intelligent system, since it is necessary for further processing such as feature extraction or object and face recognition, among others.

The research work here described has obtained very good experimental results and would contribute to the development of practical recognition and path planning systems. The use of vision improves the system, since once the visual information has been interpreted in...
order to provide a basic world representation, then the objects may be modelled to
determine a free path.

2. Mathematical Morphology Overview

The term Mathematical Morphology commonly denotes a branch of Biology that deals with
the shape and the structure of plants and animals. We use here the same word as a tool to
extract image components that are useful in the representation and description of regions,
such as contours, skeletons and convex forms.

The basis of Mathematical Morphology is the set theory. Sets in Mathematical Morphology
represent the shapes of the objects in an image. The morphological operations are based,
therefore, in geometric relations between the points of such sets.

This discipline focuses on the morphological transformations of images, i.e., erosion, dilation and its combinations, when some local operators, called structuring elements, are applied. The shape and the position of the origin of the structuring elements have a decisive influence on the final result of the morphological operator (Serra, 1992).

Mathematical Morphology describes objects as subsets of the Euclidean space. The
fundamental structure in Mathematical Morphology is the complete reticulum (Serra, 1982),
that is, a set \( \mathbb{R} \) where for all the elements \( \{X_i\} \in \mathbb{R} \), two fundamental laws exist: the
supremum (sup) or the minimum upper level (\( \lor (X_i) \)) and the infimum (inf) or the maximum
lower level (\( \land (X_i) \)). This structure explains the most common processes used in
Mathematical Morphology. A binary image can be modelled as a set belonging to a boolean
grid. A gray-level image is modelled as a function belonging to the set of the upper semi-
continuous functions.

The two basic operations defined in Mathematical Morphology, erosion and dilation, are
described below.

Set Erosion

The erosion of a set \( X \) using a symmetrical structuring element \( B \) is the locus of the centre of the
structuring element \( B \), when \( B \) is included in \( X \) (Serra, 1982). It can be written as:

\[
\varepsilon_b(X) = \{ y, \forall b \in B, y + b \in X \} = \bigcap_{b \in B} (X + b)
\]

(1)

Set Dilation

The dilation of a set \( X \) by a symmetrical structuring element \( B \) is the locus of the centre of the
structuring element \( B \), when \( B \) hit \( X \) (Serra, 1982). It can be written as:

\[
\delta_b(X) = \{ x + b, x \in X, b \in B \} = \bigcup_{b \in B} (X + b)
\]

(2)

The erosion and the dilation are dual operations: the dilation of a binary image is the
complemented erosion of the complementary image. The erosion of a binary image is the
complemented dilation of the complementary image.

From these basic operations, dilation and erosion, more complex transformations are
constructed, as opening (dilate the result of an erosion) and, its dual operation, closing
(eroe the result of a dilation), to implement basic filters.
3. Path Planning Applications of Morphological Filtering

3.1 A General Review of Path Planning

Robot path planning has proven to be a hard problem. There is strong evidence that its solution requires exponential time in the number of dimensions of the configuration space, i.e., the number of degrees of freedom (DOF) of the robot. This result is remarkably stable: it still holds for specific robots, e.g., planar linkages consisting of links serially connected by revolute joints (Joseph & Plantiga, 1985), and sets of rectangles executing axis-parallel translations in a rectangular workspace (Hopcroft & Wilfong, 1986). Though general and complete algorithms have been proposed (Canny, 1988), their high complexity precludes any useful application. This negative result has led some researchers to seek heuristic algorithms. While several of such planners solve difficult problems, they often fail or take much more computation times than simpler ones. The fact that their behavior is not well characterized is a major drawback: they cannot be used as black boxes in larger robot control systems.

Collision-free path planning, which assumes perfect knowledge of the world and stationary obstacles, is only the most basic motion-planning problem in robotics. Clearly, we would ultimately like robot planners to deal with issues such as uncertainties, moving obstacles, movable objects, and dynamic constraints. But every extension of the basic problem adds to computational complexity. For instance, allowing moving obstacles makes the problem grow exponentially with the number of moving obstacles (Canny, 1988). Before we can effectively investigate such extensions in large configuration spaces, it seems that we must better understand how to practically solve basic path planning.

Path-planning applications are so diverse that it is infeasible to design a tailor-made algorithm for every possible robot. Instead, we need general path-planning algorithms not bound to the specifics of any particular robot. We believe that between the two extreme types of planners suggested above—complete and heuristic—there is a place for practically efficient general planners achieving a weaker form of completeness. In other words, we may perhaps trade a limited amount of completeness against a major gain in computing efficiency. Full completeness requires the planner to always answer a path-planning query correctly, in asymptotically bounded time. A weaker, but still interesting form of completeness is the following: if a solution path exists, the planner will find one in bounded time, with high probability. We call it probabilistic completeness. This weaker completeness becomes particularly interesting if we can show that the planner’s running time grows slowly with the inverse of the failure probability that we are willing to tolerate.

3.2 Layout of Paths

Let us consider an application of the morphological primitives that are described in section 2, a method for the accomplishment of maps for gray-tone images (that should have been captured by the robot camera) as a previous step to collision-free path planning. Supposing that the obstacles in the picture room are represented in dark tones, the point is that images that have to be processed should be represented in a proper way. In a real situation a perspective transformation ought to be made.

Let \( I \) be the original image and \( I' \) the result image, then the followed algorithm to draw up the land map of an image that shows a room, and using a dilation with a \( n \times n \) SE, is the next one:
\[ U = \text{mean}(I) \]

threshold = \( U \)

\text{Binarize}(I)

\text{while} \ (\text{not} \ \text{End condition}) \ \text{do}

\text{Dilation}(I, I')

\text{Change black tone to gray tone in } I'

\text{end while}

Table 1. Land map algorithm

The result of this operation results in black objects surrounded by gray tones of greater intensity, until becoming into a white color, as we show in figure 1. White color is considered the tone in which the probability of collision with an object in a path followed by the robot is minimum. Consequently, in the 3D view (figure 2) the high zones would be a low risk for the robot.

(a) Picture of the room                        (b) Land map

Fig. 1. Land map for an image.

Fig. 2. 3D View of the land map.
Once determined the land map of a room, we want to show an example of high security trajectories that an autonomous vehicle could follow through it, to avoid the collision with the obstacles located in this workspace. So we are going to see some of the possible paths that the vehicle would decide to execute in figure 3, where o is the origin of the path and e is the end of the path. The algorithm is:

```
Select origin
while ((origin<>edge) OR (origin_value<>black)) do
    Find maximum in neighborhood 3x3
    if (maximum_value > origin_value)
        origin = maximum
    else, if (maximum_value = origin_value)
        origin = maximum
    else
        Follow path
end while
```

Table 2. Path planning algorithm

![Possible trajectories followed by a mobile robot.](image)

3.3 Improving the Path Planning Algorithm

In this section we propose an improved Mathematical Morphology-based path planning algorithm. Let us consider that the vision-based system works with grey-scale images. First, a map that separates obstacles from free-space is obtained; this initial processing method can be described as:
1. Apply a Gaussian smoothing to the original image.
2. Associate a set of symbols for each pixel. These symbols are extracted from:
   a. The gradient of the image.
   b. The variance considering each pixel's neighborhood.
3. Merge the results by means of the creation of a new image, where lower intensity pixels represent a higher probability of being classified as an obstacle.
4. Binarize the image, where obstacle pixels are labelled as '0' and free-space pixels as '1'.
5. Repeat the following steps:
   a. Implement the SE-decomposed morphological dilation.
   b. Change black tones to a grey tone, increasing the grey intensity for each iteration.
6. Obtain a map where higher intensity pixels constitute the obstacle-free zones.

Table 3. Improved land map algorithm

As soon as this map is obtained, the path planning algorithm can be executed. From the starting point of the path, the algorithm chooses the pixel with the lowest probability of collision in a 3x3 neighbourhood; once selected, the first movement is performed. This choice is carried out considering two criteria:
- The closeness to the obstacles. It is preferred to move to positions with high intensity, as they have a low probability of collision.
- The accomplishment of the task. This criterion makes the robot follow its path towards the destination point in case there are several similar or equal intensity values in the neighbourhood.

Hence, the selection of the next pixel in the robot's path will be completed by using a normalized weight which ensures that the new pixel has the lowest probability of collision in the neighbourhood. To do this, we define a set of weights $w_i$ that are estimated as:

$$ w_i = e^{a(dist_{old} - dist_{new})} $$

where $dist_{old}$ is the Euclidean distance from the current pixel to the destination one, $dist_{new}$ is the Euclidean distance from the selected new pixel to the destination one and $a$ is a real constant, so that $0 \leq a \leq 1$.

As a consequence, the robot will move to the pixel with the highest weight $w_{i_0}$; this operation will be repeated until it arrives to the destination or there is some failure due to a collision with non-detected obstacles. Therefore, the higher factor $a$ is, the bigger differences among the weights $w_i$ exist; obviously, factor $a$ is a critical element for the robot to follow an optimal path to the destination, as it provides the best weights $w_i$ to complete the path planning task.
In relation to this, the following section analyzes practically how to achieve a fast, powerful operation in some example situations.

3.4 Experiments

Let us consider now the results of some experiments that will test the suitability of our model. First of all, Fig. 4 (a) shows a world created for the robot to wander throughout it. We assume that there is no perspective distortion; this situation occurs when the camera’s optical axis is not perfectly perpendicular to the presentation surface. This distortion, which is part independent, can be compensated for using a homogeneous transformation.

At this point, the algorithm developed before is applied; thus, the resulting map after the initial processing method, using a 3x3 square SE, is depicted in Fig.4 (b).

From this map, a trajectory is followed after defining its origin and destination pixels (i.e., locations in real world). To do this, the weights wi must be estimated using Eq. (3) and, additionally, factor a should be defined. In Fig. 5 some example paths for a 3x3 square SE (where factor a varies from 0.1 to 1.0) are shown. Note that i refers to the initial point of the path, and f indicates the final point of the path.

The analysis of these examples shows that if factor a has a high value (Fig. 5 (a)), the algorithm will select a path that easily reaches the destination point, although it is not collision-free as the approaching to the obstacles can be sometimes dangerous. This will lead to the incorporation of some other sensing capabilities to prevent from collision. On the contrary, when factor a has a low value (Fig. 5 (c)), the robot may stop before completing its task, since it is preferred not to move close to the obstacles. As a consequence, factor a has a better behaviour when it makes the robot follow an optimal or semi-optimal path that keeps it away from collision (in this example, a = 0.1, see Fig. 5 (b)).
Nevertheless, the choice of an appropriate factor will depend mainly on the map of the environment produced after the initial method. Moreover, there must be a training process to determine a well-suited factor before a real operation.

4. Image Segmentation with Mathematical Morphology

Designing an image segmentation scheme needs the consideration of two features of visual recognition: cost and uncertainty. Visual recognition is generally costly because the image data is large. Visual information contains uncertainty from many sources such as discretization. In general, the more observation a vision system performs, the more information is obtained, but the more cost is required. Thus, a trade-off must be considered between the cost of visual recognition and the effect of information to be obtained by recognition. In relation to this, some of the most popular approaches which provide low computation times and good information are the threshold techniques and the edge-based methods (Ouadfel & Batouche, 2003), (Pal, & Pal, 1993).

Threshold techniques, which make decisions based on local pixel information, are effective when the intensity levels of the objects are exactly outside the range of the levels in the background. These thresholding algorithms are simple and give very good results, but deciding the threshold values is not easy. Specially, this is a really serious problem for an automated vision system, as the system should decide the threshold values taking its own decision.

On the other hand, edge-based methods (e.g., gradient operators) focus on contour detection. They involve finding the edges of objects in the image and using this edge information to achieve the complete boundaries for the main objects in the image. Edge detection has many problems, especially when working with noisy images, since it could even fragment the true edges.

To overcome these problems, we propose a method that combines both thresholding and gradient operators: the so-called Morphological Gradient Threshold (MGT) segmentation, as described in Table 4. It consists of 7 main steps, where the gradient and the Laplacian are calculated in terms of Mathematical Morphology operations and the optimal threshold value is selected by measuring the lowest distance between the ideal segmentation and a collection of MGT segmented images.
Step 1. Image smoothing.
Step 2. Global dilation and erosion.
Step 3. For every pixel, create a list of symbols by means of the Morphological gradient and the Morphological Laplacian.
Step 5. Binarization of the pixel-symbol map.
Step 6. Computation of a suitable measure to obtain the optimal threshold.
Step 7. Obtention of the MGT segmented image.

Table 4. MGT segmentation algorithm

4.1 Construction of a Pixel-Symbol Map
In every digital image there is a certain amount of white noise. To avoid the noise effects, which only consume computation time and affect the real image features, an initial filtering process has to be applied. There are many algorithms to accomplish this task; in our approach a Gaussian filter has been chosen, since it preserves many of the image features while its computational cost can be assumed in a real-time environment. For more information see (Basu & Su, 2001).

Once the noise is eliminated, the point is how to create the pixel-symbol map. To do this, let us consider first the computation of some derivative-based operations, i.e., the gradient and the Laplacian.

Edge detection is a main problem in image analysis. There are many approaches to obtain edges by means of the gradient of an image (e.g., Prewitt or Sobel operators). Among all of these methods we find the morphological gradient, which uses the Mathematical Morphology operators.

Therefore, one can define the morphological gradient of an image $X$ by a structuring element (SE) $B$, $\rho_B(X)$, as:

$$\rho_B(X) = \frac{1}{2} (\delta_B(X) - \varepsilon_B(X)) \quad (4)$$

where $\delta_B(X)$ and $\varepsilon_B(X)$ are, respectively, the dilation and the erosion of an image $X$ by a SE $B$.

The following step is to calculate the second derivative, the Laplacian. Again, we have chosen a morphological implementation for the Laplacian, as we can use with costless time the previously pre-calculated erosion and dilation. Thus, the morphological Laplacian of an image $X$ by a SE $B$, $\Lambda_B(X)$, is defined as:

$$\Lambda_B(X) = \frac{1}{2} (\delta_B(X) + \varepsilon_B(X)) - X \quad (5)$$

The results for a gray-scale image after these initial steps are shown in Fig. 6, where the SE $B$ is a 3x3 square.
The next task is building a map that characterizes properly the pixels for a good segmentation. Thus, the pixel-symbol map \( m(x, y) \) is obtained as follows:

\[
m(x, y) = \begin{cases} 
128 & \text{if } \rho_B(x, y) < \text{MGT} \\
255 & \text{if } \rho_B(x, y) \geq \text{MGT} \text{ and } \Lambda_B(x, y) \geq 0 \\
0 & \text{if } \rho_B(x, y) \geq \text{MGT} \text{ and } \Lambda_B(x, y) < 0 
\end{cases}
\]  

where MGT is the morphological gradient threshold and \((x, y)\) is a pixel in \( X \). The resulting image has three different gray-levels, according to if a pixel belongs to an object, to the background or to the borders.

The choice of the threshold value is one of the most difficult tasks, since the final result is highly dependent on many factors, such as lighting conditions, objects texture or shading. Fig. 7 shows the results of the construction of the pixel-symbol map for the image in Fig. 6, with several different MGT values.

Though many practical systems utilize an experimentally obtained threshold, in this work we consider the use of an automated thresholding system. This method takes into account binary image metrics to compare the segmentation results and, afterwards, to establish the quality level of the obtained segmentation, as it is described in the following section.

### 4.2 A Measure of the Quality of the Segmentation

A main problem in computer vision is to be able to compare the results using a proper metrics. This will quantify the differences between two images and, if binary images are used, the method would be both easily implementable and low computationally complex. In
our system we are interested in measuring the distance between image $G$ (the map after gradient thresholding) and image $A$ (the ideal segmentation). Thus, it will establish the optimal MGT value.

Hence, the map must be binarized first. To do this, we must recall that $m(x,y)$ has only 3 gray-levels (Eq. (6)): 0, 128 and 255. For simplicity, let us consider that the threshold is the same as in the construction of $m(x,y)$, i.e., the gradient threshold MGT. The results of this process are shown in Fig. 8.

![Fig. 8. Binarization with different MGT values: (a) Gradient mean. (b) MGT = 0.9 * max($p_B(X)$). (c) MGT = 0.8 * max($p_B(X)$).](image)

Next, a reliable measure to compare the obtained image segmentation with an ideal segmentation must be selected.

As proved in (Pujol et al., 2000), a good error measurement for binary images is $\Lambda^p(A,G)$, defined as the $p^{th}$ order mean difference between the thresholded distance transforms of two images: $A$ (the ideal segmentation) and $G$ (the binary pixel-symbol map). Let us define first some previous terms:

- Let $X$ denote the pixel raster.
- A binary image $A \subseteq X$ is a set $A = \{ x \in X : A(x) = 1 \}$.

If $o(x,y)$ is the distance between two pixels $x$ and $y$, the shortest distance between a pixel $x \in X$ and $A \subseteq X$ is defined as:

$$d(x,A) = \inf \{ o(x,a) : a \in A \}$$  \hspace{1cm} (7)

Then, for $1 \leq p \leq \infty$, we define:

$$\Delta^p(A,G) = \left[ \frac{1}{N} \sum_{x \in X} w(d(x,A)) - w(d(x,G)) \right]^{\frac{1}{p}}$$  \hspace{1cm} (8)

where $N$ is the total number of pixels in $X$ and $w(t) = \min(t,c)$, for $c>0$.

Intuitively, $\Delta^p(A,G)$ measures the suitability of an estimated image to be used instead of the real one.

Now, we can evaluate the goodness of our segmentation scheme.

### 4.3 Experiments

Let us show now the results of some experiments completed for our model. The tests have been performed with a set of real images, whose pixel-symbol maps have been calculated.
for different MGT values. Then, after applying the binarization process, the distance $\Delta'(A,G)$ has been computed.

Table 5 shows the results for the image in Fig. 8, where $p = 2$, $c = 5$.

<table>
<thead>
<tr>
<th>MGT value</th>
<th>Distance $\Delta'(A,G)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{MGT} = 0.95 \times \max(\rho_B(X))$</td>
<td>0.2722</td>
</tr>
<tr>
<td>$\text{MGT} = 0.9 \times \max(\rho_B(X))$</td>
<td>0.1988</td>
</tr>
<tr>
<td>$\text{MGT} = 0.85 \times \max(\rho_B(X))$</td>
<td>0.3412</td>
</tr>
<tr>
<td>$\text{MGT} = 0.8 \times \max(\rho_B(X))$</td>
<td>0.3704</td>
</tr>
<tr>
<td>$\text{MGT} = 0.75 \times \max(\rho_B(X))$</td>
<td>0.4966</td>
</tr>
</tbody>
</table>

Table 5. Results obtained after the segmentation process.

As shown, the lowest distance is obtained when $\text{MGT} = 0.9 \times \max(\rho_B(X))$. Fig. 9 compares the ideal segmentation and the MGT segmentation with the lowest $\Delta'(A,G)$ distance. Intuitively, if we compare the previous results in Fig. 8, the selected MGT value is quite similar to the ideal segmentation.

Let us consider now a more complex real image in order to confirm the accuracy of our technique to give an automated extraction of the threshold value with the best behavior. Fig. 10 and Table 6 show the results.
Table 6. Results obtained after the segmentation process.

The minimum distance is obtained again when $MGT = 0.9 \times \max(p_b(X))$ and, as a consequence, we can conclude that the parameters used for this segmentation are near optimal, as they have a behavior very close to ideal segmentation. Nevertheless, the threshold could be adaptively updated so as to assume the real conditions in which every image has been taken by the vision system.

5. Conclusion

In general terms, the path planning process for suitable morphological gradient threshold. To do this, global morphological operators have been used to compute the gradient and the Laplacian and, after a proper binarization, the distance between the ideal segmentation and the MGT segmentation has been computed. As a consequence, the gradient threshold with the lowest distance has been selected as the optimal threshold value. Experimental results show that our model is fast and robust and could be applied for real-time imaging.

As a future work, to fully appreciate the implications of incorporating a path planner into a robot system it is necessary to consider a real robot system. The use of simulations can give a good idea of the ability to solve the basic problem but it is also necessary to consider how the planner will receive input data and how the output path will be used to generate a trajectory and be implemented by a physical robot. This will make possible a more accurate designing method so that the robot internal hardware and software could be efficiently implemented.

Finally, the results of our research could be extended to object classification and recognition. It would be also an interesting task to consider new simulation experiments with different environments, such as image sequences obtained from a camera placed in a robot platform, where real-time constraints have a great influence a mobile robot is strongly influenced by the precision of the acquisition process. Thus, it can be modified both by the quality of the information obtained from the environment, and the attributes of the system and the environment in which it works.

In this chapter, we have developed a proposal of a model for the generation of a map in unknown environments. To do this, we have described a path planning technique for autonomous robots that uses morphological filtering. In this method, some high security paths for a robot to follow are computed; the experimentation shows that the prototype is robust and can be applied in real time for many robotic applications, since it is a very quick algorithm to compute free paths with high probability of no collision.
On the other hand, image segmentation is an essential issue since it is the first step for image understanding, and any other step, such as feature extraction and recognition, heavily depends on its results. In this chapter, we have also described a novel approach to image segmentation based on the selection of a in the final recognition results.

6. References


Research in computer vision has exponentially increased in the last two decades due to the availability of cheap cameras and fast processors. This increase has also been accompanied by a blurring of the boundaries between the different applications of vision, making it truly interdisciplinary. In this book we have attempted to put together state-of-the-art research and developments in segmentation and pattern recognition. The first nine chapters on segmentation deal with advanced algorithms and models, and various applications of segmentation in robot path planning, human face tracking, etc. The later chapters are devoted to pattern recognition and covers diverse topics ranging from biological image analysis, remote sensing, text recognition, advanced filter design for data analysis, etc.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: