IMAGE SEGMENTATION FOR CODING

MD. MAHBUBUL ISLAM CHOWDHURY

0-612-55494-5

IMAGE SEGMENTATION FOR CODING

BY

MD. MAHBUBUL ISLAM CHOWDHURY

A THESIS SUBMITTED TO THE SCHOOL OF GRADUATE

STUDIES IN PARTIAL FULFILMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENGINEERING

FACULTY OF ENGINEERING AND APPLIED SCIENCE

MEMORIAL UNIVERSITY OF NEWFOUNDLAND

SEPTEMBER 2000

ST. JOHN'S NEWFOUNDLAND CANADA

Abstract

The segmentation of regions is an important first step for a variety of image analysis and visualization tasks. There is a wide range of image segmentation techniques in the literature. Conventional segmentation techniques for monochromatic images can be categorized into two distinct approaches. One is region based, which relies on the homogeneity of spatially localized features, whereas the other is based on boundary finding, using discontinuity measures. Based on one or both of these properties, diverse approaches to image segmentation exhibiting different characteristics have been suggested in the literature.

The research of this thesis was aimed at combining region growing and edge detection methods to provide better segmentation results. Existing schemes that use region-based processing provide unambiguous segmentation, but they often divide regions that are not clearly separated, while merging regions across a break in an otherwise strong edge. Edge-based schemes are subject to noise and global variation in the picture (e.g. illumination), but do reliably identify strong boundaries. The proposed combined algorithm begins by using region growing to produce an over-segmented image. This phase is fast (order N, where N is the number of pels in the image). The over-segmented output of the region growing is then modified using edge criteria such as edge strength, edge straightness, edge smoothness and edge continuity. Two techniques – line-segment subtraction and line-segment addition – have been investigated. In the subtraction technique, the weakest edge (based on a weighted combination of the criteria) is removed

at each step. Every time that a weakest edge is removed, the combined edge strengths of the remaining edges are recalculated. In the addition technique, the strongest edge (based on the weighted combination of all criteria) of all the edges is calculated first. It is used to seed a multi-segment line that grows out from it at both ends. At each end of the strongest edge, a binary tree containing four branches is investigated. The adjoining edge that has the highest edge strength is appended to the seed. This process of appending continues until a closed loop or a boundary is reached. The overall procedure for both techniques for segmentation has been developed.

In order to investigate the performance of the proposed segmentation techniques, a segmentation evaluation method is demonstrated. Since a human is the ultimate judge, a subjective evaluation method is developed. Segmentation produced by a human is compared to segmentation produced by the algorithm and correlation is calculated between the human method and the algorithms. Subjective tests performed on the algorithms and the results confirm that the proposed algorithms can be used to produce better image segmentation than the segmentation produced by existing region-based techniques.

To my parents

Acknowledgements

I would like to offer my sincere thanks and gratitude to my supervisor Dr. John A. Robinson, for his invaluable guidance and the work he has put into helping me with my masters program. I would also like to thank him for the time he has spent assisting with revisions to this thesis despite his busy schedule.

I am grateful to my colleague Mr. Li-Te Cheng, whose assistance and advice was helpful in all stages of my masters program. I am also grateful to my friend Mr. Ahsan Shamim and all other colleagues at Multimedia Communications Laboratory for their useful comments, suggestions and cooperation throughout my research program.

I would like to offer my special thanks to Dr. John A. Robinson, the School of Graduate studies and the Faculty of Engineering and Applied Science at Memorial University of Newfoundland for providing financial assistance for my masters program.

I am very grateful to my elder brother whose support and encouragement motivated me to be an Engineer and thank my family in Bangladesh for their love, encouragement and moral support.

Table of Contents

List of Figures

Chapter 1 Introduction

With the development of computer technology, image segmentation has become an important tool in many computer vision applications and image processing applications. The goal of image segmentation is to find regions that represent objects or meaningful parts of objects. Based on the applications of an image, division of the image into regions corresponding to objects of interest is necessary before any processing can be done at a level higher than that of the pixel. Identifying real objects, pseudo-objects, and shadows or actually finding anything of interest within the image requires some form of segmentation.

There are several types of images, namely, light intensity (visual) images, range images (depth images), nuclear magnetic resonance images (commonly known as MRI), thermal images and so on. Light intensity (LI) images, the most common type of images we encounter in our daily experience, present the variation of light intensity on the scene. A range image (RI), on the other hand, is a map of depth formation at different points on the scene. In a digital LI image the intensity is quantized, while in the case of a RI the depth value is digitized. Nuclear magnetic resonance images represent the intensity variation of radio waves generated by biological systems when exposed to radio frequency pulses. Biological bodies (humans/animals) are built up of atoms and molecules. Some of the nuclei behave like tiny magnets [27]. Therefore, if a patient is placed in a strong magnetic field, the magnetic nuclei tend to align with the applied magnetic field. For MRI the patient is subjected to a radio frequency pulse. As a result of this, the magnetic nuclei pass into a high energy state, and then immediately relieve themselves of this stress by emitting radio waves through a process called relaxation. This radio wave is recorded to form MRI. In digital MRI, the intensity of the radio wave is digitized with respect to both intensity and spatial co-ordinates.

Thus any image can be described by a two-dimensional function f (x, y), where (x, y) denotes the spatial co-ordinate and f (x, y) the feature value at (x, y). Depending on the type of image, the feature value could be light intensity, depth, and intensity of radio wave or temperature. A digital image on the other hand, is a two-dimensional discrete function f (x, y) which has been digitized both in spatial co-ordinates and magnitude of feature value. I shall view a digital image as a two dimensional matrix whose row and column indices identify a point, called a pixel, in the image and the corresponding matrix element value identifies the feature intensity level.

Segmentation algorithms are generally based on one of two basic properties of grey-level values: discontinuity and similarity. In the first category, an image is partitioned based on abrupt changes in grey level. The principal areas of interest within this category are the detection of isolated points, and the detection of lines and edges in an image. The principal approaches in the second category are based on thresholding, region growing and region splitting and merging. The concept of segmenting an image based on discontinuity or similarity of the grey-level values of its pixels is applicable to both static and dynamic images.

Applications of image segmentation are numerous. For example, in a vision guided car assembly system, the robot needs to pick up the appropriate components from the bin. For this, segmentation followed by recognition is required. Other application areas range from the detection of cancerous cells to the identification of an airport from remote sensing data.

Hundreds of segmentation techniques are present in the literature, but there is no single method which can be considered good for all images, nor are all methods equally good for a particular type of image. Moreover, algorithms developed for one class of image (say ordinary intensity images) may not always be applied to other classes of images (such as MRI and RI). This is particularly true when the algorithm uses a specific information model. For example, some visual image segmentation algorithms are based on the assumption that the grey level function f (x, y) can be modelled as a product of an illumination component and a reflectance component [1]. On the other hand, the grey-level distributions have been modelled as Poisson distributions [4], based on the theory of formation of visual images. Such methods should not be applied to MRI and RIs.

Since none of the proposed segmentation algorithms are generally applicable to all images and different algorithms are not equally suitable for a particular application [1], performance evaluation of segmentation algorithms is complicated and is an important subject in the study of segmentation. More generally, performance evaluation is critical for all computer vision algorithms from research to applications [5], while image segmentation is an essential and important step of low level vision.

Considering the above factors, the aim of this research is to meet the following objectives:

- to provide a new segmentation algorithm that produces a better segmentation result than the conventional techniques based on region growing and merging
- to provide an evaluation method to evaluate the automated segmentation result produced by the newly developed algorithm with the segmented image produced subjectively by human being.

To meet the above objectives, a new algorithm has been developed. By using region growing an over-segmented image is produced which is then modified using different edge criteria such as edge strength, edge smoothness, edge straightness and edge continuity. The newly developed algorithm then shows how weighted combination of different edge criteria [6] such as edge strength, edge smoothness, edge straightness and edge continuity produces better segmentation than that would have been obtained by any of these criteria alone. Beginning from a segmented image allows definition of a subjectively – weighted objective measure of performance. This is based on the order in which a human would merge segments. The performance measure is described then used to compare the segmentation performance between the manually segmented image and the segmented image produced using the new algorithm.

This thesis includes six chapters. Besides the introductory chapter, chapter two develops the necessary background and discusses some previous work in the broad area of image segmentation. Chapter three, "The Algorithm" is the heart of this thesis that discusses the newly developed "Addition" and "Subtraction" algorithms. Chapter four discusses some previous work in the area of segmentation evaluation and demonstrates an evaluation technique to evaluate the segmentation result produced by the proposed algorithms. The capability of the algorithms is then demonstrated in an experimental result and discussion chapter followed by conclusions and suggestions for further improvement in the last chapter.

Chapter 2 Background

2.1 Overview

Image segmentation methods identify objects that either have some measure of homogeneity or have some measure of contrast with neighbouring objects. Most image segmentation algorithms are modifications, extensions or combinations of these two basic concepts. Homogeneity and contrast measures the quantities such as grey level, color and texture. After performing preliminary segmentation, higher-level object properties, such as perimeter and shape, may be incorporated into the segmentation process.

2.2 Segmentation Techniques and Previous Work

There are many challenging issues related to the development of a unified approach to image segmentation, which can (probably) be applied to all kinds of images. Even the selection of an appropriate technique for a specific type of image is a difficult problem. Up to now, there is no universally accepted method of evaluating a segmented output.

There are a wide variety of segmentation techniques in the literature; some are considered general purpose while others are applicable to a specific class of images. Using the basic properties of grey-level values, classical segmentation techniques have been developed. The classical segmentation techniques are based on histogram thresholding, edge detection, iterative pixel classifications, semantic and syntactic approaches [10]. In addition to this, there are certain methods which do not fall clearly in any one of the above classes. In addition to the classical techniques, there are also methods based on fuzzy mathematics [1]. The fuzzy mathematical approach has methods based on edge detection, thresholding and relaxation. Some of these methods, particularly the histogram-based methods, are not at all suitable for noisy images. Several methods have also been developed using neural network models. These algorithms work well even in a highly noisy environment and they are capable of producing output in real time.

The following review consolidates the work of Haralick and Shapiro [12], Pal and Pal [1] and Sahoo et al [13] in classifying segmentation techniques. First I identify the main contributions of each of these three sources.

Haralick and Shapiro [12] classified image segmentation techniques as: (1) measurement space guided spatial clustering (2) single linkage region growing schemes (3) hybrid linkage region growing schemes (4) centroid linkage region growing schemes (5) spatial clustering schemes and (6) split and merge schemes. According to them, the difference between clustering and segmentation is that in clustering, the grouping is done in measurement space; while in image segmentation, grouping is done in the spatial domain of the image. Pal and Pal [1] emphasized that segmentation tries to do the groupings in the spatial domain but it can be achieved through groupings in the measurement space, particularly for multispectral images. For multispectral data, instead of clustering in the full measurement space, Haralick and Shapiro [12] suggested working in multiple lower order projection spaces, and then reflecting these clusters back to the full measurement space. In my classifications, I concentrate on spatial domain grouping for image segmentation, thus accepting Haralick and Shapiro's argument.

Sahoo et al [13] surveyed only segmentation algorithms based on thresholding and attempted to evaluate the performance of some thresholding algorithms using some uniformity and shape measures. They categorized global thresholding techniques into two classes: point dependent techniques and region dependent techniques. They also reviewed several methods of multi-thresholding techniques.

Pal and Pal [1] offered the following comments about the previous reviews on image segmentation:

- None of these surveys [11][12][13] considers fuzzy set theoretic segmentation techniques.
- Neural networks based techniques are also not included.
- The problem of objective evaluation of segmentation results has not been adequately dealt with except in Sahoo et al [13].
- Segmentation of range images/magnetic resonance images has not been considered at all.

Pal and Pal [1] attempted to incorporate all these points to a limited but reasonable extent. They reviewed all previous segmentation techniques and categorized the segmentation techniques into the following categories.

Figure 2.1 lists several classical image segmentation techniques [1][9] consolidating the structures offered by [1][12][13].

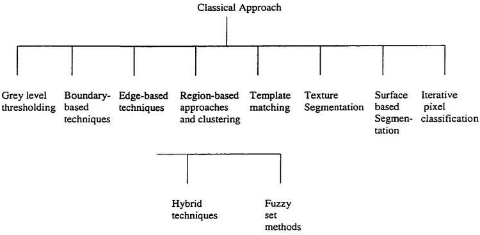

Figure 2.1: Image segmentation techniques

A brief summary of image segmentation techniques is given in the next few subsections.

2.2.1 Grey level thresholding

Thresholding is one of the old, simple and popular techniques for image segmentation. Grey level thresholding is useful whenever the grey level features sufficiently characterize the object [9]. The appropriate grey values are calibrated so that a given grey level interval represents a unique object characteristic. Thresholding can be done based on global information (e.g. the grey level histogram of the entire image) or it can be done using local information (e.g. the co-occurrence matrix) of the image. Under each of the schemes, if only one threshold is used for the entire image then it is called global thresholding. On the other hand, when the image is partitioned into several subregions and a threshold is determined for each of the subregions, the scheme is referred to as local thresholding. Some authors [14] refer to these local thresholding methods as adaptive thresholding schemes. Thresholding can also be classified as bilevel thresholding and multithresholding. In bilevel thresholding the image is partitioned into two regions: object (black) and background (white). When the image is composed of several objects with different surface characteristics (for a light intensity image, objects with different coefficient of reflection, for a range image there can be objects with different depths and so on) one needs several thresholds for segmentation. This is known as multithresholding. Bilevel thresholding is equivalent to classifying the pixels into two classes: object and background.

If the image is composed of regions with different grey level ranges, i.e. the regions are distinct, the histogram of the image usually shows different peaks, each corresponding to a region, and adjacent peaks are likely to be separated by a valley. For example, if the image has a distinct object on a background, the grey level histogram is likely to be bimodal with a deep valley. In this case the bottom of the valley is taken as the threshold for object background separation. Therefore, when the histogram has one (or a set of) deep valley(s), selection of threshold(s) becomes easy because it involves detecting valleys. Normally the situation is not like this and threshold selection is not a trivial job. There are various methods available [13] for this. For example, Otsu [58] maximized a measure of class separability. He maximized the ratio between the class variance to the local variance to obtain thresholds.

Pun [15] assumed that an image is the outcome of an L symbol source. He maximized an upper bound of the total a posteriori entropy of the partitioned image for the purpose of selecting the threshold. Kapur et al [16], on the other hand, assumed two probability distributions, one for the object area and the other for the background area. They then, maximized the total entropy of the partitioned image in order to arrive at the threshold level. Though these methods use only the histogram, they produce good results due to the incorporation of the image formation model.

All these methods have a common drawback; they take into account only the histogram information (ignoring the spatial details). As a result such algorithms may fail to detect thresholds if these are not properly reflected as valleys in the histogram, which is normally the case. There are many thresholding schemes that use spatial information, instead of histogram information. All these methods threshold the histogram, but since they make use of spatial details, they result in a more meaningful segmentation than the methods which use only the histogram information.

The philosophy behind grey level thresholding, "pixels with grey level <= T fall into one region and the remaining pixels belong to another region", may not be true on many occasions, particularly, when the image is noisy or the background is uneven and illumination is poor. In such cases the objects will still be lighter or darker than the background, but any fixed threshold for the entire image will usually fail to separate the objects from the background. This leads one to the methods of adaptive thresholding. In adaptive thresholding normally the image is partitioned into several non-overlapping blocks of equal area and a threshold for each block is computed independently (sometimes use overlapping regions and blend thresholds). The sub histogram of each block is used to determine local threshold values for the corresponding cell centers. These local thresholds are then interpolated over the entire image to yield a threshold surface.

2.2.2 Boundary based techniques

Boundary extraction techniques segment objects on the basis of their profiles. Thus, contour following, connectivity, edge linking and graph searching, curve fitting, Hough transform and other boundary extraction techniques are applicable to image segmentation

[9]. Difficulties with boundary-based methods occur when objects are touching or overlapping or if a break occurs in the boundary due to noise or artefacts in the image.

2.2.3 Edge based techniques

Segmentation can also be obtained through detection of the edges of regions, normally by locating points of abrupt changes in grey level intensity values. There are various types of edge detection operators in use today. Many are implemented with convolution masks, and most are based on discrete approximations to differential operators. Some edge detection operators return orientation information (information about the direction of the edge), whereas others only return information about the existence of an edge at each point.

Different edge operators [10] like Sobel, Prewitt, Marr-Hildreth, Canny produce an edgeness value at every pixel location. However not all of them are valid candidate for edges. Normally, edges are required to be thresholded. The selection of the threshold is very crucial as for some parts of the image low intensity variation may correspond to edges of interest while in other parts there is high intensity variation. Adaptive thresholding is often taken as a solution to this. Obviously it cannot eliminate the problem of threshold selection. A good strategy to produce meaningful segments would be to fuse region segmentation results and edge outputs. Incorporation of psychovisual phenomena may be good for light intensity images but not applicable for range images. Semantics and a priori information about the type of image are critical to the solution of the segmentation problem.

Since edges are local features, they are determined based on local information. Davis [21] classified edge detection techniques into two categories: sequential and parallel. In the sequential technique the decision whether a pixel is an edge pixel or not is dependent on the result of the detector at some previously examined pixels. On the other hand, in the parallel method the decision whether a point is an edge or not is made based on the point under consideration and some of its neighboring points. As a result of this the operator can be applied to every point in the image simultaneously. The performance of a sequential edge detection method is dependent on the choice of an appropriate starting point and how the results of previous points influence the selection and result of the next point.

There are different types of parallel differential operators such as Roberts gradient, Sobel gradient, Prewitt gradient and the Laplacian operator. These difference operators respond to changes in grey level or average grey level [1]. The gradient operators not only respond to edges but also to isolated points. For Prewitt's operators the response to diagonal edges is weak while Sobel's operator gives greater weights to points lying close to the point (x, y) under consideration. However, both Prewitt's and Sobel's operators possess greater noise immunity than other difference operators. The preceding operators are called first difference operators. On the other hand, the Laplacian is a second difference operator.

The digital Laplacian being a second difference operator, has a zero response to linear ramps. It responds strongly to corners, lines and isolated points. Thus for a noisy picture, unless the picture has low contrast, the noise will produce higher Laplacian values than the edges.

According to Canny [22] a good edge detector should be a differential operator and should have the following three properties: (1) Low probability of wrongly marking non-edge points and low probability of failing to mark real edge points (2) points marked as edges should be as close as possible to the center of true edges (i.e. good localization) (3) one and only one response to a single edge point (single response). Good detection can be achieved by maximizing signal to noise ratio (SNR), while for good localization Canny used the reciprocal of an estimate of the r.m.s. distance of the marked edge from the center of the true edge. To maximize simultaneously both good detection and localization criteria Canny maximized the product of SNR and the reciprocal of standard deviation of the displacement of edge points. The maximization of the product is done subject to a constraint which eliminates multiple responses to single edge points. Canny then described an edge detector that can successfully detect an edge by providing low probability of wrongly marking non-edge points, good localization and single response. In the case of noise free images, the edge angle can be measured accurately, but in real life images, noise cannot be avoided and it makes it difficult to estimate the true angles. There are many post-Canny operators, but for successful detection of an edge, a good edge detector should have above three criteria.

An iterative algorithm has been developed by Gokmen and Li [23] using regularization theory. The energy functional in the standard segmentation has been modified to spatially control the smoothness over the image in order to obtain the accurate location of edges. An algorithm for defining a small, optimal kernel conditioned on some important aspects of the imaging process has been suggested by Reichenbach et al [34] for edge detection. This algorithm takes into account the nature of the scene, the point spread function of the image gathering device, the effect of noise etc. and generates the kernel values which minimize the expected mean square error of the estimate of the scene characteristics. Pal and Pal [1] discussed various operators to get edge values. All the edges produced by these operators are, normally, not significant (relevant) edges when viewed by human beings. Therefore, one needs to find out prominent (valid) edges from the output of the edge operators.

2.2.4 Region based techniques

The main objective in region-based segmentation techniques is to identify various regions in an image that have similar features. One class of region-based techniques involves region-growing [9]. As implied by its name, region growing is a procedure that groups pixels or subregions into larger regions. The simplest of these approaches is pixel aggregation, where the growing process starts with "seed" points and from these grow regions by appending to each seed point those neighboring pixels that have similar properties (e.g. grey level, texture, color). If the absolute difference between the grey level of the neighboring pixel and the grey level of the seed is less than a threshold then that neighboring pixel is added to the seed.

Although this growing procedure is simple in nature, it suffers some important problems in region growing. Two immediate problems are the selection of initial seeds that properly represent regions of interest, and the selection of suitable properties for including points in the various regions during the growing process. Another important problem in region growing is the formulation of a stopping rule. Basically, a region growing process is stopped when no more pixels satisfy the criteria (e.g. intensity, texture) for inclusion in that region. Additional criteria that increase the power of a region-growing algorithm incorporate the concept of size, likeness between a candidate pixel and the pixel grown thus far (e.g. a comparison of the intensity of a candidate and the average intensity of the region), and the shape of a given region being grown.

S.A. Hojjatoleslami and J. Kittler presented [40] a new idea for region growing by pixel aggregation, which used new similarity and discontinuity measures. A unique feature of their proposed approach is that in each step at most one candidate pixel exhibits the required properties to join the region. This makes the direction of the growing process more predictable. The procedure offered a framework in which any suitable measurement can be applied to define a required characteristic of the segmented region. The authors used two discontinuity measurements called average contrast and peripheral contrast to control the growing process. Local maxima of these two measurements identify two nested regions, average contrast and peripheral contrast regions. The method first finds the average contrast boundary of a region, and then a reverse test is applied to produce the peripheral contrast boundary. Like other existing methods, this method is not universal, but it does appear to have a fairly wide application potential.

Another method of region based segmentation is region splitting and merging. An image is initially subdivided into a set of arbitrary, disjointed regions. The adjacent regions are merged if they are identical, otherwise split. This splitting and merging process continues and stop when no further merging or splitting is possible.

2.2.4.1 Region growing method used in this thesis

The region growing method used in this thesis is a new clustering region growing algorithm of John Robinson. In this algorithm clusters are grown by absorbing adjacent clusters in the spatial domain. Adjacent clusters are joined according to their distance in grey value at the join point, and some other criterion of similarity. The alternatives for the other criterion are: (1) no other criterion: just specify the desired number of clusters (2) difference in average between the two clusters below a given threshold (3) the effect on representation accuracy of using the mean of the two clusters instead of the mean of each must be below a threshold.

Mathematically,

C1 = d(x_i, x_j) where d() measures in grey-level difference, x_i and x_j are neighbors in real space. x_i belongs to cluster1 and x_j belongs to cluster2.

C2 = Xd(m1, m2) where m1 is mean of cluster1 and m2 is mean of cluster2 (in color space). X is the modifier. It can be one the following:

0	cluster means have no effect
1	cluster means compared
2*n1*n2/(n1+n2)	where n1 is the number of points in cluster1 and n2 is the

number of points in cluster2. This gives increase in representation error.

C1s are ordered in ascending order. For each in turn, the corresponding clusters are found and merged if C2<T. Stop this process when the desired number of clusters is reached.

2.2.5 Template matching

One direct method of segmenting an image is to match it against templates from a given list. The detected object can then be segmented out and the remaining image can be analyzed by other techniques [9]. This method can be used to segment busy images, such as journal pages containing text and graphics. The text can be segmented by template-matching techniques and graphics can be analyzed by boundary following algorithms.

2.2.6 Texture segmentation

Texture segmentation becomes important when objects in a scene have a textured background. Since texture often contains a high density of edges, boundary-based and region-based techniques may become ineffective unless the texture is filtered out [9]. Clustering and region-based approaches applied to textured features can be used to segment textured regions. In general, texture classification and segmentation is quite a difficult problem. Use of a priori knowledge about the existence and kinds of textures that may be present in a scene can be of great utility in practical problems.

2.2.7 Surface based segmentation

This section mainly discusses a few selected techniques for range image segmentation [18][19][20]. Besl and Jam [18] developed an image segmentation algorithm based on the assumption that the image data exhibits surface coherence, i.e. image data may be interpreted as noisy samples from a piece-wise smooth surface function. Though this method is most useful for range images, it can be used to segment any type of image that can be modelled as a noisy sampled version of a piece-wise smooth graph surface. This method is based on the fact that the signs of Gaussian and mean curvatures yield a set of eight surface primitives: peak, pit, ridge, saddle ridge, valley, saddle valley, flat and minimal. These primitives possess some desirable invariant properties and can be used to decompose any arbitrary smooth surfaces. In other words, any arbitrary smooth surface can be decomposed into one of those eight possible surface types. These sample surfaces can be well - approximated, for the purpose of segmentation, by bivariate polynomials of order 4. The first stage of the algorithm creates a surface type label image based on the local information (using mean curvature and Gaussian curvature images). The second stage takes the original image and the surface type image as input and performs an iterative region growing using the variable order surface fitting. In the variable order surface fitting, first it had been tried to represent the points in a seed region by a planar surface. If this simple hypothesis of planar surface is found to be true then the seed region is grown on the planar surface fit. If this simple hypothesis fails, then the next most complicated hypothesis of biquadratic surface fit is tried. If this is satisfied, the region is grown based on that form otherwise, the next complicated form is tried. The process is terminated when either the region growing has converged or when all preselected hypothesis fail.

Hoffman and Jain [19] developed a method for segmentation and classification of range images. They used a clustering algorithm to segment the image into surface patches. Different types of clustering algorithms including methods based on minimal spanning tree, mutual nearest neighbor, hierarchical clustering and square error clustering were attempted. The square error clustering was found to be the most successful method for range images. In order to make the method of classification more effective the authors combined three different methods, namely, "non-parametric trend test for planarity", "curvature planarity test" and the "eigenvalue planarity test". In the final stage, boundaries between adjacent surface patches were classified as crease or non-crease edges and this information was used to merge adjacent compatible patches to result in reasonable faces of the object. For this type of method, the choice of the neighborhood to compute the local parameters is an important issue and no theoretical guideline has been provided for this.

Yokoya and Levine [20] also used a differential geometric technique like Besl and Jain [18] for range image segmentation. Yokoya and Levine combined both region and edge based considerations. They approximated object surfaces using biquadratic polynomials. Two edge maps are formed: one for the jump edge and the other for the roof edge. The jump edge magnitude is obtained by computing the maximum difference in depth between a point and its eight neighbors, while the roof edge magnitude is computed as the maximum angular difference between adjacent unit surface normals. These two edge maps and the curvature sign map are then fused to form the final segmentation. This method too requires selection of threshold levels for the maps and the curvature sign map. Improper choice of these parameter values is likely to impair the quality of the segmentation output. For range images, detection of jump edges can be done with ordinary gradient operators, but detection of crease edges with ordinary gradient operators become difficult. Thus for edge detection in range images, one needs to account for both crease and jump edges separately.

2.2.8 Iterative pixel classification

According to Pal and Pal [1], iterative pixel classification includes MRF based approach and Neural network based approaches. A brief description of these approaches is given in the following subsections.

2.2.8.1 MRF based approach

There are many image segmentation methods which use the spatial interaction models like Markov Random Field (MRF) or Gibbs Random Field (GRF) to model digital images. Geman and Geman [17] have proposed a hierarchical stochastic model for the original image and develop a restoration algorithm, based on stochastic relaxation (SR) and annealing for computing the maximum a posterior estimate of the original scene given a degraded realization. Due to the use of annealing, the restoration algorithm does not stop at a local maximum of the a posterior probability. Pal and Pal [1] mentioned that the probabilistic relaxation (also known as relaxation labelling (RL)) and stochastic relaxation (SR), although they share some common features like parallelism and locality, are quite distinct. RL is essentially a non-stochastic process which allows jumps to states of lower energy. On the other hand, in SR transition to a configuration which increases the energy (decreases the probability) is also allowed. In fact, if the new configuration decreases the energy, the system transits to that state, while if the new configuration increases the energy the system accepts that state with a probability. This helps the system to avoid the local minima. RL usually gets stuck in a local minima. Moreover, in

RL, there is nothing corresponding to an equilibrium state or even a joint probability law over the configurations. Derin et al [31] extended the one-dimensional Bayes smoothing algorithm to two dimensions to get the optimum Bayes estimate for the scene value at every pixel. In order to reduce the computational complexity of the algorithm, the scene is modelled as a special class of MRF models, called Markov mesh random fields which are characterized by causal transition distributions. The processing is done over relatively narrow strips and estimates are obtained at the middle section of the strips. These pieces together with overlapping strips yield a sub-optimal estimate of the scene. Without parallel implementation these algorithms become computationally prohibitive. At the top level a Gibbs distribution (GD) is used to characterize the clusters of the image pixels into regions with similar features. At the bottom level, the feature or textural properties of region types are modelled by a second set of GD, one for each type of class. The segmentation algorithms are derived by using the maximum a posterior probability (MAP) criterion.

2.2.8.2 Neural network based approaches

For any artificial vision application, one can desire to achieve robustness of the system with respect to random noise and failure of processors. Moreover, a system can (probably) be made artificially intelligent if it is able to emulate some aspects of the human information processing system. Another important requirement is to have the output in real time. Neural network based approaches are attempts to achieve these goals.

Neural networks are massively connected networks of elementary processors [1]. Architectures and dynamics of some networks are claimed to resemble information processing in biological neurons. The massive connectionist architecture usually makes the system robust while the parallel processing enables the system to produce output in real time. Several authors have attempted to segment an image using neural networks. Blanz and Gish [32] used a three-layer feed forward network for image segmentation, where the number of neurons in the input layer depends on the number of input features for each pixel and the number of neurons in the output layer is equal to the number of classes. Babaguchi et al [33] used a multilayer network trained with back propagation, for thresholding an image. The input to the network is the histogram while the output is the desirable threshold. In these methods, at the time of learning a large set of sample images with known thresholds, which produce visually, suitable outputs are required. But for practical applications it is very difficult to get many sample images.

2.2.9 Hybrid Techniques

Hybrid technique combines two or more techniques for segmentation to produce much better results than what would have been obtained by either technique alone. There are few earlier papers that have described edge-based techniques for improving image segmentation. Following is a brief review of hybrid techniques that combine region-growing and edge detection techniques to produce better results.

Bajcsy [35] showed that both edge detection and region growing processes could be unified by making the decision whether a point is on a boundary or on a homogeneous surface. Anderson and Bajcsy [36] showed a combination of edge detection and region growing where they used edge detection to initialize a region growing process based on a local similarity threshold which was used to check whether two points belong to the same region. Montanari [39] considered smoothness and contrast and applied a measure to minimize the weighted sum of smoothness and contrast. The recent paper that integrated region growing and edge detection to produce better segmentation is described in [37]. The authors showed that any region growing process suffers from three kind of errors [37]: (1) a boundary is not an edge and there is no edges nearby (2) a boundary corresponds to an edge but it does not coincide with it (3) there exist edges with no boundaries near them. They performed first region growing and then edge detection without iterating and they had little choice but to use over-segmented images (an image that has too many segments).

The focus of this thesis is to develop a hybrid technique that combines different edge criteria such as edge strength, edge straightness, edge smoothness and edge continuity in order to produce a better segmentation than what would have been obtained by either criterion alone.

2.2.10 Methods based on fuzzy set theory

The relevance of fuzzy set theory in pattern recognition problems has adequately been addressed in the literature [24]. It is seen that the concept of fuzzy sets can be used at the feature level in representing an input pattern as an array of membership values denoting the degree of possession of certain properties and in representing linguistically phrased input features; at the classification level in representing multi-class membership of an ambiguous pattern, and in providing an estimate of missing information in terms of membership values. In other words, fuzzy set theory may be incorporated in handling uncertainties (arising from deficiencies of information: the deficiencies may result from incomplete, imprecise, ill defined, not fully reliable, vague, contradictory information) in various stages of the pattern recognition system. While the application of fuzzy sets in cluster analysis and classifier design is in the process of development [1], an important and related effort in fuzzy image processing and recognition is evolving more or less in parallel with the aforesaid general developments.

Conventional approaches to image analysis and recognition consists of segmenting the image into meaningful regions, extracting their edges and skeletons, computing various features properties (e.g. area, perimeter, centroid etc.) and primitives (e.g. line, corner, curve etc.) of and relationship among the regions, and finally, developing decision rules, grammars for describing, interpreting and or classifying the image and its subregions. In a conventional system each of these operations involves crisp decisions (i.e. yes or no, black or white, 0 or 1) about regions, features, primitives, properties, relations and interpretations.

Since the regions in an image are not always crisply defined, uncertainty can arise within every phase of the aforesaid tasks. Any decision made at a particular level will have an impact on all higher level activities. A recognition (or vision) system should have sufficient provision for representing and manipulating the uncertainties involved at every processing stage: i.e. in defining image regions, features, matching and relations among them, so that the system retains as much of the "information content" of the data as possible. If this is done, the ultimate output of the system (result) of the system will possess minimal uncertainty [1] (and unlike conventional systems, it may not be biased or affected as much by lower level decision components).

For example, consider the problem of object extraction from a scene [1]. It is difficult to define exactly the target or object region in a scene when its boundary is ill defined. Any hard thresholding made for the extraction of the object will propagate the associated uncertainty to subsequent stages (e.g. thinning, skeleton extraction, primitive selection) and this might, in turn, affect feature analysis and recognition. Consider, for example, the case of skeleton extraction of a region through medial axis transformation (MAT). The MAT of a region in a binary picture is determined with respect to its boundary. In a grey tone image, the boundaries are not well defined. Therefore, errors are more likely if we compute the MAT from the hard-segmented version of the image.

Thus, it is convenient, natural and appropriate to avoid committing [1] to a specific (hard) decision (e.g. segmentation thresholding, edge detection and skeletonization), by allowing the segments or skeletons or contours to be fuzzy subsets of the image, the subsets being characterized by possibility to which each pixel belong them. Similarly, for describing and interpreting ill-defined structural information in a pattern, it is natural to define primitives (line, corner, curve etc.) and relations among them using labels of fuzzy sets. Thus, a few methods of fuzzy segmentation (based on both grey level thresholding and pixel classification) and edge detection have been developed using global and or local information of an image space.

2.3 Summary

This chapter reviews and summarizes some existing methods of image segmentation and their drawbacks. The literature has also discussed the scope for the fuzzy set theoretic approaches and the neural network model based algorithms to segmentation. Moreover, these algorithms are robust. It is well known that no method is equally good for all images and all methods are not good for a particular type of image. Selection of an appropriate segmentation technique largely depends on the type of images and application areas. The important problem is how to make a quantitative evaluation of segmentation results. Such a quantitative measure would be quite useful for vision applications where automatic decisions are required. It is very difficult to find a single quantitative index for this purpose because such an index should take into account many factors like homogeneity, contrast, compactness, continuity, psycho-visual perception etc. Possibly the human being is the best judge to evaluate the output of any segmentation algorithm. However, it may be possible to have a small vector of attributes that can be used for objective evaluation of results. The next chapter discusses a new algorithm to produce a better segmentation result that incorporates some edge criteria such as edge strength, edge straightness, edge smoothness and edge continuity with region growing segmentation, evaluation scheme is also demonstrated.

Chapter 3 The Algorithm

3.1 Introduction

A wide range of segmentation techniques in the literature have been discussed in chapter two. Edge based techniques produce sharp edges, on the other hand, region-based techniques produce some ambiguous regions. This chapter discusses a hybrid technique, which combines different edge criteria to improve region based image segmentation.

Edge detection operators are based on the idea that edge information in an image is found by looking at the relationship a pixel has with its neighbors. If a pixel's grey-level value is similar to those around it, there is probably not an edge at that point. However, if a pixel has neighbors with widely varying grey-levels, it may represent an edge point. In other words, an edge is defined by a discontinuity in grey-level values.

Figure 3.1 shows a natural image, lenna256, and figure 3.2 shows an edge detection image produced by a Sobel operator. The Sobel edge detection masks look for edges in both the horizontal and vertical directions and then combine this information into a single metric. These masks (row and column masks) are each convolved with the image (here lenna256). At each pixel location there are two numbers: s1, corresponding to the result from the row mask, and s2, from the column mask. These two numbers are used to compute two metrics, the edge magnitude and the edge direction.

ROW MASK

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

COLUMN MASK

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$\text{Edge Magnitude} = \sqrt{s1^2 + s2^2}$$

$$\text{Edge Direction} = \tan^{-1}(s1 / s2)$$

Figure 3.1: Original Image (lenna256)

Figure 3.2: Edge Image produced by Sobel operator

The edge magnitude image produced by Sobel operator shows the sharp edges but it does not show some structurally important edges. For example, the edges in the left and right jaws and in the forehead of the original image are missing in the edge image produced by Sobel operator.

Figure 3.3: Segmented Image (lenna256)

Figure 3.4: Candidate edges from the segment boundaries of 3.3

Region-based methods provide unambiguous segmentation, but often divide regions that are not clearly separated by a strong boundary [6]. Figure 3.3 shows a segmented image of the original image (lenna256) produced by region growing methods and figure 3.4 shows the candidate edges from the segment boundaries of figure 3.3. In this segmented image there are some regions that are not clearly separated by a meaningful boundary. The central idea of this thesis is to begin with an image (as shown in figure 3.3) in which segmentation has already been done but not gone too far. This image is then going to be incrementally improved by using different edge measures along its boundaries. Although similar in aim to some previous hybrid methods discussed in section 2.2.9, the approach is new, and, as I describe later, leads itself to comparison with human-generated segmentations and therefore to quantitative evaluation.

The goal of this chapter is to show how weighted combination of different edge criteria such as edge strength, edge smoothness, edge straightness and edge continuity can produce better segmentation than would have been obtained by any of these criteria alone. It also describes two new algorithms for progressively applying the combined criteria to an over-segmented image (an image that has too many segments) previously produced by region growing. These are line-segment subtraction and line-segment addition.

The concept of the newly developed algorithm is presented in the next section.

3.2 The Algorithm

An over-segmented image is produced by region growing. The proposed algorithm acts on this over-segmented image. The output of the over-segmented image is an image that contains the segment boundaries where variation of intensities is found among the adjacent regions. In this image (that contains segment boundaries) there are many points where two or more edges meet. These points on the edges are known as vertices. Each edge starts from one vertex or from a boundary and ends at another vertex or at another or the same boundary. Each edge has different criteria such as edge strength, edge straightness, edge smoothness and edge continuity. Edge strength is the most important criterion of an edge. It is calculated by measuring the absolute intensity difference between two pels across the edge in the original image. If a and b are two points on each side of an edge as shown in figure 3.5, then the edge strength E_s of that particular point is

$$E_s = | a - b| \quad .... (3.3)$$

If N is the total number of points in that edge, then the total edge strength E_{st} is

$$E_{st} = \sum_{i=1}^{N} E_{si} \quad .... (3.4)$$

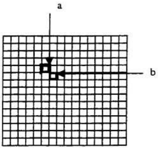

Figure 3.5: Edge strength calculation (single point)

The edge strength calculated in equation 3.4 is not a local measure. It is defined by stringing together local measures along a trajectory defined by the initial segmentation.

More than one point can be considered on each side of the edge. In that case the local averages of the intensities are calculated. If a and c are two points on one side and b and d are two points in the opposite side of the edge as shown in figure 3.6, then the edge strength of that particular point is

$$E_s = |(a+c)/2 - (b+d)/2| \quad ....(3.4)$$

Then the total edge strength can be calculated using using equation (3.4). Weighted averages could also be used in equation 3.4.

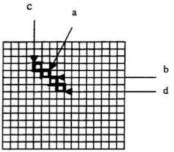

Figure 3.6: Edge strength calculation (double point)

Pixel positions shown in figure 3.5 and 3.6 are for a particular position of a pixel in the segment boundary only. Pixel positions are always taken in a line orthogonal to the boundary direction.

Straightness is the second most important criterion of an edge. It is calculated based on the Euclidean distance between the starting and the end points of the edge. If (x1, y1) and (x2, y2) are the starting and the end points of an edge respectively, Euclidean distance

$$D_{Euclidean} = \sqrt{(x1 - x2)^2 + (y1 - y2)^2} \quad ..... (3.5)$$

So, the straight distance is the Euclidean distance between two end points of an edge and the straight distance is

$$E_{straight\_distance} = D_{Euclidean} \quad .... (3.6)$$

The straightness can be defined as $D_{Euclidean}$ / N. Where N is the total number of points in that edge.

Smoothness is calculated based on the "local straightness" of a pel in an edge. The phrase "local straightness" means that every successive pel in an edge is compared with its two previous pels to check whether it is in a straight line with them or not. Figure 3.7 (a) shows a smooth portion of an edge and figure 3.7 (b) shows a portion of an edge that is not smooth. If the incoming pel is in a straight line with the previous two pels, it is considered as locally straight, otherwise it is not locally straight.

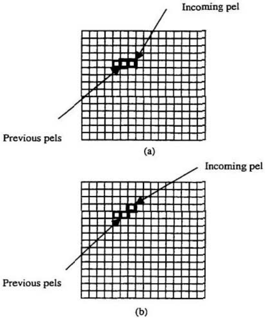

Figure 3.7: Smoothness calculation (a) smooth portion of an edge (b) nonsmooth portion of an edge

Every time a pel is not locally straight, it increases a counter. If an edge is fairly smooth its counter value is lower and if it is not smooth its counter value is higher. So, the smoothness of an edge is

$$E_{smooth} = \text{No. of count} / N \quad \dots\dots\dots\dots\dots\dots \quad (3.7)$$

Where N is the number of pels in the edge. The lower the E_{smooth} value, the more smooth the edge. Smoothness can also be measured over a larger window, i.e. more pels can be considered along the boundary for checking local straightness.

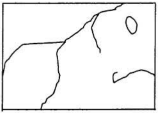

Figure 3.8: Typical discontinuous edges

Sometimes there are edges in an image that start from an image boundary or from a vertex and end without meeting another vertex or boundary or an edge. These edges are called discontinuous edges. These discontinuous edges are created when one or more edges are removed from the region or added to the strongest edge while processing the algorithms. Figure 3.8 shows some typical discontinuous edges. Discontinuous edges do not contribute to the combined edge strength. If an edge is continuous its edge continuity, $E_{continuity}$ is set to 1; otherwise $E_{continuity}$ is set to 0.

The weighted combination of all these criteria is used to calculate the combined edge strength of a particular edge. So, the combined edge strength $E_{combined}$ is

$$E_{combined} = \text{Edge\_strength\_factor} * E_{st} + \text{Straightness\_factor} * E_{straight\_distance}$$
$$- \text{Smoothness\_factor} * E_{smooth} + \text{Continuity\_factor} * E_{continuity} \quad .....(3.8)$$

Where the Edge_strength_factor, Straightness_factor, Smoothness_factor and Continuity_factor are the scaling parameters. Straightness can be used in place of $E_{straight\_distance}$ for calculating combined edge strength. In preliminary experiment, it is found that straightness does not work very well. When straightness is used for calculating combined edge strength, some regions from the output image are eliminated those are straight and important. So, $E_{straight\_distance}$ is used for calculating combined edge strength.

Based on this combined edge strength two techniques, line-segment subtraction and line-segment addition, have been developed.

3.2.1 Subtraction Technique

Based on a weighted combination of all the criteria, the weakest edge is found among all the edges in the region and removed if it is below a threshold. The threshold is chosen based on the highest and the lowest value of the combined edge strength. After removing the weakest edge, the edges that are adjacent to the weakest edge's end points are joined so that they appear as a single edge. If one end of the weakest edge is in the boundary, then the algorithm simply removes that edge.

Figure 3.9 shows an example of this technique. If edge number 4 is removed from this region, then edge number 1, edge number 2 and edge number 3 are joined together so that all the three edges appear as a single edge. If edge number 1 is the weakest edge, then the algorithm simply removes it. Every time that a weakest edge is removed, the combined edge strengths of the remaining edges are recalculated. This procedure is repeated until all the edges below the threshold are removed.

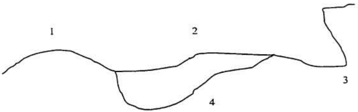

Figure 3.9: Subtraction technique

The subtraction technique therefore works by removing segmentation boundaries, i.e. merging adjacent segments, according to the edge criteria along their shared boundary.

3.2.2 Addition Technique

In this technique, the strongest edge (based on a weighted combination of all the criteria) of all the edges is calculated first. It is used to seed a multi-segment line that grows out from it at both ends. At each end of the strongest edge, the algorithm looks at a binary tree containing four branches. The adjoining edge that has the highest edge strength (based on the weighted combination of all criteria) is appended to the seed.

Figure 3.10 shows the addition algorithm at one end of the seed. There are four possible branches: 24, 25, 36 and 37. The strongest branch of these four, or its first segment only,

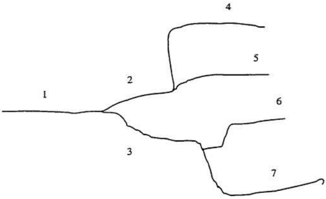

Figure 3.10: Addition algorithm

is appended to 1 if the combined edge strength (based on the weighted combination of all criteria) is greater than the threshold. A similar procedure is followed at the other end. This process of appending continues until a closed loop or a boundary is reached. The same process is repeated, finding the next strongest edge (previously unused) until a certain number of lines has been found. Every time the strongest edge of the remaining edges is considered as seed. These lines are selected by visual observation. Every time the output image is compared with the original image to see whether the segmentation matches with the original image or not. If the segmentation perfectly matches, the algorithm stops processing, otherwise it continues processing.

The addition technique can also be performed by adding one edge at a time without looking ahead on the binary tree. In that case, the strongest edge (based on weighted combination of all criteria) of edge number 2 and edge number 3 in figure 3.10 is added to edge number 1. In preliminary experiments it was found that looking ahead on the binary tree produces better results than adding single edge at a time. That is why the binary tree algorithm is recommended for the addition technique.

3.4 Summary

The goal of this chapter was to produce a better segmentation using region-growing methods. Two techniques – line-segment subtraction and line-segment addition have been developed using different edge criteria. The segmented output produced by these algorithms is discussed in chapter 5. Segmented output produced by objective and subjective measures is discussed in chapter 4.

Chapter 4 Segmentation Evaluation

4.1 Introduction

While development of segmentation algorithms has attracted significant attention, relatively fewer efforts have been spent on their evaluation, although many newly developed algorithms are (most often subjectively) compared with some particular algorithms with few particular images. Moreover, most efforts spent on evaluation are just for designing new evaluation methods and only few authors have attempted to characterize the different evaluation methods. Zhang [42] reviewed different existing methods for segmentation evaluation.

This chapter discusses the existing segmentation evaluation techniques and develops a new idea for segmentation evaluation. Existing segmentation evaluation techniques are discussed in section 4.2 and a new technique is demonstrated in section 4.3 followed by a brief summary of all evaluation methods.

4.2 Previous Works

According to Zhang [42], segmentation algorithms can be evaluated analytically or empirically. The analytical methods directly examine and assess the segmentation algorithms themselves by analysing their principles and properties. The empirical methods indirectly judge the segmentation algorithms by applying them to test images and measuring the quality of segmentation results. Various empirical methods have been proposed. Most of them can be classified into two types [42]: goodness methods and discrepancy methods. In the first category some desirable properties of segmented images, often established according to human intuition, are measured by "goodness" parameters. The performances of segmentation algorithms under investigation are judged by the values of goodness measures. In the second category some references that present the ideal or expected segmentation results are first found. The performances of segmentation algorithms under investigation are then assessed according to the discrepancy measures. Following this discussion, three groups of methods can be distinguished.

The above classification for evaluation methods can be seen more clearly in figure 4.1, where a general scheme for segmentation and its evaluation is presented. The input image obtained by sensing is first (optionally) preprocessed to produce the segmenting image for the segmentation procedure. The segmented image can then be (optionally) post-processed to produce the output image. In figure 4.1 the parts enclosed by the rectangular boxes with thin lines correspond to the segmentation procedure in its narrow sense, while the parts enclosed by the rectangular boxes with solid lines correspond to the segmentation procedure in its general form. The black arrow indicates the processing directions of segmentation. The access points for the three groups of evaluation methods are depicted with dotted arrows.

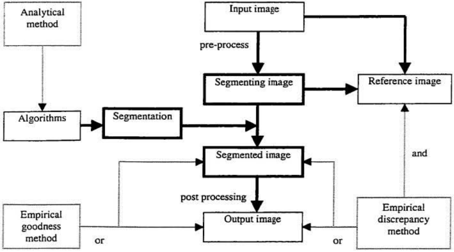

Figure 4.1: General scheme for segmentation and its evaluation

The analysis methods treat the algorithms for segmentation directly. The empirical goodness methods judge the segmented image or output image so as to indirectly assess the performance of algorithms. For applying empirical discrepancy methods, the reference image is necessary [42]. It can be obtained manually or automatically from the input image or segmenting image. The empirical discrepancy methods compare the segmented image or output image to the reference image and use their difference to assess the performance of algorithms.

Each method group has its own characteristics. In the following three subsections a brief description of the methods belonging to the three groups will be provided.

4.2.1 Analytical method

The analytical methods directly treat the segmentation algorithms themselves by considering the principles, requirements, utilities, and complexity of algorithms. Using analytical methods to evaluate the segmentation algorithms avoids the concrete implementation of these algorithms. The results are exempt from methodological influences in evaluation experiments [42]. However, not all properties of segmentation algorithms can be obtained by analytical studies. Until now, the analytical methods work only with some particular models or desirable properties of all algorithms.

One analytical method has been proposed by Liedtke et al [43]. They presented an evaluation study of several algorithms by taking into account the type and amount of a priori knowledge that has been incorporated into different segmentation algorithms. Such knowledge is usually heuristic information and different types of a priori knowledge are hardly comparable. The information provided by this method is rough and qualitative. On the other hand, not only "the amount of relevant a priori knowledge that can be incorporated into the segmentation algorithm is decisive for the reliability of the segmentation methods", but it is also very important for the performance of the algorithm how such a priori knowledge has been incorporated [44].

Other properties of segmentation algorithms that can be obtained by analysis include the processing strategy, processing complexity and efficiency and segmentation resolution.

These properties could be helpful for selecting suitable algorithms in particular applications.

4.2.2 Empirical goodness methods

The methods in these groups evaluate the performance of algorithms by judging the quality of segmented images. To carry out this work certain quality measures should be defined. Most measures are established according to human intuition [42] about what conditions should be satisfied by an "ideal" segmentation. In other words, the quality of segmented images is assessed by some "goodness" measures [42]. These methods characterize different segmentation algorithms by simply computing the goodness measures based on the segmented image without the a priori knowledge of the correct segmentation. Zhang [42] proposed different types of goodness measures such as goodness based on intra-region uniformity, goodness based on inter-region contrast and goodness based on region shape. Different types of goodness measures proposed by Zhang are discussed in the following subsections.

4.2.2.1 Goodness based on intra-region uniformity

Weszka and Rosenfield [45] proposed a threshold evaluation method that uses a busyness measure as the criterion to judge the thresholded images. To apply the busyness measure the authors assume that the images are composed of objects and background of compact shapes and are not strongly textured. Under these assumptions, the thresholded images should look smooth rather than busy. In practice, they compute the amount of busyness for a thresholded image by using the grey-level co-occurrence matrix representing the percentage of object background adjacencies. The lower the busyness, the smoother the thresholded images, the better the segmentation results and the higher the performance of applied algorithms.

Similar to Weszka and Rosenfield, Nazif and Levine also believe that an adequate segmentation should produce images having higher intra-region uniformity, which is related to the similarity of property about region elements [46]. The uniformity of a feature over a region can be computed on the basis of the variance of that feature evaluated at every pixel belonging to that region.

The intra-region uniformity, as a desired property of segmented images, can also be measured by the higher order local entropy based on information theory [47]. Pal and Pal proposed a thresholding method that maximizes the second order local entropy of the object and background regions [47].

4.2.2.2 Goodness based on inter-region contrast

In addition to intra-region uniformity, Levine and Nazif also believe that an adequate segmentation should also produce images having higher contrast across adjacent regions

[48]. Contrast is computed on the basis of the average values of features of adjacent regions. The contrast between two regions R_a and R_b is thus given by

$$C_{ab} = |f_a - f_b| \;/\; (f_a + f_b) = C_{ba} \qquad \ldots\ldots\ldots\ldots\ldots (4.1)$$

where f_a and f_b are the average grey level of regions R_a and R_b.

Each neighboring region R_a contributes to the summation with a value that is proportional to the adjacency between it and the region R_b for which the measure is computed. If the region R consists of a set of regions R_a, R_b, R_c..... R_n resulting from the segmentation of an input image, then two regions R_a and R_b are said to be adjacent if there exists at least a pair of points $p \in R_a$ and $q \in R_b$ which is four-connected. The measure for the region R_b is given by

$$C_b = \sum_{AdjR_a} p_{ab}\, C_{ab} \qquad \ldots\ldots\ldots (4.2)$$

Where p_{ab} is the region adjacency between two regions defined in [57]. It takes a value 1 if two regions are adjacent and a value 0 otherwise.

In order to obtain a single contrast measure for the whole area, a weighted sum of the contributions of each region in that area is computed. In general, the contrast measure for area α is given by

$$C_\alpha = \sum_{R_b \in \alpha} V_b\, C_b \;/\; \sum_{R_i \in \alpha} V_b \qquad \ldots\ldots\ldots\ldots (4.3)$$

Where V_b is the weight assigned to region R_b defined in [48].

The weight assigned to the contribution of each region to the contrast measure depends on its size, but not linearly. This contrast measure must be high to indicate a good segmentation.

4.2.2.3 Goodness based on region shape

Not only the grey level, but also the form of a segmented image can be taken into account to design goodness measures for satisfying human intuition about an "ideal" segmentation. A measure, called the shape measure SM, is used for the measurement of the shape of the object in the images. Sahoo et al [49] proposed a shape measure (SM) for evaluating several threshold selection algorithms for a given image, which is defined as:

$$SM = 1 / C\{\sum_{(x,y)} Sgn[f(x,y) - f_{s(x,y)}]q(x,y)Sgn[f(x,y) - T]\} \quad \text{........(4.4)}$$

where $f_{s(x,y)}$ is the average grey value of the neighborhood N (x, y) of a pixel located at (x, y) with grey level f (x, y) and gradient value q(x, y). T is the threshold value selected for segmentation, C is a normalization factor and Sgn() is the unit step function. Based on the calculated shape measure values for different threshold selection algorithms, the best threshold selection algorithm can be found and the segmentation produced by that algorithm would be considered as best segmentation among them. The shape measured value that selects the best threshold is considered as best shape.

4.2.3 Empirical discrepancy methods

In practical segmentation applications, some errors in the segmented image can be tolerated. On the other side, if the segmenting image is complex and the algorithm used is fully automatic, the error is inevitable [42]. The disparity between the actual segmented image and a correctly ideally segmented image (reference image) which is the best possible result, can be used to assess the performance of the algorithms. The methods in this group take into account the difference (measured by various discrepancy parameters) between the actually segmented and reference images, i.e. these methods try to determine how far the actually segmented image is from the reference image. A higher value of the discrepancy measure would imply a bigger error in the actually segmented image relative to the reference image and this indicates the lower performance of the applied segmentation algorithms.

In image encoding, the disparity between the original image and the decoded image has often been used to objectively assess the performance of coding algorithms. A commonly used discrepancy measure is the mean-square signal-to-noise ratio. Many other discrepancy measures have been proposed and used.

4.2.3.1 Discrepancy based on the number of mis-segmented pixels

Considering image segmentation as a pixel classification process, the percentage of pixels mis-classified is the discrepancy measure that comes most readily in mind [49]. Weszka and Rosenfield [45] used a similar approach to measure the difference between an "ideal" (correct) image and a thresholded image. Under the assumption that the image consists of objects and background each having a specified distribution of grey level, they compute, for any given threshold value, the probability of misclassifying an object pixel as background, vice versa. This probability in turn provides an index of segmentation results, which can be used for evaluating threshold selection algorithms. In their work, such a probability is minimized in the process of selecting an appropriate threshold.

The idea of computing discrepancy based on the number of error pixels is also reflected in some edge detection evaluation schemes. Such a measure could be readily extended to measure what fraction of the segmented object pixels were actually object pixels so as to evaluate the segmentation.

4.2.3.2 Discrepancy based on the position of mis-segmented pixels

The discrepancy measures based only on the number of mis-segmented pixels do not take into account the spatial information of these pixels [42]. It is thus possible that images segmented differently can have the same discrepancy measure values if these measures only count the number of mis-segmented pixels. To address this problem, some discrepancy measures based on pixel position error have been proposed.

One way is to use the distance between the mis-segmented pixel and the nearest pixel that actually belongs to the mis-segmented class [42]. Let N be the number of mis-segmented pixels for the whole image and d(i) be a distance metric from the ith mis-segmented pixel and the nearest pixel that actually is of the mis-classified class. A discrepancy measure based on this distance is defined by Yasnoff et al [50] as:

$$D = \sum_{i=1}^{s} d^2(i) \quad .... (4.5)$$

This measure is further normalazied (ND) to exempt the influence the image size and to give a suitable value range by

$$ND = 100\sqrt{DA} \quad .... (4.6)$$

Where A is the total number of pixels in the image (a measure of area).

In the evaluation of edge detectors a commonly used discrepancy measure is the mean square distance figure of merit (FOM) proposed by Pratt [51].

$$FOM = 1 / N \sum_{i=1}^{N} 1/(1 + p * d^2(i)) \quad .....(4.7)$$

Where $N = \max(N_i, N_a)$ and N_i and N_a denote the number of ideal and actual detected edge pixels respectively, d (i) denotes the distance between the ith detected edge pixel and its correct position and p is a scaling parameter. This measure has been shown insensitive to correlation in false alarms and missed edges.

4.2.3.3 Discrepancy based on the number of objects in the image

For perfect segmentation a necessary condition is that there should be an equal number of objects of each class in the reference image and the segmented image [42]. A substantial disagreement in the number of objects indicates a large discrepancy between the reference and segmented images. Yasnoff and Bacus [52] proposed to compute the object-count-agreement (OCA) based on probability theory. Let R_i be the number of objects of class I in the segmented image, the authors used the probability F_{OCA} that the two numbers R_i and S_i represent samples from the same distribution for measuring the OCA:

$$F_{OCA} = \int_{L}^{z} 1/(2^{M2}\,\Gamma(M-2)) * z^{(M-2)2}\, e^{-z2}\, dz \quad \ldots\ldots. (4.8)$$

In equation (4.6), $M = N\text{-}1$ denotes the number of degrees of freedom. $\Gamma()$ denotes the Gamma function and L can be computed by:

$$L = \sum_{t=1}^{N} (St - Rt)/(p * Rt)$$

where N is the number of object classes and p is a correlation parameter.

4.2.3.4 Discrepancy based on the feature values of segmented object

One fundamental question in image analysis is whether a measurement made on the objects from segmented images is as accurate as one made on the original images. According to this measure, a segmented image has the highest quality if the object features extracted from it precisely match the features in the reference image [42]. Here the reference image is not the original image, but is one that is ideally segmented from the original image. So, if the object features extracted from the actually segmented image perfectly match the features in the reference image, then the actually segmented image and the reference image would be same. In practice, an image has high quality if the decision made on it is unchanged from that made on the original image. The ultimate goal of image segmentation in the context of image analysis is to obtain the measurements of object features [53]. The accuracy of these measurements obtained from the segmented image with respect to the reference image provides useful discrepancy measures. This accuracy can be termed "ultimate measurement accuracy" [42] (UMA) to reflect the ultimate goal of segmentation. The UMA is feature dependent and denoted as UMA_f. Let R_f denote the feature value obtained from the reference image and S_f denote the feature value measured from the segmented image, the absolute UMA_f ($AUMA_f$) and relative UMA_f ($RUMA_f$) are defined as [54]:

$$AUMA_f = | R_f - S_f | \qquad \text{..........} \quad (4.9)$$

$$RUMA_f = (| R_f - S_f | / R_f) * 100 \qquad \text{......}(4.10)$$

Both $AUMA_f$ and $RUMA_f$ can represent a number of discrepancy measures when different object features are used.

4.3 Comparison of previous works

The three method groups for segmentation evaluation described in the above sections have their own characteristics. In the following, their advantages and limitations are discussed.

4.3.1 Generality for evaluation

One desirable property of an evaluation method is its generality to be applied for studying various properties of algorithms [42]. To apply analytical methods some formal models of an image should be first defined. The behavior of the algorithm on such an image can then be analysed (mathmatically) in terms of the parameters of the image and the algorithms [55]. Certain properties of segmentation algorithms can be easily obtained just by analysis, such as the processing strategy of the algorithms and the resolution of segmentation results. However, some other properties cannot be precisely analysed since no formal model exists. For instance, there is no quantitative measure for a priori knowledge about images that can be incorporated into segmentation algorithms [43]. In addition, there are methods that can only be applicable to certain segmentation algorithms. For instance, the method based on detection probability ratio is merely suitable for studying simple edge detection.

Empirical methods described in the previous sections are mainly used to study the correctness of segmentation algorithms [42] by taking into account the accuracy of segmentation results. One reason is that other properties of algorithms, such as computational cost have been partially overcome by the progress of technology. Another reason is that the accuracy of segmentation is often the primary concern in real applications and is difficult to study by analytical methods. From the point of view that only one property is studied, the empirical methods can be thought of as somewhat limited. However, most of them can be considered as relatively general, because they can evaluate different types of segmentation algorithms.

4.3.2 Qualitative versus quantitative and subjective versus objective

Two more properties of an evaluation method are the abilities to evaluate segmentation algorithms in a quantitative way and on an objective basis. Quantitative study can provide precise results reflecting the exactness of evaluation [54]. Objective study will exempt the influence of human factor and provide consistency and unbiased results. Generally, analytical methods are more ready to apply, but they often provide only qualitative properties of algorithms. Empirical methods are normally quantitative as the values of quality measures can be numerically computed. Among them, goodness methods based on subjective measures of image quality are less suitable for an objective evaluation of segmented algorithms. Discrepancy methods can be both objective and quantitative.

4.3.3 Consideration of segmentation applications

The effective use of domain –dependent knowledge in computer vision can help to make different processes reliable and efficient. To effectively evaluate segmentation algorithms, the consideration of segmentation applications in which algorithms are applied is also important.

The above three method groups are different in the extent to which they explicitly consider the applications for which the segmentation algorithms are used. At one extreme are the analytical studies that do not consider the nature and goal of application. The evaluation results depend only on the analysis of algorithms themselves. The empirical discrepancy methods, which take both the reference and segmented images into consideration, attempt to capture the application through the discrepancy measures. The need to have a reference forces the evaluation to be connected to the applications [54].

4.3.4 Common problems for most existing methods

There are still two main problems associated with most of existing evaluation methods.

(I) Each evaluation method determines the performance of algorithms according to certain criteria. If the same criterion used for segmentation is also used for evaluation then some biased results will be produced [54]. For example, the second order local entropy that was maximized for selecting threshold values in the new algorithm proposed by Pal and Pal [47] and was also computed for comparing the performance of this algorithm with that of other algorithms in Pal and Bhandari [56]. It is expected that the new algorithm should produce a high performance value. In many applications images are modeled as a mosaic of regions of uniform intensity corrupted by additive Gaussian white noise. Therefore, region homogeneity is a commonly used criterion for designing various segmentation algorithms. Using a goodness measure based on uniformity takes the same criterion for evaluation.

(II) To strengthen certain aspects in the quality measures, some scaling weighting parameters are often used. For example, the parameter p in FOM provides a relative penalty between smeared edges and isolated but offset edges, while the parameters p and q in FOC determine the contribution of the large deviation relative to a small deviation. There exists no suitable guideline or rule for choosing these parameters. In practice, they are often selected on the basis of human intuition or judgement. This turns a supposedly objective evaluation into one that is influenced by subjective factors.

4.4 New Approach for segmentation evaluation

Several evaluation techniques of segmented images and their limitations have been discussed in the previous section. Objective evaluation of segmented images is a difficult task and possibly the human being is the ultimate judge to make a qualitative assessment of outputs from different segmentation algorithms. A subjective (segmentation produced by human being) evaluation method is described in this section

Starting from an over-segmented image, a human observer looks at the original image, removes the weakest edge (subjectively) and set a value for that one. He then looks at the second weakest edge comparing with the original image, removes that one and sets another value for it. This process is continued until all edges in the picture are removed. After removing all edges from the region, a list of numbers has been found according to the order in which the edges were removed. Based on that list, individual edge criteria can be calculated objectively from the original image and their correlation with human performance calculated. Correlation is a method of determining a degree of association between two variables. In this method the subject provides ordinal values (first, second, third and so on) that are then represented by cardinal values (1,2,3 and so on) in order to calculate correlation between the human method and the algorithms. In fact, the quality measurements of the proposed segmentation algorithms are defined in terms of their correlation with human performance. This method gives a well-defined mapping from human judgement to a value for each segment boundary. This is possible because both the algorithm and the human are beginning from a segmented image.

4.5 Summary

In this chapter most methods for segmentation evaluation and comparison so far are reviewed. A new evaluation technique has been explained. Each method studied in this chapter has advantage and limitations. From an application point of view, those that belong to different groups are more complementary than competitive. Besides, the performance of segmentation algorithms is influenced by many factors, so only one evaluation method would be not enough to judge all properties of an algorithm and different methods should be combined.

Segmentation evaluation is indispensable for improving the performance of existing segmentation algorithms and for developing new powerful algorithms. This study attempts to stimulate the work in this direction. To make segmentation get off trial-and-error status further studies and more efforts for segmentation evaluation are needed.

Chapter 5 Experimental Results and Discussion

5.1 Introduction

In the previous two chapters of this thesis, several existing segmentation techniques and segmentation evaluation techniques have been discussed and a new technique for segmentation and segmentation evaluation have been presented. The capability of the new algorithm is demonstrated in this chapter. The next section shows some experimental results of the combined algorithms and a brief summary of the results is discussed in the last section.

5.2 Results

In order to investigate the performance of the algorithms, natural images and medical images were used. All images have a resolution of 256*256 and up to 256 different grey levels. For the time being, only a single pixel on each side of the edge is used for calculating the edge strength to save computation time. The original images together with their results are shown below.

Figure 5.1 shows a natural image (lenna256). An over-segmented image is produced by using region-growing method, which is shown in figure 5.2. Figure 5.3 shows the candidate edges that come from the over-segmented image. The newly developed algorithms have been applied on this image shown in figure 5.3.

Figure 5.1: Original image (lenna256)

In order to investigate the performance of the combined algorithms some of the edge criteria have been applied individually to the input image. Figure 5.4 shows an example output for a threshold value 16, when only the edge strength criterion is applied. It preserves only the regions that are separated by a strong boundary from their neighboring regions, that is, if the average intensity difference across the edge in the original image is higher than the threshold then that edge is preserved. Otherwise that edge has been removed from the output. Figure 5.5 shows another example output for edge strength for a threshold value 30. This shows that when the threshold value is changed from 16 to 30, the output is changed, because the average intensity differences are less than this threshold value. Similarly, figure 5.6 shows the example output when the edge straightness criterion is applied. In this case, the threshold value was 17. When the straightness criterion is applied alone, it preserves only the straight portion of the regions. As a result, the example output of straightness preserves some regions that are not significant in segmentation but straight. Figure 5.7 shows another example output for edge straightness when threshold value 23 is applied. The effect of changing threshold in this case is to lose some regions in the lips, nose, eyes and hat in lenna256, since these regions are made by short straight lines. Other edge criteria can also be applied individually.

Figure 5.2: Over-segmented image produced by region growing method

Figure 5.3: Candidate edges from the segmented image of figure 5.2

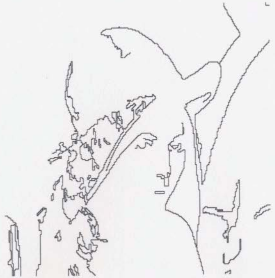

Figure 5.4: Example output of edge strength for threshold 16

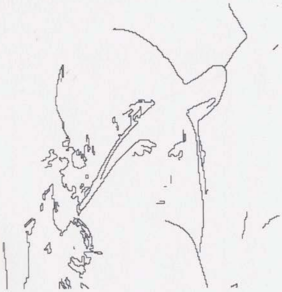

Figure 5.5: Example output of edge strength for threshold 30

Figure 5.6: Example output of straightness for threshold 17

Figure 5.7: Example output of straightness for threshold 23

Figure 5.8: Example output of subtraction technique when first weakest edge is removed

Figure 5.9: First weakest edge

After combining all the edge criteria, the subtraction technique and the addition technique were investigated. Figure 5.8 shows the example output of subtraction technique when the first weakest edge from the picture is removed and figure 5.9 shows the first weakest edge in the picture alone. This appears as a dot. Figure 5.10 shows the example output of subtraction techniques for a threshold value 0.55. The weighting factors for edge strength, edge straightness, edge smoothness and edge continuity are 0.32, 0.05, 2.5 and 0.1 respectively. In this output image there are some regions that have been preserved, but are not significant in segmentation, such as the regions in her hair and her shoulder. Figure 5.11 shows another example output for subtraction technique for a threshold value 0.65. The weighting factors remained same in this case. The weighting factors are set by visual observation. Every time the weighting factors are set, the output is observed and compared with the original image to check whether the segmented output is properly segmented or not. If the output is not properly segmented, the weighting factors are changed by repeated observations. Proper segmentation was not found in this case but it was attempted to find an accurate segmentation. With the increment of threshold value, some regions in the lips, cheek and some straight and strong regions are lost.

Figure 5.12 shows an example output of subtraction technique when the edge straightness factor is changed to 0.1 where all other weighting factors and threshold remained same as in figure 5.10. Figure 5.13 shows an example output of subtraction technique when edge smoothness factor is changed to 4. It is observed that increasing the straightness factor preserves some straight portions of the picture in the hat and shoulder of lenna256 that are not significant and are not shown in figure 5.10. On the other hand, when the smoothness factor increases, the output picture could not preserve some important regions in the lips,

Figure 5.10: Example output of subtraction technique for threshold 0.55

Figure 5.11: Example output of subtraction technique for threshold 0.65

Figure 5.12: Example output of subtraction technique when edge straightness factor is increased to 0.1

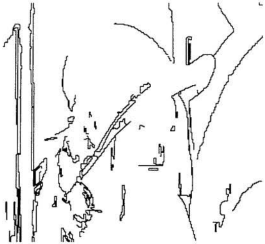

Figure 5.13: Example output of subtraction technique when edge smoothness factor is increased to 4

hat and some strong regions in lenna256 but it lost some insignificant regions in the hair as well. When the continuity factor is increased, it preserved the regions shown in figure 5.10 but it added some additional regions in the eyes, shoulder and hat. Figure 5.14 shows the example output when the continuity factor increases to 0.45.

The above demonstrations show that the various edge criteria are independently controllable and that their effect is in line with expectations. Similar results have been observed on other images.

Figure 5.14: Example output of subtraction technique when edge continuity factor is increased to 0.45

The addition technique that starts from the strongest edge in the region is also applied to the input image shown in figure 5.3. Figure 5.15 shows the first strongest edge (based on the weighted combination of all criteria) among all edges in the picture shown in figure 5.3. After finding the first strongest edge, it grows out from both ends until a closed loop or a boundary is reached as shown in figure 5.16. Figure 5.17 shows the example output of the addition technique for a threshold value 2. The weighting factors for edge strength, edge straightness, edge smoothness and edge continuity are 0.35, 0.18, 2.25 and 0.1 respectively. This output image preserves all significant regions in segmentation though it preserves a few regions that are not significant. When the threshold value is increased to 4, some strong regions are lost but some insignificant regions in the hat are added as shown in figure 5.18.

Figure 5.15: Example output of first strongest edge

Figure 5.16: Example output of first strongest edge when it is grown from both ends after addition technique

Figure 5.17: Example output of addition technique for threshold 2

Figure 5.18: Example output of addition technique for threshold 4

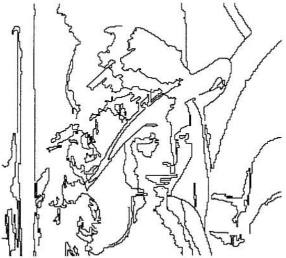

Figure 5.19: Example output for addition technique when edge straightness factor is increased to 0.3

Figure 5.20: Example output of addition technique when edge smoothness factor is increased to 4.5

Figure 5.21: Example output of addition technique when continuity factor is increased to 0.3

Figure 5.19 shows the example output of the addition technique when the edge straightness factor is increased to 0.3 keeping all other parameters same as in figure 5.17. When the edge straightness factor is increased to 0.3, some insignificant regions in the eyes, hat and shoulder are added. With the increment of edge smoothness factor, some important regions in the eyes and some strong edges are lost as shown in figure 5.20. When the continuity factor is increased to 0.3, the output picture shown in figure 5.21 preserved some additional regions that are not meaningful in segmentation.

It is observed that different edge criteria add different regions in the output picture according to their weighting factors. The weighting factors should be chosen carefully, considering the importance of the various features for particular applications.

In order to evaluate the performance of the proposed segmentation techniques a subjective measure, described in section 4.4, has been investigated. Figure 5.22 shows an example output of segmented image with edges removed by a human. In this picture, only fifty edges were removed by the human. Figure 5.23 shows another example output of segmented image with more edges removed than that of figure 5.22, by the same subject. Both images were produced by a single subject. A full method would involve controlled subjective tests with several subjects but here the purpose was to prove the concept. Therefore a single subject (the author) performed the boundary removals on two separate occasions. The orderings from the two cases were then compared and found to be consistent. In this method the subject provides ordinal values (first, second, third and so on) and these are compared with the ordinal values yielded by the algorithm. To do this, they were represented by cardinal values (1, 2, 3 and so on) in order to calculate correlation.

Figure 5.22: Example output of segmented image produced by a human

Figure 5.23: Another example output of segmented image produced by a human

Figure 5.24 shows the correlation graph between the *edge goodness* value (obtained subjectively) and *edge strength* that was obtained objectively. The correlation coefficient for this case is 0.41; that is, the edge strength is correlated with the subjective measure.

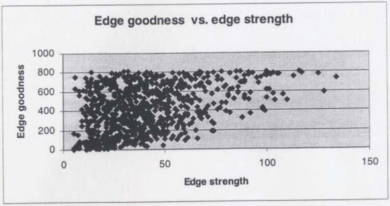

Figure 5.24: Correlation graph - edge goodness vs. edge strength

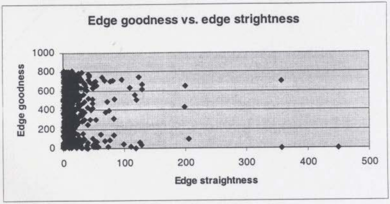

Figure 5.25: Correlation graph - edge goodness vs. edge straightness

Figure 5.25 shows the correlation graph between the subjective measure - *edge goodness* and the objective measure *edge straightness*. The correlation coefficient for this case is

0.07. Figure 5.26 and figure 5.27 show the correlation graphs between the subjective measure *edge goodness* and objective values *edge smoothness* and *edge continuity* respectively. In these cases, the correlation coefficients are 0.11 and 0.18. It is observed that there is a little correlation between human method and the algorithms for edge straightness, edge smoothness and edge continuity. The human has selected the edges for removal by visual observation. This selection procedure is dominated by edge strength rather than considering all criteria with equal importance. That is why there is a little correlation between human method and the algorithms for edge straightness, edge smoothness and edge continuity.

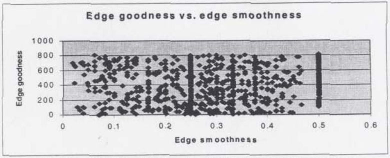

Figure 5.26: Correlation graph - edge goodness vs. edge smoothness

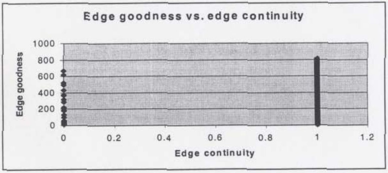

Figure 5.27: Correlation graph – edge goodness vs. edge continuity

Figure 5.28: Less segmented image of lenna256 than that of figure 5.2

Figure 5.29: Candidate edges of figure 5.28

The proposed segmentation algorithms have also been applied to the same image (lenna256) but less segmented than that shown in figure 5.2. Figure 5.28 shows this less segmented image and figure 5.29 shows the candidate edges that come from it.

Some of the edge criteria are also applied individually to this image and their outputs have also been observed. Figure 5.30 shows the example output when the edge criterion – edge strength is applied alone with threshold value 20. Figure 5.31 shows the example output for straightness alone. For this case, threshold value was 18. Output for edge strength preserves the regions that are significant based on their intensities but output for edge straightness preserves the straight portions of the picture though some of them are not supposed to be present in the segmented image.

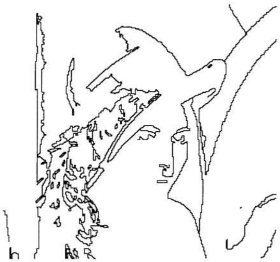

Figure 5.30: Example output for edge strength only for threshold 20

Figure 5.31: Example output for straightness only for threshold 18

Figure 5.32: Example output of subtraction technique when the first weakest edge is removed

After combining all edge criteria (edge strength, edge straightness, edge smoothness and edge continuity) the subtraction technique is applied to this image. Figure 5.33 shows the example output of the subtraction technique when the first weakest edge is removed from the region and figure 5.9 shows the first weakest edge alone. It is seen that the first weakest edge removed is the same for both input images (over segmented and less segmented). Figure 5.33 shows the final output of subtraction technique for threshold value 1.5. For this output image, the weighting factors for edge strength, edge straightness, edge smoothness and edge continuity are 0.25, 0.8, 3.5 and 0.25 respectively.

Figure 5.33: Example output of subtraction technique with less segmented image

When the subtraction technique was applied to the more segmented image shown in figure 5.3, it preserved the regions such as those in the hat and shoulder and some other regions that are separated by meaningful boundaries. These regions have significant values for edge strength, edge straightness, edge smoothness and edge continuity. So, after combining all these criteria, the combined edge strength became significant and these regions were preserved. Similar regions were also preserved when the subtraction technique was applied to the less segmented image shown in figure 5.29. But there are some small regions in lenna256, like the regions in the eyes, nose and lips, that are important for segmentation. The combined edge strength for these regions do not offer significant values. But when the subtraction technique was applied to this over-segmented image in figure 5.3 it preserved these small but important regions. On the other hand, when the subtraction technique was applied to a less segmented image shown in figure 5.29, it could not preserve the sitter's (lenna256) eyes, nose and lips as preserved in figure 5.10.

Figure 5.34: Example output of addition technique with less segmented image

Like the subtraction technique, the addition technique has also been applied to the less segmented image of figure 5.29. Figure 5.34 shows the example output of addition technique for threshold 1.5 and the weighting factors are 0.2, 0.7, 3 and 0.25 respectively.

The segmented image shown in figure 5.29 produced by the proposed algorithms has been evaluated. A subjective measure has also been used in this case. Figure 5.35, 5.36, 5.37 and 5.38 show correlation graphs between the subjective value obtained by a human observer and objective values - edge strength, edge straightness, edge smoothness and edge continuity respectively. The correlation coefficient for each of the above cases has been calculated. The correlation coefficients are 0.35, 0.06, 0.18 and 0.17 respectively. It is apparent that human method and the algorithm are correlated for edge strength criterion. But there is a little correlation between human method and the algorithm for edge straightness, edge smoothness and edge continuity. The reason for this little correlation in this case is similar to that of an over-segmented image.

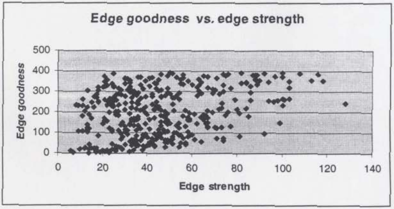

Figure 5.35: Correlation graph – edge goodness vs. edge strength

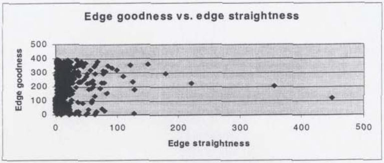

Figure 5.36: Correlation graph - edge goodness vs. edge straightness

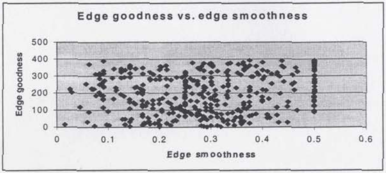

Figure 5.37: Correlation graph – edge goodness vs. edge smoothness

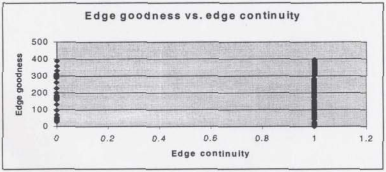

Figure 5.38: Correlation graph – edge goodness vs. edge continuity

The proposed algorithms are also applied to a medical image (mri256). Figure 5.39 shows the original image (mri256) and figure 5.40 shows the starting segmented image. The candidate edges that from this segmented image are shown in figure 5.41.

Figure 5.39: Original medical image (mri256)

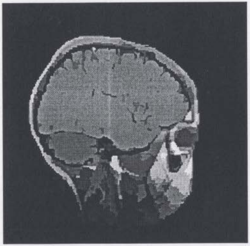

Figure 5.40: Segmented image (mri256)

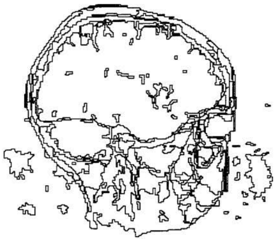

Figure 5.41: Candidate edges that come from figure 5.40

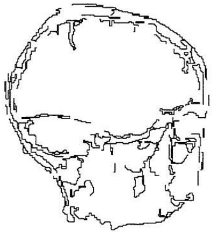

Figure 5.42: Example output of subtraction technique for weighting factors 0.25, 0.9, 3 and 0.25 respectively

Figure 5.43: Example output of subtraction technique for weighting factors 0.25, 0.6, 3 and 0.25 respectively

Figure 5.44: Example output of addition technique for weighting factors 0.2, 0.7, 3 and 0.25 respectively

Figure 5.45: Example output of addition technique for weighting factors 0.2, 0.7, 5 and 0.25 respectively

The combined algorithms have been applied to the image shown in figure 5.41. Figure 5.42 shows an example output of the subtraction technique for threshold 1.75 and the weighting factors for edge strength, edge straightness, edge smoothness and edge continuity were 0.25, 0.9, 3 and 0.25 respectively. When the edge straightness factor is decreased to 0.6 and edge smoothness factor is increased to 4, the example output of the subtraction technique removed some non-smooth regions from the picture as shown in figure 5.43. Figure 5.44 shows an example output of addition technique for threshold 1.5. The weighting factors for this output were 0.2, 0.7, 3 and 0.25. Figure 5.45 shows another example output of addition technique when the edge smoothness factor is increased to 5 and all other parameters remained same as in figure 5.44.

5.3 Comparison with pure region-based scheme

The addition and subtraction algorithms start with an over-segmented image containing L edges. The processing of each algorithm produces an image with N edges, where L is greater than N. On the other hand, it is possible to generate an output image with N edges using region-based segmentation alone. The key question is, does the proposed algorithm generate a better, more useful and more reliable segmentation than that obtained using region-based segmentation alone. To answer this question, several output images that have been obtained by region growing and by the new algorithm are discussed below.

Figure 5.46 shows a segmented image produced by region-based segmentation. The candidate edges that come from this segmented image are shown in figure 5.47. The number of edges in this image is 275. When the addition and the subtraction techniques are applied to an over-segmented image (with 856 edges) shown in figure 5.3, the resulting outputs preserve 275 edges shown in figure 5.48 and 5.49. The output of the region-based method removes some insignificant regions and preserves 275 edges, but it could not remove some structurally unimportant regions in the eyes, shoulder and hat of lenna256. On the other hand, the output of the addition technique preserves some structurally important regions and removes insignificant regions though it preserves a few insignificant regions, like the region in her forehead and loses a few strong boundaries. The output of subtraction technique removes most of the insignificant regions and

Figure 5.46: Segmented image produced by region-based method alone

Figure 5.47: 275 candidate edges that come from figure 5.46

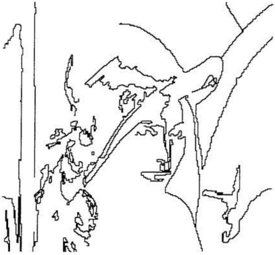

Figure 5.48: Example output of addition technique with 275 edges (from 856 edges)

Figure 5.49: Example output of subtraction technique with 275 edges (from 856 edges)

Figure 5.50: Example output of addition technique with 275 edges (from 423 edges)

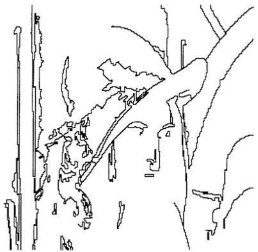

Figure 5.51: Example output of subtraction technique with 275 edges (from 423 edges)

preserves all structurally significant regions though it preserves a few unimportant regions, like the regions in the forehead and shoulder.

To justify the acceptance of this algorithm, the addition and the subtraction techniques are also applied to a less segmented version of the same image, which is shown in figure 5.29. The number of edges in this image is 423. Figure 5.50 shows an example output when the addition technique is applied to figure 5.29. This output image contains 275 edges. Figure 5.51 shows the example output with 275 edges when the subtraction technique is applied to figure 5.29. The output of the addition technique preserves structurally significant regions and removes unimportant regions, though it preserves some unimportant regions in her hat and shoulder. Similarly, the output of the subtraction technique preserves most significant regions. It also preserves some insignificant regions and boundaries.

The proposed algorithm starts with an over-segmented image where the region growing method ends. The number of segments in the picture can be changed by setting different threshold values in the region growing method. The results demonstrated in this section and in the previous section show that the proposed algorithm produces better results starting with an over-segmented image containing 856 edges. These 856 edges are produced by region growing method for a threshold value 15.

It is observed that though it is possible to produce an output image with a certain number of edges using region-based segmentation only, this does not preserve all significant regions and remove unimportant regions. On the other hand, the proposed algorithm can achieve this goal as the above demonstrations show.

5.4 Discussion

To investigate the performance of the proposed algorithms, natural images and medical images are used. The proposed algorithms – subtraction and addition techniques – have been applied to an over-segmented image first and then applied to a less segmented version of the same image. When the subtraction technique is applied to an over-segmented image, it removes some regions from the image that are not separated by meaningful boundaries and preserves some regions that are important for segmentation such as the regions in the eyes, nose and lips in lenna256. It also preserves a few non-structural regions like the regions in her forehead and hair. But when the subtraction technique is applied to a less segmented version of the same image, it does not preserve some important regions such as eyes, nose and lips while it preserves some insignificant regions. So, it is observed from the above results that the subtraction technique produces better segmentation for natural and medical images when it is applied to an over-segmented version of that image.

Like the subtraction technique, the addition technique is also applied to the over-segmented and a less segmented version of the same image. In the over-segmented image it preserved all significant regions though it preserved a few regions that are not important like the shoulder and forehead artifactual regions. But in a less segmented version of the same image, though the addition technique preserved these insignificant regions, it could not preserve some meaningful regions such as the eyes, nose and lips.

Different correlation coefficients are obtained for an over-segmented and less segmented version of the same image. For an over-segmented image, the correlation coefficients for edge strength, edge straightness, edge smoothness and edge continuity are 0.41, 0.07, 0.12 and 0.18 whereas, for a less segmented image these are 0.35, 0.06, 0.11 and 0.17 respectively. The correlation coefficient for edge strength is higher for an over-segmented image than a less segmented version of the same image. This shows that human method is correlated with the algorithm for edge strength criterion. On the other hand, the correlation coefficients of edge straightness, edge smoothness and edge continuity for the over-segmented image are higher than for the less segmented image, but are all very small. From the results shown above, it is apparent that the proposed algorithms produce better results from an over-segmented image and when the algorithms are applied to an over-segmented and a less segmented image, they stop with a comparable number of edges left after processing. Different edge criteria – edge strength, edge straightness, edge smoothness and edge continuity are used in the proposed algorithms. The above calculated correlation coefficients show that edge strength is the only criterion that is correlated with the subjective measure. Other edge criteria have very little correlation. But there is a limitation to this subjective measure, since it is related to the involvement of visual perception. In fact, the quality measurements are defined in terms of their correlation with human performance. Two problems arise here. First, the results heavily rely on the judgement of a particular observer and it may vary from one observer to another. Second, although such a procedure may be acceptable for certain image processing tasks such as image enhancement, it is not suitable for an automatic image analysis where the objective judgement is mandatory. Subjective human judgement frequently differs from objective computer measurements. Finally, experimental results are important in order to illustrate the effectiveness of proposed segmentation methods.

Chapter 6 Conclusions and further improvements

This thesis reviews and summarizes some existing approaches to image segmentation. So far, image segmentation techniques are application dependent. Selection of an appropriate segmentation technique largely depends on the type of images and application areas. Semantic and a priori information about the type of images are critical to the solution of the segmentation problem.

A new image segmentation method based on region-growing has been presented in this thesis. The region-growing method provides unambiguous segmentation but often divides regions that are not clearly separated by a strong boundary. The goal of this thesis was to produce a better segmentation using region-growing methods. For this purpose, an attempt has been made to include more information about an edge. Different edge criteria – edge strength, edge straightness, edge smoothness and edge continuity have been used. The combined algorithm aimed to eliminate the regions that are not clearly separated by a meaningful boundary in the segmented image. The proposed algorithms acted better on an over-segmented image than an under segmented image. The addition and the subtraction techniques produced better segmentation on an over-segmented image though a few small curly regions produced by less significant regions were not eliminated. But both techniques could not preserve some significant regions and produced some insignificant regions when applied to an under segmented image.

Several existing segmentation evaluation and comparison methods have also been reviewed in this thesis and a new segmentation evaluation method has been presented. A subjective evaluation method has been demonstrated and a correlation was found between subjective judgement and the segmentation generated automatically. Since the quality measurements were defined in terms of their correlation with human performance, the results heavily rely on the judgement of a particular observer. Such a procedure may be acceptable for certain image processing tasks such as image enhancement, but it is not suitable for automatic image analysis where the objective judgement is mandatory.

Various improvements can be made with this algorithm. In the binary tree method (described in chapter 3), every time at one end of the seed, the strongest tree is found and added to the seed. Before comparing the trees at one end, the straightness, smoothness and continuity criteria of each tree can be checked again before adding to the strength. This can eliminate adding small curly regions to the seed.

References

[1] N. R. Pal and S. K. Pal, "A review on image segmentation techniques", *Pattern Recognition*, Vol. 26 No. 9 pp 1277-1294, 1993.

[2] A. Rosenfield and A. C Kak, *Digital Picture Processing*, Academic Press, New York, 1982.

[3] E. L. Hall, *Computer Image Processing and Recognition*, Academic Press, New York, 1979.

[4] N. R. Pal and S. K. Pal, "Image model, Poisson distribution and object extraction", *International Journal on Pattern Recognition and Artificial Intelligence*, Vol. 5, pp 459-483, 1991.

[5] R. M. Haralick, "Performance characterization in computer vision", *CVGIP-IU* 60, pp 245-249, 1994.

[6] M. I. Chowdhury and J. A Robinson, "Improving image segmentation using edge information", *Canadian Conference on Electrical and Computer Engineering 2000*, Halifax, Canada.

[7] Rolf Adams and Leanne Bischof, "Seeded Region Growing", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 16, No. 6 pp 641-647, June 1994.

[8] Scott E Umbaugh, *Computer Vision Image Processing: A Practical Approach using CVIPtools*, Prentice-Hall, Inc. 1998.

[9] Anil K. Jain, *Fundamental of Digital Image Processing*, Prentice-Hall,Inc. 1989.

[10] Rafael C. Gonzalez, Paul Wintz, *Digital Image Processing*, Addison-Wesly Publishing Company, 1987.

[11] K. S. Fu and J. K. Mui, "A survey on image segmentation", *Pattern Recognition*, Vol. 13, pp 3-16, 1981.

[12] R. M. Haralick and L. G. Shapiro, "Survey, image segmentation techniques", *Computer Vision Graphics Image Processing*, Vol. 29, pp 100-132, 1985.

[13] P. K. Sahoo, S. Soltani, A. K. C. Wong and Y. C. Chen, "A survey of thresholding techniques", *Computer Vision Graphics Image Processing*, Vol. 41 pp 233-260, 1988.

[14] S. D. Yanowitz and A. M. Bruckstein, "A new method for image segmentation", *Computer Vision Graphics Image Processing*, Vol. 46, pp 82-95, 1989.

[15] T. Pun, "A new method for grey level picture thresholding using the entropy of histogram", *Signal Processing*, Vol. 2 pp 223-237, 1980.

[16] J. N. Kapur, P. K. Sahoo and A. C. K. Wong, "A new method for grey level picture thresholding using the entropy of histogram", *Computer Graphics Vision Image Processing*, Vol. 29, pp 273-285, 1985.

[17] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images", *IEEE Transactions on Pattern Analysis and machine Intelligence*, Vol. 6, pp 721-741, 1984.

[18] P. C. Besl and R. C. Jain, "Segmentation using variable surface fitting", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 10, pp 167-192, 1988.

[19] R. Hoffman and A. K. Jain, "Segmentation and classification of range images", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 9, pp 608-620, 1987.

[20] N. Yokoya and M. D. Levine, "Range image segmentation based on differntial geometry: a hybrid approach", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 11(6), pp 643-694, 1989.

[21] L. S. Devis, "A survey of edge detection techniques", *Computer Graphics Image Processing*, Vol. 4, pp 248-270, 1975.

[22] J. F. Canny, "A computational approach to edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 8, pp 679-698, 1986.

[23] M. Gokeman and C. C. Li, "Edge detection with iteratively refined regularization", *Proceedings of 10^{th} ICPR*, pp 690-693, 1990.

[24] J. C Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms,* Plenum Press, New York, 1981.

[25] S. K. Pal and R. A. King, "Image enhancement using fuzzy sets", *Electronics letter,* Vol. 16, pp 376-378, 1980.

[26] M. Trivedi and J. C. Bezdek, "Low-level segmentation of aerial images with fuzzy clustering", *IEEE Transactions on System Man and Cybernetics*, Vol. 16(4), pp 589-598, 1986.

[27] M. Osteaux, K. De Meirleir and M. Shahabpour, eds, "Magnetic Resonance Imaging and Spectroscopy in Sports Medicine", *Springer, Berlin*, 1991.

[28] W. Doyle, "Operations useful for similarity-invariant pattern recognition", *J. Ass. Computer*, March 9, pp 259-267, 1962.

[29] R. J. Yoo and T. S. Huang, "Image segmentation by unsupervised clustering and its applications", *TR-EE 78-19, Purdue University, West Lafayette, Indiana* 1978.

[30] M. Hucckel, "An operator which locates edge on digital pictures", *J. Ass. Computer,* March 18, pp 113-125, 1971.

[31] H. Derin, H. Elliott, R. Cristi and D. Geman, "Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 6, pp 707-720, 1984.

[32] W. E. Blanz and S. L. Gish, "A connectionist classifier architecture applied to image segmentation", *Proceedings of 10^{th} ICPR,* pp 272-277, 1990.

[33] N. Babaguchi, K. Yamada, K. Kise and T. Tezuka, "Connectionist model binarization", *Proceedings of 10^{th} ICPR,* pp 51-56, 1990.

[34] S. E. Reichenbach, S. K. Park and R. A. Gartenbarg, "Optimal, small kernels for edge detection", *Proceedings of 10^{th} ICPR,* pp 57-63, 1990.

[35] M. Bajcsy, M. Mintz and E. Liebman, "A common framework for edge detection and region growing", *University of Pennsylvania GRASP Lab. Technical Report* 61, 1986.

[36] H. L. Anderson, R. Bajcsy and M. Mintz, "A modular feedback system for image segmentation", *University of Pennsylvania GRASP Lab. Technical Report* 110, 1997.

[37] Theo Pavlidis and Yuh-Tay Liow, "Integrating region growing and edge detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 12, No. 3, March 1990.

[38] S. Grinakar, "Edge based segmentation and texture separation", *Proceedings of 5^{th} International Conference on pattern Recognition, Miami Beach, FL,* pp 554-557, 1980.

[39] U. Montanari, "On the optimal detection of curves in noisy pictures", *Communication ACM,* Vol. 14, No. 5, pp 335-345, May 1971.

[40] S. Hojjatoleslami and J. Kittler, "Region growing: a new approach", *IEEE Transactions on Image Processing*, Vol. 7, No. 7, July 1998.

[41] J. Kittler, "A locally sensitive method for cluster analysis", *Pattern Recognition*, Vol. 8, pp 23-33, 1976.

[42] Y. J Zhang, "A survey on evaluation methods for image segmentation", *Pattern Recognition*, Vol. 29, pp 1335-1346, 1996.

[43] C. F. Liedtke, T. Gahm, Kapper and B. Aeikens, "Segmentation of microscopic cell scene", *AQCH 9*, pp 197-211, 1987.

[44] Y. J. Zhang and J. J. Garbrands, "Transition region determination based thresholding", *Pattern Recognition Letter*, Vol. 12, pp 13-23, 1991.

[45] J. S. Weszka and A. Rosenfield, "Threshold evaluation techniques", *IEEE Transactions SMC* 8, pp 622-629, 1978.

[46] A. M. Nazif and D. M. Levine, "Low-level image segmentation: an expert system", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, pp 555-577, 1984.

[47] N. R. Pal and S. K. Pal, "Entropy thresholding", *Signal Processing*, Vol. 16, pp 97-108, 1989.

[48] M. D. Levine and A. Nazif, "Dynamic measurements of computer generated image segmentations", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 7, pp 155-164, 1985.

[49] P. K. Sahoo, S, Soltani, A. K. C. Wong and Y. C. Chen, "A survey of thresholding techniques", *CVGIP* 41, pp 233-260, 1988.

[50] W. A. Yasnoff, J. K. Mui and J. W. Bacus, "Error measures for scene segmentation", *Pattern Recognition*, Vol. 9, pp 217-231, 1977.

[51] W. K. Pratt, *Digital Image Processing*, John Wiley and Sons, New York 1978.

[52] W. A. Yasnoff and J. W. Bacus, "Scene segmentation algorithm development using error measures", *AOCH 6*, pp 45-58, 1984.

[53] M. D. Levine and A. Nazif, "An experimental rule based system for testing low level segmentation strategies, in Multi-Computers and Image Processing: Algorithms and Programs", *K. Preston and L. Uhr, eds, pp 149-160. Academic Press, New York* 1982.

[54] Y. J Zhang and J. J. Gerbrands, "Objective and quantitative segmentation evaluation and comparison", *Signal Processing*, Vol. 39, pp 43-54, 1994.

[55] L. J. Kitchen and J. A. Main, "The effect of spatial discretization on the magnitude and direction response of simple differential edge operators on a step edge", *CVGIP* 47, 243-258, 1989.

[56] N. R. Pal and D. Bhandari, "Image thresholding: some new techniques", *Signal Processing*, Vol. 33, pp 139-158, 1993.

[57] M. Suk and S. Oh, "Region Adjacency and its application to object detection", *Pattern Recognition*, Vol. 19, pp 161-167, 1986.

[58] N. Otsu, "A threshold selection method from grey level histograms", *IEEE Transaction on System Man Cybernetics* 9, 62-66, 1979.