OORJA

# *Image Segmentation Techniques: A Survey*

Ms. Navjeet Kaur


## Abstract

An image is a spatial representation of a two-dimensional or three-dimensional scene. Segmentation of an image is the process that subdivides an image into its constituent regions or objects. The goal of image segmentation is to cluster pixels into salient image regions, i.e., regions corresponding to individual surfaces, objects, or natural parts of objects. There are now a wide variety of image segmentation techniques, some considered general purpose and some designed for specific classes of images. Since there is no general solution to the image segmentation problem, these techniques often have to be combined with domain knowledge in order to effectively solve an image segmentation problem for a problem domain. This paper provides a survey of achievements, problems being encountered, and the open issues in the research area of image segmentation and usage of the techniques in different areas. We consider the techniques under the following two groups: Low level image segmentation techniques and high level image segmentation techniques.


## Introduction

The goal of image segmentation is the detection of meaningful structures from a cluttered scene. Most current segmentation techniques take a bottom-up approach, where local image properties such as feature similarity (brightness, texture, motion etc) are used to detect coherent units. Unfortunately, image segmentation becomes very difficult in poor data conditions like shadows, occlusions and noise. In such situations, the detected coherent units often do not coincide with our perception of objects in a scene. The automatic detection and recognition of visual objects in images, on the other hand, has been among the prime objectives of computer vision for several decades. Segmentation algorithms might help to solve the object detection task by partitioning the image into meaningful parts that might serve as the inputs of a classification system. Meaningful segmentation is the first step from low-level image processing transforming a greyscale or colour image into one or more other images to high-level image description in terms of features, objects, and scenes. The success of image analysis depends on reliability of segmentation, but an accurate partitioning of an image is generally a very challenging problem.

Segmentation of an image can be broadly classified into: Low level segmentation and High level segmentation.

## Low level segmentation

In this type one employs only information present in image. Segmentation algorithms under this category are based on one of two basic properties of grey-level values:

Discontinuity- Image is partitioned based on abrupt changes in grey-level (isolated points, lines and edges). Under this we have edge detection using Sobel operator, Robert's cross operator, Perwitt's operator, Laplacian of guassian.

Similarity- The process of segmentation involves partitioning the entire scene in a finite number of regions. Under this we have thresholding, region growing and region splitting and merging.

**Ms. Navjeet Kaur**
Lecturer
Amity Institute of
Information Technology,
Amity University,
Rajasthan

The easiest way to detect edges in an image is to look for places in the image where the intensity changes rapidly, using one of these two criteria:

(1) Places where the first derivative of the intensity is larger in magnitude than some threshold.

(2) Places where the second derivative of the intensity has a zero crossing. (Refer Figure I, II, III)

### Sobel Operator

The operator consists of a pair of 3×3 convolution kernels as shown in Figure IV. One kernel is simply the other rotated by 90° .

It is sensitive to both horizontal and vertical edges. The kernels can be applied separately to the input image, to produce separate measurements of the gradient component in each orientation (say Gx and Gy).These can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient. The gradient magnitude is given by:

$$|G| = \sqrt{Gx^2 + Gy^2}$$

The angle of orientation of the edge (relative to the pixel grid) giving rise to the spatial gradient is given by:

$$\theta = \arctan(Gy/Gx)$$

However, when there is noise with high magnitude, this method may fail because first and second derivatives are very sensitive to the existence of noise.

### Thresholding

Thresholding is the simplest method of image segmentation. From a grayscale image, thresholding can be used to create binary images. During the thresholding process, individual pixels in an image are marked as "object" pixels if their value is greater than some threshold value and as "background" pixels otherwise. An object pixel is given a value of "1" while a background pixel is given a value of "0." Finally, a binary image is created by coloring each pixel white or black, depending on a pixel's labels. The key parameter in the thresholding process is the choice of the threshold value. An initial threshold (T) is chosen; this can be done randomly or according to any other method desired. For an input image f (i, j) and output image g (i, j) (Refer Figure V):

For each pixel (i, j)

g(i, j) ={ 1 for f(i,j) > = T

0 for f( i,j) < T }

It is a simple technique and more often used. It is easy to use in hardware, intrinsically parallel. But in this technique, threshold is a parameter which is difficult to adjust automatically in general.

### Canny's Edge Detection Algorithm

The Canny edge detection algorithm is known to many as the optimal edge detector. Canny proposed a method to detect edges that can be found in his paper, "A Computational Approach to Edge Detection". In his paper, he followed a list of criteria to improve current methods of edge detection. The first and most obvious is low error rate. It is important that edges occurring in images should not be missed and that there be no responses to non-edges. The second criterion is that the edge points be well localized. In other words, the distance between the edge pixels as found by the detector and the actual edge is to be at a minimum. A third criterion is to have only one response to a single edge.

### Split-and-Merge Segmentation

The top-down split-and-merge algorithm considers initially the entire image to be a single region and then iteratively splits each region into sub regions or merges adjacent regions until all regions become uniform or until the desired number of regions have been established. A common splitting strategy for a square image is to divide it recursively into smaller and smaller quadrants until, for any region R, the uniformity predicate P(R) is TRUE. The strategy builds a top-down quadtree: if P (image) is FALSE, the image is divided into four quadrants; if P(quadrant) is FALSE, the quadrant is divided into subquadrants; and so on (Refer Figure VI) :

The splitting stage alternates with a merging stage, in which two adjacent regions $R_i$ and $R_j$ are combined into a new, larger region if the uniformity predicate for the union of these two regions, P($R_i$ *″ $R_j$), is TRUE.

There are problems with regional segmentation of any form. Meaningful regions may not be uniform. Surface properties of a solid body will vary in brightness or colour dependent on the existence of slowly varying gradients due to lighting conditions Boundary segmentation is much more widely applied than regional segmentation for several reasons. Algorithms are usually less complex. They tend to use local properties and software and hardware implementations are readily available. Humans may use edge detection: there is evidence of links between edge detection and early human visual processing, which lead to the observation that contoured images are more easily identified than regional images, particularly when degraded in some form.

All the methods of edge detection given above have certain limitations.The edges extracted using the classical methods often do not necessarily correspond to boundary objects.The edge detection techniques depend on the information contained in the local neighbourhood of the image. Most of the edge detection techniques do not consider model-based information embedded in an image. After the edge points are extracted from the image, these points are linked in order to determine boundaries.The edge linking process sometimes leads to discontinuities and gaps in the image. The edge linking methods often resort to arbitrary interpolation in order to complete boundary gaps. It is often difficult to identify and classify spurious edges.

**High Level Segmentation**

To overcome limitations of low level segmentation techniques we have high level segmentation techniques. In this technique a priori knowledge about shape, texture, colour, or position of the object in question is somehow included in the search procedure. High level segmentation can be performed by using various deformable models.

A deformable model can be characterised as a model, which under an implicit or explicit optimization criterion, deforms the shape to match a known object in a given image.

**Deformable models are divided into two classes:**

Free-form models - These models can represent any arbitrary shape as long as some constraints like continuity, smoothness etc. are satisfied.These models are often called active contours.

Parametric models- These can encode a specific shape and its variation where the shape can be characterised by a parametric formula or a prototype and its deformation modes.They can be of two kinds:

Analytical deformable templates

Prototype-based deformable templates

**Free-form deformable models (or snakes)**

Free-form deformable models, also called active contour models, assume very little structure about the object shape except for some regularisation constraints like continuity and/or smoothness of the boundary. Snakes are the most popular free form deformable models.

Snakes are computer-generated curves that move within images to find object boundaries. Kass developed this novel technique for image segmentation. It is an elastic contour which is fitted to features detected in an image. The nature of its elastic energy draws it more or less strongly to certain preferred configurations, representing prior information about shape which is to be balanced with evidence from an image.

A gradient descent procedure is used to slowly bring an initial contour to the edges of the object of interest in the image.These models do not encode any specific shape information about, and are very sensitive to the initial position of the snake and to image noise. To overcome these limitations, more constrained energy functions with balloon forces, attractors and tangent constraint, region information and gradient vector flow have been proposed (Refer Figure VII).

The blue dot marks the starting point and end point of the snake curve. Ideally the snake will have minimized its energy when it has positioned itself on the contour of the object.

Snakes are able to solve a large class of segmentation problems that had eluded more conventional techniques.

Snakes have several advantages over classical feature extraction techniques. They can be controlled interactively by using appropriately placed springs and volcanoes.They are easy to manipulate because the external image forces behave in an intuitive manner.They are autonomous and self-adapting in their search for a minimal energy state. Snakes are not without their drawbacks.They can often get stuck in local minima states; this may be overcome by using simulated annealing techniques at the expense of longer computation times. They often overlook minute features in the process of minimizing the energy over the entire path of their contours.

**Parametric Deformable Models**

This approach employs prior knowledge about the shape of the object in a direct manner. This prior shape information is specified as a sketch, binary template or a parametric prototype. The a priori information is then encoded either in the form of edge information from a binary template or the parameter vector. That information does not need to be exact in the sense that it matches the boundaries of the image exactly. They are categorised into

Analytical deformable templates

Prototype-based deformable templates

### Analytical Templates

Analytical deformable templates are defined by a set of analytical curves. Hence, the geometrical shape is parameterised directly. The template is represented by a set of curves which is uniquely described by some parameters. The geometric shape of the template can be changed by varying the parameters in the analytical expressions, and the variations in the shape are determined by the distribution of the parameter values. This representation requires that the geometrical shape is well structured.

### Prototype-Based Templates

This approach was first presented by Grenander et al. who described a systematic framework to represent and generate patterns from a class of shapes. A shape is represented by a set of parameters and a probability distribution on the parameters is specified to allow a flexible bias towards a particular shape. The active shape model, suggested by Cootes, is a prototype based deformable template.

Disadvantages: None of the deformable template models have the generality which is comparable with that of snakes. Deformable templates are strongly constrained by a priori shape knowledge and it is not allowed to take any arbitrary shape. The models that are hand-crafted for some specific problems are of course not general. Some models deform templates in a rigid and random way which is not always suited for medical images.

### Image Segmentation using Fuzzy Logic

Fuzzy inference system is useful in edge detection in digital image processing. A fuzzy inference system consists of three steps fuzzification; modification in inputs by applying rule based system and then defuzzification in order to convert the fuzzy output into crisp values. Fuzzy logic provides a good method of thresholding an image. But the disadvantage of this method is that many difficulties in image processing arise because the data/tasks/results are uncertain.

### Conclusions

Due to the limitations of low level segmentation techniques, the practical usefulness of low-level segmentation algorithms for the purpose of object recognition is questionable and, indeed, the currently best approaches to object recognition do not employ low-level segmentation. Despite the fact that bottom-up image segmentation is sometimes considered as an important pre-processing step for object recognition, the actual usefulness of such an approach in real-world recognition scenarios is still an ongoing debate. For real-world scenes it is often difficult to do segmentations of sufficiently high quality that would allow reliably extracting object-specific information like shape, coloring distribution, and texture features etc. It is, thus, not surprising that the currently best object recognition systems do not use low-level segmentations. Deformable models offer a unique and powerful approach to image analysis that combines geometry, physics, and approximation theory. They have proven to be effective in segmenting, matching, and tracking anatomic structures by exploiting (bottom-up) constraints derived from the image data together with (top-down) a priori knowledge about the location, size, and shape of these structures. Deformable models are capable of accommodating the significant variability of biological structures over time and across different individuals.

### References


- R C Gonzalez and RE Woods (2006). "*Digital Image Processing*", Addison Wesley Publishing company Inc., Second edition.
- M. Kass, A. Witkin, and D. Terzopoulos (1988). *"Snakes: Active contour models", International Journal of Computer Vision, 1(4):321–331, 1988.*
- Nick Efford (2000). "*Digital Image Processing: A Practical Introduction Using Java™*", Pearson Education, 2000.
- Canny, J. (1986). "A Computational Approach To Edge Detection", *IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679–698, 1986.*
- A.K. Jain, Y. Zhong and M-P. Dubuisson-Jolly (1998). "Deformable template models: A review", *Signal Processing 71 (1998), pp. 109–129.*
- Eikvil L. (2000). "Deformable Models, Norsk Regnesentral Anv Dataforskning", Nov., 2000.

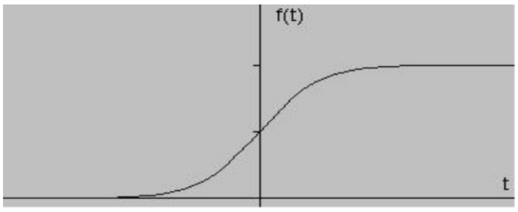


Figure I: Edge is Shown by Jump in Intensity

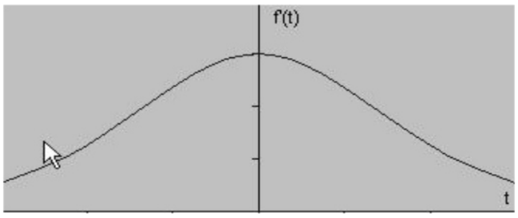


Figure II: Gradient of the Signal in Fig I

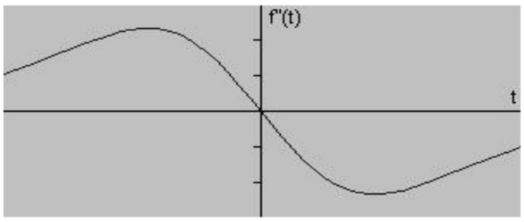


Figure III: Second Derivative of the Signal

| | | |
|---|---|---|
| -1 | 0 | +1 |
| -2 | 0 | +2 |
| -1 | 0 | +1 |

Gx

| | | |
|---|---|---|
| +1 | +2 | +1 |
| 0 | 0 | 0 |
| -1 | -2 | -1 |

Gy

Figure IV : Sobel Operator

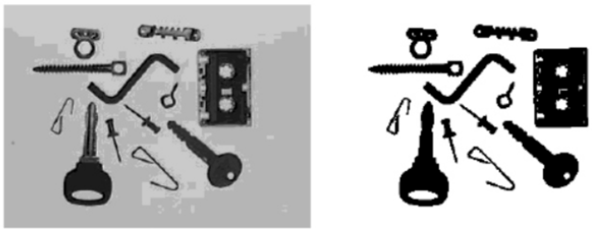

Figure V : Threshold Influence

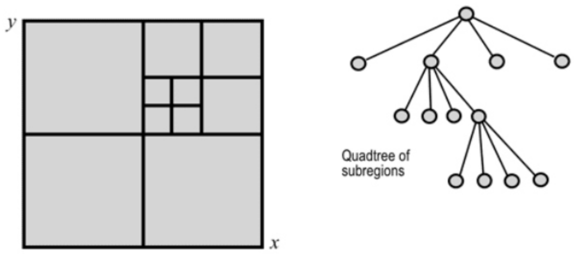


Figure VI : Split-and-Merge Segmentation

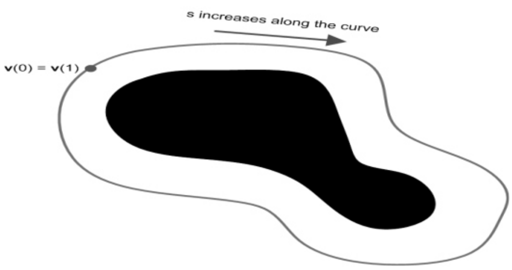


Figure VII : Illustration of a Parametric Snake Curve V(s)