Object Segmentation Using Active Contours: A Level Set Approach

Farnaz Shariat

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation
https://scholar.uwindsor.ca/etd/336

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
Object Segmentation Using Active Contours:
A Level Set Approach

By
Farnaz Shariat

A Thesis
Submitted to the Faculty of Graduate Studies through the School of Computer Science in Partial Fulfilment of the Requirements for the Degree of Master of Science at the University of Windsor

Windsor, Ontario, Canada
2009
© 2009 Farnaz Shariat
Object Segmentation Using Active Contours: A Level Set Approach

By:
Farnaz Shariat

APPROVED BY:

__________________________
Dr. Yung Tsin, Internal Reader
School of Computer science

__________________________
Dr. Jonathan Wu, External Reader
Electrical Engineering Department

__________________________
Dr. Boubakeur Boufama, Advisor
School of Computer science

__________________________
Dr. Imran Ahmad Chair of Defense
School of Computer science

15 June 2009
Author’s Declaration of Originality

This thesis includes one original paper that has been previously published for publication in peer reviewed journals, as follows:

<table>
<thead>
<tr>
<th>Thesis Chapter</th>
<th>Publication title/full citation</th>
<th>Publication status*</th>
</tr>
</thead>
</table>

I certify that I have the copyright of the paper so that I can include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as graduate student at the University of Windsor.

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone’s copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.
Abstract

Image segmentation is responsible for partitioning an image into sub-regions based on a preferred feature. Active contour models have widely been used for image segmentation. The use of level set theory has enriched the implementation of active contours with more flexibility and simplicity. The past models of active contours rely on a gradient based stopping function to stop the curve evolution. However, when using gradient information for noisy and textured images, the evolving curve may pass through, or stop far from the salient object boundaries.

Therefore, we propose using a polarity based stopping function. Comparing to the gradient information, the polarity information accurately distinguishes the boundaries or edges of the salient objects more precisely. Hence, with combining the polarity information with the active contour model, we obtain a fast and efficient active contour model for salient object detection. Experiments are performed on several images to show the advantage of the polarity based active contour.

**Keywords:** Computer vision, image segmentation, active contours, level set theory, object detection, variational level set, polarity
Acknowledgements

First I would like to thank Dr. Boubakeur Boufama, my advisor and thesis supervisor. A special acknowledgement goes to Dr. Riadh Ksantini for his guidance and assistance as co-supervisor. I would also like to thank Dr. Tsin, Dr. Wu and Dr. Ahmad for serving on my committee.

Lastly, and most importantly, I wish to thank my parents. They bore me, raised me, supported me, taught me, and loved me. To them, I dedicate this thesis.
Table of Contents

AUTHOR’S DECLARATION OF ORIGINALITY ................................................iii

ABSTRACT ........................................................................................................iv

ACKNOWLEDGEMENT .......................................................................................v

1 INTRODUCTION ..............................................................................................1

1.1 Overview of computer vision .................................................................1

1.2 Definition of computer vision ...............................................................2

1.3 Motivations of the thesis .................................................................2

1.4 Overview of the thesis .................................................................4

2 IMAGE SEGMENTATION .................................................................................6

2.1 Definition .................................................................................................7

2.2 What is a good segmentation? ..............................................................7

2.3 Classification of algorithms ..............................................................8

2.3.1 Thresholding .....................................................................................9

2.3.2 Edge-based segmentation ...........................................................12

2.3.3 Region-based segmentation .........................................................16
2.3.4 Hybrid methods ................................................................. 19

2.4 Summary .................................................................................. 21

3 ACTIVE CONTOURS ......................................................................... 22

3.1 Snakes ....................................................................................... 22

3.2 Level set methods ....................................................................... 27

3.2.1 Level set concept ................................................................. 27

3.2.2 Level set dictionary and technology ...................................... 29

3.2.3 Numerics ............................................................................... 31

3.2.4 Level set in segmentation ...................................................... 31

3.3 Summary ..................................................................................... 37

4 AN IMPROVED TEXTURE-RESISTANT ACTIVE CONTOUR MODEL ................................................................................................................. 39

4.1 Polarity definition ........................................................................ 40

4.2 Method description ....................................................................... 43

4.3 Summary ..................................................................................... 46

5 EXPERIMENTS .............................................................................. 48
5.1 Comparison results.................................................................48

5.2 Summary...............................................................60

6 CONCLUSION .................................................................61

REFERENCES .................................................................63

VITA AUCTORIS .............................................................76
List of Figures

Figure 1.1 objects on a uniform background .......................................................... 3

Figure 1.2 Object on a noisy background .............................................................. 3

Figure 1.3 Object on a textured background .......................................................... 4

Figure 2.1 Different ways to segment an image ...................................................... 8

Figure 3.1 Active contour’s movement ................................................................... 23

Figure 3.2 The construction of level set function .................................................. 28

Figure 3.3 showing all the cases for the Chan & Vese method. ....................... 34

Figure 4.1 Active contour result using Li’s algorithm (clear Background) .......... 39

Figure 4.2 Active contour result using Li’s algorithm (not clear background) .. 40

Figure 4.3 Different values of polarity .................................................................. 41

Figure 4.4 Comparison between polarity and gradient information ................. 42

Figure 5.1 Comparison Results set no.1 ............................................................... 49

Figure 5.2 Comparison Results set no.2 ............................................................... 50

Figure 5.3 Comparison Results set no.3 ............................................................... 51

Figure 5.4 Comparison Results set no.4 ............................................................... 52

Figure 5.5 Comparison Results set no.5 ............................................................... 52

Figure 5.6 Comparison Results set no.6 ............................................................... 53

Figure 5.7 Comparison Results set no.7 ............................................................... 53

Figure 5.8 Comparison Results set no.8 ............................................................... 54

Figure 5.9 Comparison Results set no.9 ............................................................... 54
Figure 5.10 Comparison Results set no.10 ............................................................ 54
Figure 5.11 Comparison Results set no.11 ............................................................ 55
Figure 5.12 Comparison Results set no.12 ............................................................ 55
Figure 5.13 Comparison Results set no.13 ............................................................ 56
Figure 5.14 Comparison Results set no.14 ............................................................ 56
Figure 5.15 Comparison Results set no.15 ............................................................ 56
Figure 5.16 Comparison Results set no.16 ............................................................ 57
Figure 5.17 Comparison Results set no.17 ............................................................ 57
Figure 5.18 Comparison Results set no.18 ............................................................ 58
Figure 5.19 Comparison Results set no.19 ............................................................ 58
Figure 5.20 Comparison Results set no.20 ............................................................ 58
Figure 5.21 Comparison Results set no.21 ............................................................ 59
Figure 5.22 Comparison Results set no.22 ............................................................ 59
1 Introduction

1.1 Overview of computer vision

Computer vision, as a relatively new discipline, has the goal to enable computers to “see”. The focused study of computer vision started in 1970 and it is still being investigated today. Computer vision is regarded as one of the branches of artificial intelligence. Artificial intelligence intends to simulate human behaviour in such a way that computer systems become capable of performing functions that normally require human intelligence e.g. reasoning, problem solving and learning from experience. Artificial intelligence researches combine the elements of computer science and cognitive psychology. Because of the difficulty of cognitive psychology and human intelligence, most of the times computation stream is an alternative which makes the machine to look intelligent. In computer vision the aim is developing artificial vision systems that simulate human vision.

Computer vision is a multidisciplinary research field and has been overlapping with other fields such as computer graphics, image processing, pattern recognition, and photogrammetry. These fields have significant techniques and applications in common, while more briefly computer graphics deals with creating images, image processing concerns low level processing, pattern recognition extracts information from signals mainly based on statistical approaches and photogrammetry is obtaining highly accurate measurement using photographic images.

A growing number of applications exist for computer vision. One prominent usage is in medical image analysis where data is in the form of x-ray, ultrasonic images, Magnetic resonance imaging (MRI), etc. Some
examples are detecting malign changes in samples or measuring the organ dimensions. Military fields have been vastly improved by computer vision advances. Missile guidance and autonomous vehicles are two instances for this application. Also, computer vision is highly used in industry for supporting a manufacturing process. An example could be automatic quality control. Some of the other applications of computer vision are robotic, surveillance and security, image data bases, virtual reality, view synthesis and so on.

1.2 Definition of computer vision

Forsyth [Forsyth, 2003] described computer vision term as “Extracting descriptions of the world from pictures or sequences of pictures”.

More detailed, computer vision can be defined as the study of enabling the computers to acquire visual information, interpret this information and act in response to this information. Computer vision studies first deal with what kind of information is appropriate for the system and should be captured from the input data i.e. from images. Second, should find out how to extract this information. Next, come across what is the most proper way to represent this information. At last decide how to use this information in a system to perform its task [Faugeras, 1993].

1.3 Motivations of the thesis

Segmentation is one of the sub domains of computer vision which has been the subject of numerous researches. In object segmentation the main purpose is to distinguish between the objects of interest and the rest of the image. Most of the existing methods do this task while the object is located on a non-textured noiseless background. In the other words, most
of techniques, as well as active contours assume that background has uniform intensity. Figure 1.1 shows an example of an object on a uniform background.

![Figure 1.1 object on a uniform background](image)

Since the gradient in background of figure 1.1 is almost zero so the contours are attracted to the edges of the flower without any problem.

One main problem in this field is the presence of noise and texture outside the object. This means that, if the background is not clear enough, the problem of object segmentation becomes more difficult. As we can see in figure 1.2, the background has noise. The contour will be deviated since it there are high gradient values in some regions other than the objects borders.

![Figure 1.2 Object on a noisy background](image)
Figure 1.3 shows an example on a textured background. Same problem may occur here. The active contour stops in the textured area before reaching the salient object.

![Figure 1.3 Object on a textured background](image)

This thesis proposes a new technique, based on active contours, that can detect objects' boundaries even when the background has noise and/or texture. The proposed active contour model is aimed at providing robust segmentation results for complicated cases with non-uniform backgrounds. They are also applicable to any image segmentation problem with clear and uniform background.

### 1.4 Overview of the thesis

The structure of the thesis is as follows: first image segmentation in literature is studied. Different types of segmentation methods are categorized in four groups: thresholding techniques, edge-based techniques, region-based techniques and hybrid ones. In each group main contributions are introduced and explained. The Third chapter discusses active contours which are curves that deform and move toward the objects' boundaries. In the context of active contours, snakes and level set methods are studied. Chapter four represents our new variational level set method by using active contours. The last chapter
contains the experiments that show the robustness of the thesis proposed method.
2 Image segmentation

Objects need to be separated from the rest of the image. This is the first step in image analysis which is the task of “image segmentation”. Image segmentation is a long standing problem in computer vision. Segmentation means organizing image content into semantically related groups which are connected and homogenous. Some of the practical applications of image segmentation are:

- Medical Imaging
  - Locate tumors and other pathologies
  - Measure tissue volumes
  - Computer-guided surgery
  - Diagnosis
  - Treatment planning
  - Study of anatomical structure
- Locate objects in satellite images (roads, forests, etc.)
- Face recognition
- Fingerprint recognition
- Traffic control systems
- Brake light detection

The result of image segmentation- the description of these objects- will be used later in object representation and in feature measurement process.

In this chapter after a brief description of image segmentation and investigating “good” segmentation, classification of image segmentation is studied. This classification is grouped in four main categories: thresholding, edge-based, region-based and hybrid methods.
2.1 Definition

Formal definition for segmentation is [Horowitz, 1976]:

Segmentation of a grid $X$ into $X_1, X_2, \ldots, X_n$ subsets must satisfy the following conditions, where $P(R_i)$ is a uniformity predicate for all elements in set $R$:

- $\bigcup_{i=1}^{n} X_i = X$
- For $i$ and $j$, if $i \neq j$, $X_i \cap X_j = \emptyset$
- $P(X_i) = \text{TRUE}$ for all $i$
- $P(X_i \cup X_j) = \text{FALSE}$ if $i \neq j$

The first condition implies that the collection of all the segments will make the whole image. The second condition shows that two different segments should not overlap. The third condition points out that the pixels in one segment have the same properties all over the segment, and the last condition presents that two different segments have dissimilar properties.

Segmentation methods may use this definition or a variation of this. However, there are cases that all these definitions are not enforced in the algorithms.

2.2 What is a good segmentation?

There are exponentially many possibilities to partition an image (figure 2.1). There is no single answer to the question: “What is a good segmentation?” It depends on what application we want to use segmentation in and what information we already have.
2.3 Classification of algorithms

By introducing the first edge detector, the image was officially decomposed into its components. This could be considered as the first image segmentation technique. The first edge detector, Roberts [Roberts, 1963], worked based on applying a $2 \times 2$ filter in sequence. The input for this operator was a gray scale image and in the output each pixel was showing estimated magnitude of the gradient. Although this detector is fast but it is very sensitive to noise. A great number of methods have emerged in segmentation since 1963, and still this topic is considered as a challenging research topic.

There are many developed algorithms and there is more than one classification of these algorithms. Image segmentation methods can be split into many groups based on what classification strategy is used. We
use the traditional classification which classifies the image segmentation methods into four categories: thresholding methods, edge-based methods, region-based methods and hybrid methods that integrate-edge and region-based ones.

2.3.1 Thresholding

Among all the techniques in image segmentation, thresholding is conceptually the simplest approach that we can take to separate the objects from the background. Thresholding convert an image into a binary image based on a threshold value ‘T’. Pixels are going to be marked as object pixels if their intensity is greater than ‘T’. Otherwise, if their intensity is less than ‘T’, they will be set as background. Thresholding works well when objects and background have dissimilar intensity distributions.

The goal in thresholding algorithms is to find an ideal threshold value for getting the best segmented image. Thresholds can be adjusted either manually or automatically. In manual threshold selection, a person should give comments whether the threshold value is correct enough or not. This process is time-consuming and objective. Errors may happen in selection of ‘T’ and later more problems are produced in image analysis. For this reason, a great number of methods have been introduced to automate the thresholding process like [Melgani, 2006] [Sezgin, 2004] [Sauvola, 2000]. We are going to review some of the automatic methods in followings.

One of the simplest thresholding methods is to find peaks and valleys in histogram and set the threshold according to them. This approach is robust because peaks can be easily found since their shape is well defined. Problem occurs when the histogram is noisy and false minimum
and maximum are detectable. To avoid false peaks and valleys, the histogram usually is smoothed. One algorithm based on peaks detection is [Sezan, 1985] which at first finds zero-crossings and later uses a peak detection signal to estimate significant peaks. One advantage of this method is that we can adjust how fine the peaks are going to be. In [Boukharouba, 1985] a variation of this method is used where the cumulative distribution function of the image is first extended; pursued by the curvature examination. More related methods can be found in [Tsai, 1995] [Carlotto, 1987] [Olivo, 1994].

Rosenfeld [Rosenfeld, 1983] finds the optimal threshold through an analysis of the convex deficiency which is calculating from deducting the histogram from its convex hull. Some other modifications of this idea are [Whatmough, 1991] [Sahasrabudhe, 1992].

One famous thresholding algorithm is Otsu’s [Otsu, 1979] method. In this automatic unsupervised thresholding algorithm, distribution of pixels value is analyzed. The basic idea here is that the pixels in each class (cluster) should be as similar as possible. This means that the variance inside each class should be minimized. Otsu defined the within-class variance as the weighted sum of the variances of each cluster. Because calculating within-class variance is expensive, we can use between-class variance in computation instead. Minimizing within-class variance is equivalent with maximizing between-class variance. Otsu’s method is most widely employed in literature and its result is robust and satisfactory. Many papers use Otsu’s method with slight modifications. Two-Stage Multi-threshold Otsu method is introduced in [Huang, 2009] developed to overcome the inefficiency of original Otsu by reducing the iterations so that, the computation time will be much more less. In [Wu, 2008] authors claim that the variation of Otsu’s algorithm is more
feasible and faster also, resulted better segments in image. More applications can be found in [Wang, 2007] [Ali, 2004] [Junwei, 2007].

The Iterative Self-Organizing Data Analysis Technique (ISODATA) [Ridler, 1978] was developed by Ridler and Calvard. Like Otsu, they use the means of foreground and background but here search for the optimum is locally, unlike in Otsu which is global. The algorithm starts by an initial threshold which is equal to the half of the maximum gray level. Binary image is generated based on this threshold. The mean value then is computed for the current background and foreground. The new threshold is then replaced with the average of the calculated means. This process is repeated until the threshold reaches convergence. This method is quite popular and does not need initial training for segmentation process. However, convergence in short period of time sometimes is inaccessible. DYNO (Dynamic Optimal Cluster seeking) algorithm [Tou, 1979] is a modification of ISODATA, with the ability of split and merge. In [Jiang, 2008], they employ ISODATA technique for building extraction in remote sensing. In medical image processing [Ding, 2004] Self-Organizing Data Analysis Technique Algorithm was applied to detect tissue damage in MRI images.

Niblack [Niblack, 1986] introduced a local thresholding algorithm which moves a window across the image and calculates the local mean and local deviation for the center of window each time. He achieved the threshold value by developing a function of the mean and the standard deviation of the neighbourhood. Trier and Jain [Trier, 1995] tested Niblack’s method and proved that it has best results comparing to any other local thresholding methods. In [Sauvola, 2000], a heuristic variation of Niblack’s formula is used in which the standard deviation is normalized. Niblack’s method has the difficulty of choosing the right
window size. The size of the window should be set so that it saves the local details and suppresses noises. In [Zhao, 2008] Niblack’s method is used in video text processing.

Some algorithms for thresholding are based on information theoretic approach which was introduced by [Pun, 1980]. The entropy-based techniques have proven to be successful and convincingly robust [Kapur, 1985] [Chang, 1994] [Luthon, 2004]. These techniques rely on maximising the total entropy of both the object and background regions to discover the suitable threshold. Some of these methods also use the pixels' spatial information. More examples are in [Chang, 2006] [Sezgin, 2004].

Although Gray level thresholding is the simple, easy to grasp and fast, but threshold selection is not always straightforward. The major drawback of threshold-based approaches is that they often fail to find the best separation between true positive signals and false positive signals (noise), e.g. if the threshold is kept too low, a lot of true positive signals maybe are not detected. Another alternative way to segment the image is edge-based segmentation.

2.3.2 Edge-based segmentation

Edge-based segmentation approaches relies on the edges found in images. These edges represent the location of the discontinuities in gray level, colour, texture, etc. Edge-based methods are based on the information previously achieved by the edges in the image.

A variety of edge detector operators, which usually are named after their inventors, exist in literature. The most famous ones are Prewitt [Prewitt, 1970], Sobel [Sobel, 1978], Laplacian [Pratt, 1991], Canny [Canny, 1983]. Because edges are not always connected and are not
always showing the objects boundaries, the images which are resulted from edge detection could not be an appropriate segmentation result by itself. Therefore, edge detection is regarded as a prepossessing step. The goal is to connect the relevant edges in such a way that the object boundaries are produced.

There exist many methods which use different approaches to locate the objects borders. These methods also use different quantity of former information. The more former information is available, the better the segmentation results will be. Otherwise, if there is not enough information about boundaries, methods should employ more local information about the image.

Sometimes some small edges appear in images, because of noise or illumination changes. One group of methods try to use threshold to eliminate these edges. The original idea is from Kundu and Mitra [Kundu, 1987]. Finding a global threshold that works all over the image is not achievable. Edges are usually thick as well. For the solution, non-maximum suppression and hysteresis thresholding can be used as it was introduced by the Canny [Canny, 1983] edge detector. Non-maximum process checks if each pixel is local maximum along gradient direction. Then suppress the points which are not local maximum. Hysteresis checks if that maximum value of gradient value is large enough. If the gradient of a pixel is above the high threshold, it will be declared as edge pixel. Otherwise, if the gradient of a pixel is less than the low threshold, it would be non-edge pixel. Any value between these two ranges is going to be edge pixel if it is connected to any edge pixel. These papers use the same ideas in image segmentation [Liu, 2000] [Tefera, 2002].

Using thresholds in finding edges usually ends in noisy results, and sometimes some important edges are missing. Edge relaxation method
considers the edges properties’ in the context of both ends of edges. Local edge strength is raised if there are adequate evidences that borders may exist. Studying the context here, means investing the local neighbourhood of the edge. Weak edges, which are located between two strong edges, are considered as a boundary. An isolated edge, even strong one, without a supporting context will not be considered as border. Hanson et al. [Hanson, 1978] introduced a conventional edge context assessment. Later Prager [Prager, 1980] modified his algorithm. According to Prager’s method, three groups of edge patterns exist, which cause the confidence in an edge to be modified: patterns in which the confidence of edge would be increased, decreased or remain unchanged. The initial confidence of an edge will be set as the normalized gradient value. Then, “edge type” would be recognized based on the confidence of edge neighbours. At last, the confidence of the edge will be modified based on previous confidence and its type. This process will be repeated until the confidence value converges to either 0 or 1. This method is comparatively simple and noise robust. However, it often slowly flows and after larger numbers of iterations, giving worse results than expected. In [Sher, 1992] another approach, which is using probabilistic distribution of edge neighbourhood, is presented. More recent applications of edge relaxation can be found in [Czuni, 2001] [Moro, 2008].

Another edge-based technique is called boundary tracing. This technique is performable after the image is over segmented; means that the background and the foreground are already separated. Tracing inner and outer boundary is part of this algorithm. Inner region border is part of the region but outer border is not. This definition indicates that two adjacent regions do not have a common border. To overcome this deficiency, Pavlidis [Pavlidis, 1977] proposed extended borders as a
hybrid technique. Extended borders utilize inner borders for the upper and left sides of the object and outer borders for the lower and right sides. Extended borders specify a common border between adjacent regions. A more advanced method for extended boundary tracing was developed in [Liow, 1991].

One way to connect edge segments is to trace from pixel to pixel through potential edge points. Decision for every edge pixel is based on the neighbour pixel gradient value and gradient orientation. Local edge linking methods usually start at some arbitrary edge point and then observe the points in neighbourhood. Edge linking is regularly followed by post processing. Farag [Farag, 1991] detected the contours in two stages: edge enhancement followed by edge linking. In [Gao, 1999] for object extraction a low-complexity edge-linking algorithm in colour images is designed. For finding global edges Hough transform, [Hough, 1962], named after Paul Hough, is a good option that decides which tokens belong to which objects. Here the input for Hough transform is a set of ‘n’ edge points, which are found formerly by an edge detector, and the output is all the lines which these edge points are placed on. Instead of x-y plane, each line is represented in a-b plane which is the slope and the intercept of lines. All points, that lie on a line ‘S’ in x-y plane, have lines in parameter space that intersect at the ‘a1’, ‘slope’, and ‘b1’, Intercept of the line S. Later Duda et al. [Duda 1972] showed a more efficient method by using polar parameterization. Generalized Hough transform introduced in [Merlin, 1975] to find arbitrary shapes with known orientation and scale. Generalized Hough transform with arbitrary orientation and scale developed in [Ballard, 1981]. Hough transform has imperfections including lack of accuracy and misleading results when objects happen to be aligned by chance. Random Hough transform [Xu,
1993] tries to find a solution to solve the mentioned problems. Mapping form image space to parameter space is replaced with converging mapping to improve time complexity and accuracy. Some Applications to the randomized Hough transform are in [Behrens, 2003] [Ding, 2005] [Jean, 2004] [Xu, 2007].

Although edge-based methods produce clean and well defined boundaries between different regions, they are likely to produce gaps between boundaries. The necessity of complicated post-processing is considered as one of the problems in this area. Another approach in segmentation is region-based segmentation that comes in follow.

2.3.3 Region-based segmentation

When we segment the image by judging only on the gray value of pixels, the pixels are grouped into objects and taking no account of connectivity property. In other words, pixels are classified independently of the context. In region-based segmentation, uniformity within a sub-region is the main issue, unlike the edge-based segmentation that discontinuity is the main concern. The uniformity may be based on different properties e.g. intensity, colour and texture. Based on the chosen property, the complexity, the form and the quantity of former information vary in segmentation method. Comparing to edge-based methods, in region-based methods more coherent regions are created. However, judgment over region membership is harder than applying edge detectors.

A simple approach in region-based segmentation is region growing. The central idea in region growing is to start from a single pixel and grow into a coherent region. The starting pixel is called the seed pixel. A similarity measure is used for comparing every other pixel to the seed. Then new pixels are added to the region if similarity measure is satisfied.
Various definitions exist for describing the similarity measure for instance using pixel’s intensity value or average intensity value. The Comparing stage could also be done in several ways, sometimes the seed is the only reference. This makes the region very sensitive to the seed selection. If every new pixel is compared with its neighbours, sensitivity to seed is removed. But region growing will become so slow and results may be far away from the original pixel. Another comparing candidate is region statistics i.e. region mean, variance, etc. One seeded algorithm in [Adams, 1994] works as follow: checks to see if a pixel touches only on region by checking all the neighbours having the same label. If so, the similarity measure between the new pixel and the region is computed. Otherwise, if the new pixel touches more than one region, the similarity to all the regions is calculated and smallest one is selected. After the similarity is retained for the pixel, the pixel is put in sequentially sorted list (SSL). This list is ordered according to similarity attribute. In this algorithm, pixels having not same labelled neighbours are labelled as boundary pixels. 3D extension of this algorithm is employed in [Justice, 1997]. An approach using seeded region growing with effective pixel labelling technique and automatic seed selection process is introduced in [Fan, 2005]. Latest applications of seeded growing segmentation are in [Wu, 2008] and [Gomez, 2007]. In region growing (merging) if the comparisons are based on fine details, it will be computationally expensive. And also, final outputs depend on seed points and search strategy.

One other viewpoint for region-based segmentation is region splitting. Contrary to region growing, the method starts with the whole image as a single region then, splits into sub regions based on homogeneity criteria. Although region splitting is the opposite of the region growing but, their
results are not the same even if both use a same similarity criterion. An early work using region splitting is in [Ohlander, 1978] where splitting of inhomogeneous regions is used to divide recursively the entire image until homogeneous regions are found. In [Shulman, 2004] recursive region splitting is used for evaluation of a single scene by testing statistical homogeneity criteria after each split. If homogeneity has got better, the split is accepted otherwise the split is undone.

The main flaw in region splitting is that the sub regions may have adjacent regions with similar properties. Solution for this problem suggests using split and merge together [Horowitz, 1976]. It is achievable to take advantage of these two methods by combination them. First the entire image is supposed as one region if, it is not homogenous it splits to sub regions. Each sub region is checked iteratively and is divided if, it is not homogenous. At the end these adjacent regions with same properties merge [Fukada, 1980] [Chen 1980]. Split-and-merge method is more efficient than split or merge. An adaptive split-and-merge method and a review of region homogeneity testing are in [Chen, 1991]. Diamand et al. [Diamand, 2003] extended the algorithm to 3D case images. They used topological maps for the representation of segmentation states and in split and merge process. In [Zhan, 2006], for detecting text on colour images split-and-merge segmentation is used after a pre-processing enhancement. For locating a diagnostic tumour from ultrasound images, a split-and-merge technique is employed [Kwak, 2003].

A drawback of algorithms in this group is that in general they create distorted boundaries since the segmentation typically is carried out at region level instead of pixel level. Next hybrid methods are studied.
2.3.4 Hybrid methods

Examining the segmentation results of both edge-based and region-based techniques leads to the conclusion that either edge-based or region-based segmentation fails to produce accurate segments. As mentioned in [Salotti, 1992], both approaches usually suffer from lack of information for segmentation. Because of the segmentation problems in complex images, using only one of these techniques will not lead to satisfactory results. Integrating both approaches looks like a good solution. Yet, achieving this goal is not easy because region-based and edge-based segmentation are based on different ideas.

Time of fusion is one main property of hybrid methods. Considering that the hybrid algorithms are grouped into embedded integration and post processing integration [Munoz, 2003]. As it is obvious from the names, in embedded segmentation, an edge-based operator segments the image first then, the output information is used in a region-based segmentation or, a region-based operator segments the image first and then, the results are used in edge-based segmentation. But, in post processing method both edge-based method and region-based method are processing the image independently. Afterwards, all the output information is used in a posterior fusion step.

The most usual way in embedded segmentation is integrating of edge information with region-based segmentation during the decision making in region growing procedure. In [Bonnin, 1989], plus the homogeneity criterion, the edge information is also considered during split and merge process. When there is no edge pixel in the regions and adjacent regions are homogenous, the region grows. Similarly, [Healey, 1992] employs the absence of edge pixels as a homogeneity criterion in 3D scenes. Besides, he claims the low edge threshold for edge detectors will increase the
accuracy since false negative results from edge detection have serious consequences on segmentation. In [Lewis, 2002], edge information is used as a decisive factor for the split and merge during sonar images processing.

One kind of post processing hybrid methods is over-segmentation. This method is about finding all the possible segments by strict region-based segmentation. At the same time, all the edges are found by edge-based segmentation. The results of region-based method are checked with edges to find out whether they are real boundaries or not. If there is no correspondence for each boundary, it will be removed. Examples of this type of hybrid segmentation are in [Pavlidis, 1990] and [Gagalowicz, 1986]. Another strategy for getting over-segmented image is to start with one boundary detection technique to over segment the image. Then the boundaries are verified by analyzing the chromatic and textural attributes on each side of the contour. If the attributes are different on sides then, the boundary is valid. This approach is used in [Philipp, 1996] and [Fjortoft, 1997]. A More current case of over-segmentation is in [Guo, 2005] where the gradient is used to find the correct boundary of the over-segmented image to prevent from the merging dissimilar regions.

In addition, post processing segmentation is a way for finding the best approach in image segmentation in the absence of ground truth data. Defining an appropriate stopping condition or setting suitable thresholds in region segmentation were some issues in traditional region-based methods. These problems can be solved using the evaluation function which measures the degree of the excellence of a region-based segmentation in line with its consistency with the edge map. If the region boundaries are corresponding most closely to the contours, the region
segmentation is selected as the best one. Examples are in [Revol-Muller, 2000] and in [Hojjatoleslami, 1998].

2.4 Summary

Image segmentation partitions the image into semantically related groups which are homogenous and connected. There are exponentially many possibilities to segment an image and there are a lot of options for getting correct segmentation. Based on prior information we have and kind of the application we want to use the results in, the approach may differ.

Generally segmentation methods can be categorized as thresholding, edge-based segmentation, region-based segmentation and hybrid segmentation which is the integration of the both edge and region segmentation techniques. Thresholding is the simplest image segmentation method. A constant called a threshold is employed to segment objects and background. In Edge-based segmentation, edges that found in an image are the basis for segmentation. On the other hand in region- based segmentation the homogeneity of the region is the main issue. By integrating edge and region information hybrid segmentation gives better results. That is the reason why some people use hybrid methods instead of choosing one of the image segmentation techniques.


3 Active contours

Studying and using active contours have leaded to promising results in context of segmentation. As the methods discussed in previous chapter are not fully capable of segmenting objects boundaries, active contours are introduced as a solution. This approach is based on using deformable contours that move under the influence of forces and are used to track boundaries and motions. The idea of using a deformable pattern for selecting particular features in image are introduced in [Widrow, 1973] and in [Fischler, 1973] for the first time. However, it was not until the work of [Kass, 1987] that the active contours became famous. The goal is to find the equation that will drive the contour to the object. In other words the curve should evolve until its boundary segments the object of interest.

There are two deformable models: parametric models (snakes) and geometric models (level sets). In parametric active contours, curves are presented explicitly during deformation. On the other hand, in level sets contours are shown as implicit level of functions which are based on curve evolution and level set method.

In the following sections snakes are studied. Then a detailed review of level set method is presented. Examples in literature for both techniques are also introduced.

3.1 Snakes

The earliest and most famous active contour method is introduced by kass [Kass, 1987]. Kass named his algorithm “snakes” because during the evolution, the contours motion toward the object resembles snakes’ movement. Given an approximation of the boundary of an object in an
image, called initial contour, snakes locate the “actual” boundary. Let us define a contour parameterized by arc length s as

\[ C(s) = \{(x(s), y(s)) : 0 \leq s \leq L\}, R \to \Omega \]

(3.1)

Where L denotes the length of the contour C, and \( \Omega \) denotes the entire domain of an image \( I(x, y) \). This algorithm is based on energy minimization scheme. The basic idea of energy minimization is minimizing the weighted sum of the internal energy, which depends on the shape of the contour i.e. smoothness of the contour, and the external energy which depends on image properties i.e. gradient.

\[ E(c) = E_{\text{int}} + E_{\text{ext}} \]

(3.2)

Minimizing the total energy yields internal forces and external forces. Internal forces keep the curve together and prevent it from bending too much. External forces draw the curve toward the desired object boundaries. Each point is moved to the place of the minimum value in \( E_{\text{E}} \) (figure 3.1). If the energy functions are chosen precisely, the contour, should move towards, and stop at, the object boundary.

![Deformable contour](image)

Figure 3.1 Active contour’s movement
If $\alpha$ controls the tension of the contour and $\beta$ controls the rigidity of the contour, a common option for internal energy function will be:

$$E_{int} = \int_{0}^{L} \left( \alpha \left| C'(s) \right|^2 + \beta \left| C'(s) \right| \right) ds$$

(3.3)

The external energy function attracts the deformable contour to interesting features, such as object boundaries, in an image.

$$E_{ext} = \int E_{img}(C(c)) ds$$

(3.4)

A common option for edge attraction function is a function of image gradient:

$$E_{img}(x, y) = \frac{1}{\lambda \left| \nabla G\sigma^* I(x, y) \right|} \Omega \rightarrow R$$

(3.5)

Where $G\sigma$ is a Gaussian smoothing filter and $\sigma$ is its standard deviation. Also $\lambda$ is a proper constant. To find the object boundary, parametric curves are initialized within the image and are moved toward the energy minima under the influence of both these forces. That is why [He, 2008] refers to original snake as an interactive method which needs expert guidance on the snake initialization and the choice of accurate deformation parameters.

The classical snake limitations motivated other snake variations to be introduced [Amini, 1990], [Cohen, 1991], [Zhu, 1996], [Xu, 1998], [Giraldi, 2000], [McInerney, 2000], [Fenster, 2001], [Delingette, 2001], and so on. Some of these limitations are as follows: since the magnitude of the external force vanishes quickly as the contour diverges to the boundaries and this makes the capture range of the snakes very small, the classical snakes provide an exact position of the edges only if the initial
contour is specified satisfactorily near the edges. Approximating a proper location of initial contours without previous knowledge is usually a tricky problem. In addition, snakes are very sensitive to the noises in the image can be easily distracted to wrong places. These were some drawbacks of the original snakes but the most important flaw of classical snake is its inability to adaptation of the model topology during the deformation. In other words, snakes maintain the same topology during the evolution stage. That is, snakes cannot split to multiple boundaries or merge from multiple initial contours.

In [Cohen, 1991] and [Xu, 1998] the focus is on reducing the dependency on initial conditions by defining new external energy for improving the snakes’ algorithm. Cohen et al. [Cohen, 1991] proposed a new snake, called Balloon snake, and added a second external force which shifts the contour out (inflation) or in (deflation) along its normal. The new defined snake has resemblance to a balloon being inflated in 2D. The balloon force enables the snake to be initialized inside the object in addition to remove the necessity for the initial curve to be close to the real edges. Comparing to the original snake by Kass [Kass, 1987], balloon snake passes over relatively weak edges so has more stable results. Also, if the object has a problematic shape, insertion a balloon inside the shape and expanding its contour will locate the desired shape. However, balloon snake has the problem with forcing the snake into concavities. Xu and Prince [Xu, 1998] introduced gradient vector flow (GVF) snake which increases the capture range and improve the snakes ability to move into boundary concavities. It still has difficulties, however, forcing a snake into lengthy, thin boundary indentations. Like the Cohen’s [Cohen, 1991], Xu’s [Xu, 1998] method enlarged the capture range from object
vicinity. GVF s are vector fields derived from images by minimizing an energy functional in a variational framework.

Both [McInerney, 2000] and [Delingette, 2001] try to help the topology change during the evolution. In [McInerney, 2000] topology-adaptive snakes or T-snakes are introduced for medical image segmentation based on an affine cell image decomposition (ACID) framework. The ACID provides a method for contour re-parameterization and that allows T-snakes to split or merge; in other words adapting to the topology of object. Yet only specific motions- inflating or deflating- is applicable. Also re-parameterization is making the method complex and expensive particularly in 3D. New physical constraints are introduced in [Delingette, 2001] to control the contour deformation. In addition to ability of splitting, the contours are kept separated by removing overlapped snake areas. So that handling the topological changes plus restoring contour separations as a procedure reduces the likelihood that contours converge toward the boundaries of other object, are Delingette’s [Delingette, 2001] method properties.

In [Fenster, 2001] first the shape of the object is achieved then the shape information is used as constraint. The Contour evolves considering this constraint so that the contour will not be captured by fake edges. Region based features are used in [Zhu, 1996]. In this method boundary deformation and region merging are done iteratively. Region based information to the accompaniment of edge based data overcome the noise in the image. But the problem in splitting the contour in multiple contours still exists. Amini el al. [Amini, 1990] and Giraldi et al. [Giraldi, 2000] used a dynamic programming approach instead of Kass’s [Kass, 1987] variational method for minimizing the energy functional in snakes. The
estimation of higher order derivatives are omitted here and the numerical stability is improved.

3.2 Level set methods

Geometric deformable models provide an elegant solution to address the primary limitations of parametric deformable models. These models are based on curve evolution theory and the level set method.

Level set method was first introduced in [Dervieux, 1980] and then devised by Osher and Sethian [Osher, 1988]. For capturing moving fronts in a wide range of problems, level set method has shown to be a robust numerical option. Some fields using level set techniques are image processing, computer vision and graphics. As mentioned by Tsai [Tsai, 2003], an implicit data representation of a hypersurface, set of PDEs that govern how the surface moves, and the corresponding numerical methods for implementing this on computers are building components of classical level set method.

3.2.1 Level set concept

The main idea of the Level set method can be described as follows. In an open region $\Omega$, $\Gamma$ is a closed interface evolving with the velocity $v$. The goal is to analyze and compute the motion of the interface. Osher and Sethian’s idea is to define an implicit smooth (Lipschitz continuous) function $\phi(x, t)$ which represents the interface as the set where:

$$
\phi(x, t) = 0 \text{ if } x \in \Gamma \\
\phi(x, t) < 0 \text{ if } x \in \Gamma_{in} \\
\phi(x, t) > 0 \text{ if } x \in \Gamma_{out}
$$

(3.6)
Where $\Gamma_{in}$ shows the area inside the interface and $\Gamma_{out}$ shows the area outside (figure 3.2).

![Figure 3.2 The construction of level set function](image)

The evolution could be described by convecting the $\phi$ with the velocity field $v$ on the interface:

$$\frac{\partial \phi}{\partial t} + v \nabla \phi = 0$$  (3.7)

If the normal component of $v$ is $v_N = v \frac{\nabla \phi}{|\nabla \phi|}$, where $|\nabla \phi| = \sqrt{\sum_{i=1}^{n} \phi_{ni}^2}$, the equation (2.1) can be written using normal velocity:

$$\frac{\partial \phi}{\partial t} + v_N |\nabla \phi| = 0$$  (3.8)

These equations are Hamilton-Jacobi equations so that with suitable restrictions the theory of viscousy solutions [Crandall, 1983] picks out unique Lipschitz continuous solution.

A useful property of this approach is that the level set function remains a valid function while the embedded curve can change its topology. Geometric active contours have many advantages over
parametric active contours, such as computational simplicity and the ability to change curve topology. Unlike the snake can start far from the boundary and will converge to boundary concavities.

### 3.2.2 Level set dictionary and technology

Key terms and some key technological advances in level set methods are [Osher, 2003]:

1. The interface boundary $\Gamma(t)$ is defined by $\{ x | \phi(x, t) = 0 \}$. The region is bounded and its exterior is defined by $\{ x | \phi(x, t) > 0 \}$.

2. The unit normal $N$ to $\Gamma(t)$ is

   \[ N = \]  

3. The mean curvature $\kappa$ of $\Gamma(t)$ is defined by

   \[ \kappa = -\nabla \cdot \left( \frac{\nabla \phi}{|\nabla \phi|} \right) \]

4. The Dirac delta function concentrated on an interface is

   \[ \delta(\phi)|\nabla \phi| \]

   where $\delta(\phi)$ is a one-dimensional delta function.

5. The characteristic function $\chi$ of a region $\Omega(t)$ is

   \[ \chi = H(-\phi) \]

   where $H$ is a one-dimensional Heaviside function and

   \[ H(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0 \end{cases} \]
6. The surface (or line) integral of a function \( f \) over \( \Gamma \) is

\[
\int_{\Gamma} f(x) \delta(\Phi) \|\nabla \Phi\| \, dx
\]

and The volume (or area) integral of \( f \) over \( \Omega \) is

\[
\int_{\Omega} f(x) H(\Phi) \, dx
\]

7. In many cases, \( \phi \) will develop steep or flat gradients which cause problems in numerical approximations. For preventing \( \phi \) from becoming too flat or too steep near the interface as well as keeping the zero location unchanged, the distance reinitialization [Sussman, 1994] procedure reshapes a general level set function \( \phi(x, t) \) by \( d(x, t) \) which is the value of the distance from \( x \) to \( \Gamma(t) \), positive outside, and negative inside.

Let \( d(x, t) \) be signed distance of \( x \) to the closest point on \( \Gamma \). The quantity \( d(x, t) \) satisfies

\[
\|\nabla d\| = 1, \quad d \text{ is positive outside and negative inside and also is the steady state solution to}
\]

\[
\phi_t + \text{sgn}(\phi_t) \|\nabla \phi\| - 1 = 0, \phi(x, \tau = 0) = \phi_0(x)
\]  \hspace{1cm} (3.9)

Here \( \phi_0 \) shows the level set function before the reinitialization. For most applications, the reinitialization is only needed for a neighbourhood around the zero level set, and the diameter of this neighbourhood depends on the discretization of the partial derivatives in the PDE. This implies that only a few time steps in \( \tau \) are needed.

8. The basic level set method concerns a function \( \phi \) which is defined all over space. Obviously this is wasteful unless one only cares about information near the zero level set. The local level set method defines
\( \phi \) only near the zero level set. We may solve (3.7) in a neighbourhood of \( \Gamma \) of width \( m \Delta x \), where \( m \) is usually 5 or 6. Points outside this neighbourhood need not be updated by this motion. Thus, this local method works easily in the existence of topological changes and for multiphase flow.

3.2.3 Numerics

Eq. (3.8) is Hamilton-Jacobi equation when normal velocity is dependant of \( x, t \) and \( \nabla \phi \), Numerical methods should be used on uniform Cartesian grid because of existence of singularities in solutions. The key ideas involve monotonicity, upwind differencing, essentially non-oscillatory (ENO) schemes, and weighted essentially non-oscillatory (WENO) schemes [Osher, 1991] [Osher, 1988] [Jiang, 2000].

3.2.4 Level set in segmentation

As mentioned before, level set method has been widely used because it lets the contour to fit in angles, corners and topological changes. A special case of the motion of the contour is based on mean curvature and \( v \) is calculated with curvature of the curve. A basic version of the speed functions that combine curvature and constant deformation were proposed in [Caselles, 1993] and [Malladi, 1995]. A famous active contour model based on mean curvature is introduced in [Caselles, 1993] using the flowing equations:

\[
\frac{\partial \phi}{\partial t} = g(\nabla u_0) \nabla \phi
\]  

(3.10)

Where \( \alpha \) is a constant pushing the curve when curvature becomes null or negative or also inside the curve increasing the speed. The
curve moves with the speed \( g(|\nabla u_0|)(div \left( \frac{\nabla \phi}{|\nabla \phi|} \right)) \). And \( g(|\nabla \phi|) \) is an edge dependent function so that the contour stops at desired boundary where \( g \) disappears. Another formula for finding the zero level sets are proposed in [Malladi, 1993].

\[
\frac{\partial \phi}{\partial t} = |\nabla \phi| (\alpha + \frac{\alpha}{M_1 - M_2} (|\nabla G_2 * u_0| - 1) - I)
\]  

(3.11)

Again \( \alpha \) is a constant. \( M_1 \) and \( M_2 \) are minimum and maximum value of the magnitude of the gradient \(|\nabla G_2|\). When the speed vanishes the evolving contour will stop and this happens at the highest gradients. Later Caselles [Caselles, 1997] proved that the minimization of the contour energy is even to the minimization of the contour length weighted by an edge detection function in the Riemannian space. He integrated the curve evolution methods with the classical energy minimization methods (snakes). Other speed functions for evolving curves can be found in [Siddiqi, 1998]. Often in level set methods the initial level set function is frequently based on the signed distance. An efficient algorithm for building of the signed distance function is called a fast marching method [Malladi, 1996], [Malladi, 1998], [Sethian, 1999]. Applying the constant deformation method may create sharp corners of the zero-level set resultant in a vague normal direction. In that situation, the deformation can be continued using an entropy condition [Sethian, 1982].

In classical geometric models, an evolution PDE for level set function is originated from a certain evolution PDE of a parameterized curve. On the other hand, in variational methods the evolution PDF of the level set function is derived from minimizing the energy function defined on the level set function. Comparing with classical methods, variational methods are more convenient for incorporating additional information.
In Zhao [Zhao, 1996] a variational level set method introduced. Suppose there are disjoint regions $\Omega$ with the boundaries $\Gamma$ so that the common boundary between $\Omega_i$ and $\Omega_j$ is $\Gamma_{i,j}$. Energy function is described as

$$E = E_i + E_2$$  \hspace{1cm} (3.12)

Where $E_1$ is the energy of the interface and $E_2$ is the bulk energy. The normal velocity is positive multiple of curvature of the interface plus the bulk differences. That can be written as

$$E_1 = \sum_{i=1}^{n} \gamma_i \int \delta(\phi_i) |\nabla \phi_i| dx$$

$$E_2 = \sum_{j=1}^{n} \gamma_j \int H(\phi_j) dx$$  \hspace{1cm} (3.13)

Where $H$ is Heaviside Function, $\delta$ is delta function. Now minimizing $E$ is the solution.

Chan et al. [Chan, 2001] came with new variational method without a stopping edge-function, unlike the other level set methods that use gradient value to stop the curve evolution. The original formulation of Chan et al. [Chan, 2001] developed for bimodal images. This was afterwards extended to multiphase images [Chan, 2002]. In the bi-modal model, it is supposed that an image $I$ is formed of two approximately piecewise-constant distinct intensity regions, and . If the region to be segmented is represented by , then a curve $C$ can be evolved to reach the boundary of by minimizing the energy:

$$F_1 + F_2$$

And
The variables $c_1$ and $c_2$ show the average intensities inside, and outside the curve respectively. It can be easily represent that the minimum of the above fitting term is the boundary. If the curve is outside the region then $F_2 < 0$ and $F_1 > 0$ (figure 3.3 a), If the curve is inside the region then $F_1 < 0$ and $F_2 > 0$ (figure 3.3 b), If the curve is both inside and outside the region, then $F_2 < 0$ and $F_1 > 0$ (figure 3.3 c). The only case that the fitting energy is minimized is when the curve is located on the boundary (figure 3.3 d).

Chan and Vese [Chan, 2001] added some terms to the explained function and introduced their energy function as:

$$F(c_1, c_2, \emptyset) = \mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(\text{inside}(C))$$

$$+ \lambda_1 \int_{\text{inside } C} |I - c_1|^2 \, dx \, dy$$

$$+ \lambda_2 \int_{\text{outside } C} |I - c_2|^2 \, dx$$

(3.16)
Here \( \mu > 0, \lambda_1, \lambda_2 \): are fixed parameters. By writing the area and volume in energy form and

\[
\int_{\text{inside}} c |I - c_1|^2 \, dx \, dy = \int_{\Omega} |I - c_1|^2 H(\phi) \, dx \tag{3.17}
\]

And

\[
\int_{\text{outside}} c |I - c_2|^2 \, dx \, dy = \int_{\Omega} |I - c_1|^2 (1 - H(\phi)) \, dx \tag{3.18}
\]

the function becomes:

\[
F(C_1, C_2, \Theta) = \mu \int_{\Omega} c(\phi) |\nabla \phi| \, dx \, dy
\]

\[
\int_{\Omega} H(\phi) \, dx \, dy
\]

\[-\lambda_1 \int_{\Omega} |I - c_1|^2 H(\phi) \, dx \, dy
\]

\[
\lambda_2 \int_{\Omega} |I - c_1|^2 (1 - H(\phi)) \, dx \tag{3.19}
\]

By keeping the \( \phi \) fixed and minimizing the F with respect to \( c_1 \) and \( c_2 \), the values of \( c_1 \) and \( c_2 \) are calculated. Then they regularized \( \delta \) and \( H \) by two smooth functions \( \delta_\varepsilon \) and \( H_\varepsilon \) to use Euler-Lagrange equation:

\[
\partial_\phi F = \delta_\varepsilon (\phi) \mu V \frac{\nabla \phi}{|\nabla \phi|} + \nu + \lambda_1 (u_0 - c_1)^2 - \lambda_2 (u_0 - c_2)^2 = 0 \tag{3.20}
\]

A way to solve this minimization problem is using gradient descent on eq. (3.20) so that \( \partial_\phi \phi = -\partial_\phi F \).

The advantage of the Chan-Vese active Contours is that it is able to segment an image even if it has smooth boundaries. The evolution of the curve does not depend on gradient information; as a result weak edges do not concern the final segmentation. However the Chan-Vese model has
initialization problems. The result of segmentation is dependent on the situation of the initial curve.

The level set function $\phi$ can develop shocks which makes additional computation vastly imprecise. A way to avoid this problem is to initialize the function $\phi$ as a signed distance function, before the evolution and then reshape the function $\phi$ to be a signed distance function periodically during the evolution. In [Li, 2005] by mentioning to reinitialization’s flaws such as the displacement of the zero level set within the reinitialization, an increase of the number of iterations, nonexistence of a known single method for reinitialization, making the computation more expensive and complex; a new formula for geometric active contours using new variational method has been discussed so that the is no need to re-initialize the function. They define an energy function which consists of internal and external energy. In order to keep the level set function as an approximate signed distance function they use special internal energy that penalizes the deviation of the level set function from a signed distance function so that the level set function will be always close to sign distance function. Sign distance function has the property $\|x\|=1$ and any function satisfying $\|x\|=1$ is signed distance function. The internal energy is

$$P(\phi) = \int_\Omega \frac{1}{2} (|\nabla \phi| - 1)^2 \, d\Omega$$  \hspace{1cm} (3.21)

So that it shows how close a function is to its distance function. The final energy function is

$$E(\phi) = E_{\text{true}} + 1 \hspace{1cm} (3.22)$$
\[ E_{\text{int}} = u \int_{\Omega} \frac{1}{2} (|\nabla \varphi| - 1)^2 \, d\Omega \quad (3.23) \]

External energy function uses gradient function; it means that the contour will be attracted to the points where the gradient is high:

\[ E_{\text{ext}} = \lambda \int_{\Omega} g \delta(\varphi) |\nabla \varphi| \, dx \, dy + u \int_{\Omega} gH(-\varphi) \, d\Omega \quad (3.24) \]

\[ g = \frac{1}{1 + |\nabla \varphi|} \quad (3.25) \]

Li’s method [Li, 2005] is computationally efficient, stable results are produced and the most important advantage is omitting the reinitialization process.

The use of level set and PDEs in computer vision has been developed in recent years. In image segmentation many algorithms has utilized the level set method to find “a collection of non-overlapping regions” of a given image. There are a large variety of applications where which geometric deformable models were employed for segmenting the image. Examples include a level set-based cortical unfolding method [Hermosillo, 1999]; cell segmentation [Sarti, 1996] and [Yang, 2005]; cardiac image analysis [Niessen, 1998], [Angelini, 2004], [Lin, 2003]; tumor tracking [Li, 2007], Biomolecular surfaces construction [Bajaj, 2008], and many others.

### 3.3 Summary

In this chapter, we have described the fundamental concepts of both parametric and geometric deformable models and shown that they can be used in finding object boundaries.
Snakes are parametric deformable models. The energy functional which is minimized is a weighted combination of internal and external forces. To find the object boundary, parametric curves are initialized within the image domain, and are forced to move toward the potential energy minima under the influence of both these forces. Snakes maintain the same topology during the evolution stage. So, snakes cannot split to multiple boundaries or merge from multiple initial contours. A traditional snake must start close to the boundary and still cannot converge to boundary concavities.

Geometric deformable models provide an elegant solution to address the primary limitations of parametric deformable models. These models are based on curve evolution theory and the level set method. Level set methods present robust numerical techniques for analyzing and calculating interface evolution problems based on partial differential equations, it is especially suitable for image segmentation without strong previous information. Also many other applications use level sets like capturing multiphase flows, graphics, control and include many others.
4 An improved Texture-resistant Active Contour Model

In this chapter the limitations of the paper of Li [Li, 2005] which was studied at the previous chapter are presented. Then we propose our method that has the same advantages and also fixes the drawbacks of Li’s [Li, 2005] method.

When background is clean and clear, [Li, 2005] algorithm works fine. As we can see in the image two objects are correctly detected using Li’s [Li, 2005] methods (Figure 4.1). Problems occur in using Li’s [Li, 2005] method when the background is noisy or textured. As the definition of the function represents, the contour is moved toward gradient based boundaries. So that wherever the gradient is high the contour will stop. Like in (Figure 4.2), the main object, which is a car, could not be correctly detected. The contour stops in bushes.

![Figure 4.1 Active contour result using Li’s algorithm (clear Background)](image)

[From “Level set evolution without re-initialization: a new variational formulation”]

What we propose to solve this problem is using “pixels polarity information” instead of gradient information.

The most common kind of edge detection techniques use gradient based operators, of which there have been several variations. To detect
objects borders, we employ pixels’ polarity. Polarity helps to smooth the texture of the image in order to help segmenting the main objects.

![Figure 4.2 Active contour result using Li’s algorithm (not clear background)](image)

4.1 Polarity definition

The pixel’s polarity [Carson, 1997] is a local image property, described as a measure of the extent to which the gradient vectors in a certain neighbourhood all point in the dominant orientation $\phi$. Which is defined as:

$$P_\sigma = \frac{|E_+ - E_-|}{E_+ + E_-}$$

(4.1)

Where

$$E_+ = \sum_{x,y} G_\sigma(x,y)[\nabla I.n]_+$$

(4.2)

and

$$E_- = \sum_{x,y} G_\sigma(x,y)[\nabla I.n]_-$$

(4.3)

---

1 Image from http://sampl.ece.ohio-state.edu/data
The gradient of the image intensity $\nabla I$ is computed by the convolution of the image with the first derivative of a Gaussian filter along each dimension difference approximation along each dimension. Scale $\sigma$ is defined to be the width of the Gaussian window within which the gradient vectors of the image are pooled. $G_{\sigma}$ is a Gaussian smoothing kernel with variance $\sigma^2$. $[+]$ and $[-]$ are the rectified positive and negative parts of their arguments, if $\phi$ is the dominant direction in the neighbourhood, $n$ is a unit vector perpendicular to $\phi$. $E^+$ measures how many gradient vectors in the window $G_{\sigma}(x,y)$ are in positive side of $\phi$ and $E^-$ measures how many gradient vectors in the window $G_{\sigma}(x,y)$ are in negative side of $\phi$. The argument of the principal eigenvector of the second-moment matrix below is representing $\phi$.

$$M_{\sigma} = \sum_{x,y} G_{\sigma}(x,y)(\nabla I)(\nabla I)^T$$  

(4.4)

Polarity values are between zero and one (Figure 4.3) Many people use this polarity value as a measure in segmentation like [Belongie, 1998] [Carson, 2002] [Lozano, 2003] [Gordon, 2004] [Pinhas 2004] [Allili, 2007] and so on.

![Figure 4.3 Different values of polarity](image-url)
We use polarity to distinguish main objects boundaries. This should be done in a way that only mitigate the inside regions without bothering the region boundaries. Polarity values vary between zero and one. The polarity value is near 1 for all scales near the region which contains edges (for instance $E^+ \neq 0$ and $E^- = 0$) (Figure 4.3.a). While the polarity vanishes with scale in textured area. (Figure 4.3.b, c). In fact, comparing to gradient based methods; the polarity information accurately discriminates the boundaries of the salient objects (Figure 4.4).

![Comparison between polarity and gradient information](image)

*Figure 4.4 Comparison between polarity and gradient information*

First row original grey level images, second row gradient images of the original images and third row polarity images of the original images
4.2 Method description

In image segmentation, active contours are dynamic curves that move toward the object boundaries. To achieve this goal, an external energy is defined to move the zero level curves toward the object boundaries. Let \( I \) be an image, and \( g \) be the classical gradient based stopping function defined by eq. 4.5:

\[
g = \frac{1}{(1 + |\nabla G_x|)^2} \quad (4.5)
\]

Most of the classical snakes and active contour models use this function as stopping criteria. This function is supposed to vanish when the active contour is very close to the boundaries. However, in practice, the discrete gradient module \(|\nabla G_x|\) can have relatively small local maximums on the object edges and then the stopping function can be comparatively far from zero on the edges, and the curve may pass through the boundary. Also, for the textured or noisy regions \(|\nabla G_x|\) can be very close or equal to those of the object edges. Therefore, the evolving curve may stop before reaching the object boundaries. Moreover, if the image is very noisy, the isotropic smoothing Gaussian (used to compute the gradient module values) has to be strong, which will smooth the edges too.

What we propose to solve this problem is using “pixels polarity information” instead of gradient information. We suggest the use of a stopping function based on the polarity information and defined by eq. 4.6:

\[
g_p = 1 - f \quad (4.6)
\]
Where in eq. 4.6, P(I) is the polarity information of the image I. As the active contour is very close to an edge inside a noisy or textured region, gp is very close to 1, and if the active contour is very close to a salient object boundary, gp is very close to zero. Therefore, the active contour will keep evolving till reaching the salient object boundaries.

Combining the proposed polarity based stopping function with the variational formulation [Li, 2005]; a new external energy (eq. 4.7) is defined as the contour will be absorbed to real boundaries:

$$E_{gp, \lambda, \nu}(\phi) = \lambda \Gamma_{gp}(\phi) + \nu \psi_{gp}$$  \hspace{1cm} (4.7)

Where \( \lambda \) and \( \nu \) are constants, \( \Gamma_{gp} \) is the level set function and the terms \( \Gamma_{gp} \) and \( \psi_{gp} \) are defined as:

$$\Gamma_{gp}(\phi) = \int_{\Omega} g \phi \delta(\phi)^{1/2} |\nabla \phi| \, d\Omega$$  \hspace{1cm} (4.8)

and

$$\psi_{gp}(\phi) = \int_{\Omega} g H(-\phi) \, d\Omega$$  \hspace{1cm} (4.9)

The energy functional calculates the length of the zero level curve of \( \phi \). The energy functional is initiated to speed up curve evolution.

While the function g is constant 1, the energy functional is the area of the region \( \Omega \). The energy functional can be interpreted as the weighted area of \( \Omega \). The coefficient can be positive or negative, depending on the position of the initial contour to the object. For instance, if the initial contours are located outside the object, the coefficient in the
weighted area term should take positive value, so the contours can get smaller faster. If the initial contours are located inside the object, the coefficient should take negative value to speed up the growth of the contours.

Here is the Dirac function and is the Heaviside function.

For the internal energy, we use the same internal energy as [Li, 2005] so that the need for reinitialization is removed:

\[ E_{\text{int}} = u \int_{\Omega} \frac{1}{2} (|\nabla \varphi| - 1)^2 \, \mathrm{d}x \]  

(4.10)

Total energy consists of internal and external energy:

\[ E(\varphi) = E_{\text{int}} + E_{g(x,y)} \]  

(4.11)

Then by using energy minimization method to minimize the total energy it can reach to:

\[ \frac{\partial E}{\partial \varphi} = -\mu [\Delta \varphi - \text{div}(\nabla \varphi)] - \lambda \, \delta(\varphi) \, \text{div}(g) - \nu \, g \]  

(4.12)

Where \( \Delta \) is a Laplacian operator.

And by using gradient descent, \( \frac{\partial E}{\partial \varphi} = -\frac{\partial E}{\partial t} \), the approximation of eq. 4.12 is:

\[ \frac{\partial \varphi}{\partial t} = \mu [\Delta \varphi - \text{div}(\nabla \varphi)] + \lambda \, \delta(\varphi) \, \text{div}(g) + \nu \, g \]  

(4.13)

Finally the level set function is:

\[ \varphi_{l,j}^{k+1} = \varphi_{l,j}^k \quad (\mu [\Delta \varphi - \text{div}(\nabla \varphi)] + \lambda \, \delta(\varphi) \, \text{div}(g) + \nu \, g) \]  

(4.14)
In eq. 4.14, \( \mu > 0 \) supervising the effect of penalizing the deviation of \( \phi \) from a signed distance function, \( \lambda \) is time step, \( \lambda > 0 \) and \( \nu \) are constants and be the edge indicator function. If the initial contours are placed outside the object, the coefficient in the weighted area term should take positive value, so that the contours can shrink faster. If the initial contours are placed inside the object, the coefficient should take negative value to speed up the expansion of the contours. Also \( \delta \) is the Dirac function (with \( \varepsilon = 1.5 \)) defined by eq. 4.15:

\[
\delta(x) = \begin{cases} 
0 & |x| > \varepsilon \\
1 - \frac{1}{2\varepsilon^2} \left[ 1 + \cos\left( \frac{x\pi}{\varepsilon} \right) \right] & |x| \leq \varepsilon 
\end{cases}
\]  

(4.15)

4.3 Summary

We have proposed a fast and efficient active contour model for salient object detection. By combining the polarity information with the active contour model of [Li, 2005] the salient objects can be easily detected. Li’s [Li, 2005] level set equation is working well in the images without background. When there are textures and objects in background the active contour cannot find the object in the scene. That’s why we use polarity information before using active contour program so that we can focus on the main object in the scene. Also as it proposed in [Li, 2005] there is no need for re-initialization procedure and level set function is no longer has to be initialized as a signed distance function. As a result the modified algorithm is able to detect the salient abject on noisy or textured background while Li’s [Li, 2005] method is not capable of that. In addition the method takes advantage of Li’s [Li, 2005] promising points. The initial contour is a fixed square in all the images so there would be no
need for finding a suitable initial contour to start the level set evolution. Comparing to the gradient information, the polarity information accurately distinguishes the boundaries or edges of the salient objects. Moreover, thanks to the use of polarity information, the ad-hoc initialization of the evolving curve inside the image object can be avoided, since the noise and texture outside the object have no more effect or removed.
5 Experiments

In this section, we compare the active contour model based on the polarity information explained in previous chapter and the active contour model based on the gradient information proposed by [Li, 2005]. The proposed method applied to a set of real images and in all of them the main object can be found with acceptable amount of error. The initial level set is the square enclosing the image. So there is no need for specific initialization for every single object.

5.1 Comparison results

In figure 5.1 our method’s result is shown on a 300 × 400 pixel image of a pot with books on the background. $\sigma$ is 0.1 for polarity and $\lambda=5$, $\mu=0.04$ and $\tau=5.0$. This result is compared to using [Li, 2005] active contour on the original image. Both images are computed using 700 iterations in active contour program. As we see in the original image the object is not detectable but in our result the object is found.

In another example\(^2\) of a bird (figure 5.2) the 204 × 157 pixel image of a bird in grass is used. The other parameters are the same as the previous example except the iterations are 200 for both images. As it is obvious the bird and leaves on the grass are detected.

\(^2\) Image from http://pascallin.ecs.soton.ac.uk/challenges/VOC/voc2008/index.html
Figure 5.1 Comparison Results set no.1

Comparison between the two active contour models: (a) original grey level images (b) initial active contours (represented by green lines), (c) segmentation results using gradient based active contour and (d) segmentation results using polarity based active contour.
Comparison between the two active contour models: (a) original grey level images (b) initial active contours (represented by green lines), (c) segmentation results using gradient based active contour and (d) segmentation results using polarity based active contour.

In figure 5.3 the car we show in (Figure 4.2) is segmented by our method.
Comparison between the two active contour models: (a) original grey level images (b) initial active contours (represented by green lines), (c) segmentation results using gradient based active contour and (d) segmentation results using polarity based active contour.

We applied both models on more grey level image and the results are illustrated in advance (Figure 5.4, 5.5, 5.6, 5.7, 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, 5.15, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21) where left image segmentation results using gradient based active; right image segmentation results using polarity based active contour.
The flower in figure 5.4 is detected using polarity. As we see with the gradient the contour stopped in the grass because of the existence of high gradients. Except the tiny edges of the flower leaf, the other parts are segmented right.

Background in figure 5.5 is regarded as noise and the polarity based method is almost perfect here. This can show that existence of noise in the background does not stop our contour.
In figure 5.6 some small errors exist in the output. But the main object is detected and the result is acceptable.

---

3 Image from http://www.vision.caltech.edu/
In this image the man with the camel is segmented. Because of using thresholds to get rid of small edges, the two men in the background are not detected. Segmentation here is acceptable except the shadow which is not part of the salient object.

---

5 Image from http://www.vision.caltech.edu/
Background in figure 5.11 is also regarded as noise and the polarity based method is perfect here. Because the cat has distinct colour from the rest of background, the polarity is comparatively high so there is no error in results.

Figure 5.12 and 5.13 are about microscopic segmentation. The noise around the desired shapes causes problem in gradient based contour while it is easily handled in polarity based contour.

---

6 http://www.emt.tugraz.at/~pinz/data/GRAZ_02/
Below are more examples using birds’ images from Robotics and Computer Vision Laboratory, university of Illinois.

![Figure 5.16 Comparison Results set no.16](http://lear.inrialpes.fr/people/jegou/data.php)

![Figure 5.17 Comparison Results set no.17](http://www-cvr.ai.uiuc.edu/ponce_grp/data/)

10 http://lear.inrialpes.fr/people/jegou/data.php

11 http://www-cvr.ai.uiuc.edu/ponce_grp/data/

12 http://www-cvr.ai.uiuc.edu/ponce_grp/data/
Of course there are some cases when even polarity based method is not successful. This is a case (Figure 5.22) when both algorithms do not work. The goal was finding the see star between the rocks. The polarity of the see star’s neighbourhood is not high comparing to the other parts of the image. That is why the contour could not locate the see star correctly.

15 http://www-cvr.ai.uiuc.edu/ponce_grp/data/
16 http://www-cvr.ai.uiuc.edu/ponce_grp/data/
17 http://lear.inrialpes.fr/people/jegou/data.php
5.2 Summary

From these figures, we can clearly notice that the proposed active contour model based on the polarity information outperforms significantly the gradient based active contour model of [Li, 2005]; in term of salient object detection. Moreover, we can see that the salient objects in each image are efficiently detected using the polarity based active contour, despite of the global and easy initialization of the evolving curve outside of the salient objects. Hence, no more need for the ad-hoc manual initialization of the evolving curves inside the salient objects.
6 Conclusion

Even though image segmentation is a primary problem in image analysis, it has been hard for traditional image segmentation techniques to construct desirable results on images with non-uniform backgrounds. We have proposed a new image segmentation technique, which can produce desired segmentation outputs on difficult image segmentation problems where traditional segmentation methods cannot create satisfying results.

We have developed a fast and efficient model for salient object detection. The proposed image segmentation methods employ the framework of active contours. Since active contours always present continuous boundaries of sub-regions, they can generate more rational segmentation results than traditional segmentation methods. The mathematical implementation of our active contour is achieved by means of level set technique. By introducing contours as a level of a topological function, merging multiple contours into one contour or splitting a contour into multiple contours is possible which provide flexibility in the use of active contours. Then via combining the polarity information with the active contour model of [Li, 2005] the salient objects can be easily detected. In fact, comparatively to the gradient information, the polarity information accurately distinguishes the boundaries or edges of the salient objects. Moreover, thanks to the use of polarity information, the ad-hoc initialization of the evolving curve inside the image object can be avoided, since the noise and texture outside the object have no more effect and they will be skipped.

For the purpose of exhibiting the improved performance of our method, we have applied them to images that other segmentation method [Li, 2005] cannot correctly segment. The segmentation results of
proposed methods and the results of Li’s [Li, 2005] active contour are compared. Our experimental results showed clearly that the proposed active contour model based on the polarity information outperforms significantly the gradient based active contour model of [Li, 2005]; in term of salient object detection.

While the proposed methods generate robust and satisfying results, there are still a few features could be enhanced. Future work will be dedicated to generalize the proposed active contour model to colour images and to improve its performance using relevant region and shape information.
References


Y. Ding, X. Ping, M. Hu and D. Wang, “Range image segmentation based on randomized Hough transform”, *Pattern Recognition*,


A.A. Farag, “Edge-based image segmentation”, Remote Sensing Reviews, Volume 6, Issue 1, Pages 95-121, 1992


C. Revol-Muller, F. Peyrin, C. Odet and Y. Carillon, “Automated 3d region growing algorithm governed by an evaluation function”,


M. Sezgin and B.Sankur, “Survey over image thresholding techniques and quantitative performance evaluation”, Journal of Electronic Imaging, Volume 13, Issue 1, Pages 46-165, 2004

D.B. Sher, “A technique for deriving the distribution of edge neighborhoods from a library of occluding objects”, *Image Analysis and Processing*, Pages 422-429, 1992


L. Xu and E. Oja, “Randomized Hough transform (RHT): Basic mechanisms, algorithms, and computational complexities”, CVGIP - Image Understanding, Volume 57, Pages 131-154, 1993


Vita Auctoris

Farnaz Shariat was born in 1984 in Tehran, Iran. In 2002, she graduated from Farzanehgan high school in Iran. From there she went to Sharif University of technology where she obtained a Bachelor of Science degree with High Honours in information technology in 2007. In the winter of 2009, she graduated with a Master of Science Degree in Computer Science from the University of Windsor in Ontario, Canada.