p - 2566

IMAGE SEGMENTATION BASED WATERSHED TRANSFORM USING TEXTURE GRADIENT

By

Sini Antony Alakaparambil

Reg.No:0720108019

of

KUMARAGURU COLLEGE OF TECHNOLOGY

COIMBATORE-641006

(An Autonomous Institution affiliated to Anna University Coimbatore)

MAHATMA GANDHI KCT CBE-6 CENTRAL LIBRARY

A PROJECT REPORT

Submitted to the

FACULTY OF INFORMATION AND COMMUNICATION ENGINEERING

In partial fulfillment of the requirements

For the award of the degree

of

MASTER OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

MAY, 2009

BONAFIDE CERTIFICATE

Certified that this project report entitled "IMAGE SEGMENTATION BASED WATERSHED TRANSFORM USING TEXTURE GRADIENT" is the bonafide work of Ms.SINI ANTONY ALKAPARAMBIL, who carried out the research under my supervision. Certified further, that to best of my knowledge the work reported here in does not from any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

K. Sivan Arul Selvan	S. Thangasamy
Signature of the Guide	**Head of the Department**
Mr.K.Sivan Arul Selvan M.E	**Dr.S.Thangasamy, Ph.D.,**
Department of Computer	Department of computer
Science and Engineering	Science and Engineering

The candidate with **University Register No. 0720108019** was examined by us in the project viva-voce examination held on 04.05.09

4/5/09	4/5/09
INTERNAL EXAMINER	**EXTERNAL EXAMINER**

ई.आर.एण्ड.डी.सी.आई. इनस्टिटयूट ऑफ टेक्नॉलजि
ER&DCI INSTITUTE OF TECHNOLOGY
प्रगत संगणन विकास केन्द्र द्वारा संचालित
An institution Managed by
CENTRE FOR DEVELOPMENT OF ADVANCED COMPUTING
(संचार एवं सूचना प्रौद्योगिकी मंत्रालय कि वैज्ञानिक संस्था, भारत सरकार)
(A Scientific Society of the Ministry of Communications and Information Technology, Govt. of India)

ER & DCI - IT
डा.पे.सः ६५२० वेल्लयंबलम
तिरुवनन्तपुरम ६९५ ०३३, इंडिया
फ़ोन / Tel: +91-471-2726106, 2723333
फ़ैक्स / Fax: +91-471-2723456, 2722230
P.B. No. 6520, Vellayambalam
Email : erdciit@cdactvm.in
http://www.erdciit.ac.in

ER&DCI-IT/M.Tech/Project 08/04/09

CERTIFICATE

Certified that **Ms.Sini Antony Alakaparambil**, Kumaraguru College of Technology, Coimbatore, has successfully completed the project titled **"Image Segmentation Using Texture Gradient Based Watershed Transform"** at the centre during the period December 2008 to April 2009, towards the partial fulfillment of the requirements for the award of M.Tech Degree in Computer Science & Engineering in Kumaraguru College of Technology, Coimbatore,

प्रिंसीपल / PRINCIPAL
ई.आर. एण्ड डी.सी.आई / ER & DCI
प्रौद्योगिकी संस्थान / Institute of Technology
तिरुवनंतपुरम / Thiruvananthapuram

Ms.Roshni V S
Principal
(Project Guide)

Departments of Computer Science and Engineering & Information Technology

CSI COLLEGE OF ENGINEERING

CET'09

Ketti, The Nilgiris- 643 215

HIRD NATIONAL CONFERENCE ON EMERGING TECHNOLOGIES - 2009

CERTIFICATE

This is to certify that ~~Dr.~~/~~Prof.~~/~~Mr.~~/Ms. SINI ANTONY ALAKA PARAMBIL of KUMARAGURU COLLEGE OF TECHNOLOGY has participated in **CET '09** and presented a paper entitled IMAGE SEGMENTATION BASED WATERSHED TRANSFORM USING TEXTURE GRADIENT at the conference held on 11th March 2009.

Authors: Ms. Sini Antony Alakaparambi)

C.S.I.C.E

O-ORDINATOR | CONVENER | S. T. Jeyaselan PRINCIPAL

ABSTRACT

The segmentation of images into meaningful and homogenous regions is a key method for image analysis within applications such as content based retrieval. The watershed transform is a well established tool for the segmentation of images.

However, watershed segmentation is often not effective for textured image regions that are perceptually homogeneous. In order to properly segment such regions the concept of the "texture gradient" is now introduced.

Texture information and its gradient are extracted using a novel non decimated form of a complex wavelet transform. A novel marker location algorithm is subsequently used to locate significant homogeneous textured or non textured regions. A marker driven watershed transform is then used to properly segment the identified regions. The combined algorithm produces effective texture and intensity based segmentation for the application to content based image retrieval.

ஆய்வுச்சுருக்கம்

உருவங்களை அர்த்தமுள்ள மற்றும் ஒத்த பரப்புகளாக பகுக்கின்ற பணியானது உருவங்களை பகுத்தாய்கின்ற பணியில் முக்கியப் பங்கு வகிக்கின்றது.இது பொருள் சார்ந்த தகவல் பெறுதலின் பயனுள்ள பணியாகும். வாடர் ஷெட் செக்மென்டேஷன் முறையானது பலதரப்பட்ட உருவங்கள் உள்ள பரப்பினை ஆய்வு செய்ய தகுதியானது அல்ல.ஆகவே இத்தகைய பரப்புகளை பகுத்தாய டெஃஸ்சர் க்ரேடிஎன்ட் முறை பயன்படுத்தப்படுகிறது. இந்த ஆய்வில், பலதரப்பட்ட தகவல்களும் க்ரேடிஎன்டும் கடினமான அலைத் தொகுப்பில் இருந்து பெறப்படுகின்றது.நொவெல் மார்கர் விதிமுறையில் பலதரப்பட்ட மட்ற்றும் ஒத்த முறையிலான பரப்புகளில் உருவங்களின் இடம் நிர்ணயிக்க உதவுகின்றது. இந்த மார்கர் ட்ரிவென் வாடெர்ஷெட் முறையானது இந்த பரப்புகளை பகுத்தாய உதவுகின்றது.பொருள் சார்ந்த தகவல் திரும்பபெருவதற்க்கு இந்த முறையானது மிகவும் பயனுள்ளதாகவும் பலதரப்பட்ட மற்றும் செம்மையான பகுத்தாய்தலுக்கும் உதவுகின்றது.

ACKNOWLEDGEMENT

I express my profound gratitude to our Chairman **Padmabhusan Arutselver Dr. N.Mahalingam B.Sc, F.I.E.,** for giving this great opportunity to pursue this course.

I thank, Principal, **Dr.Joseph V.Thanikal Ph.D., PDF., CEPIT** and vice principal Prof.R.Annamalai , Kumaraguru College of Technology, Coimbatore, for giving me this great opportunity to pursue this project.

I express my deep sense of gratitude to **Dr.S.Thangasamy, Ph.D.,** Professor and Dean of Department of Computer Science and Engineering, for his support and encouragement throughout the project.

I express my sincere thanks to Project Guide, **Mr.k.Sivan Arul Selvan, M.E.,** Computer Science and Engineering for his kind support, motivation and inspiration that triggered me for the project work.

I express my sincere thanks to Project Guide **Roshini.V.S (M.TECH), CDAC, Trivandrum** for her kind support, motivation and inspiration that triggered me for the project work.

I convey my special thanks to **Ms.V.Vanitha, M.E.,** Assistant Professor and Project Coordinator, Department of Computer Science and Engineering who support us to complete the project successfully.

I would like to convey my honest thanks to all Teaching staff members and Non Teaching staff members of the department for their support. I would like to thank all my classmates who gave me proper light moments and study breaks apart from extending some technical support whenever I needed them most.

I dedicate this Project work to my **Parents** for no reason but feeling from bottom of my heart, without their love this work wouldn't possible

TABLE OF CONTENTS

List of Figures:

List of Tables:

Abbreviations :

CWT	–	Continuous Wavelet Transform
DTCWT	_	Dual Tree Complex Wavelet Transform
DWT	–	Discrete Wavelet Transform
FB	–	Filter Banks
NDWT	–	Non Decimated Wavelet Transform
NDXWT	-	Non Decimated Complex Wavelet Transform
NDXWPT	-	Non Decimated Complex Wavelet Packet Transform
PR	-	Perfect Reconstruction property
QMF	-	Quadrature Mirror Filters
SIDWT	-	Shift Invariant Discrete Wavelet Transform
STFT	–	Short Term Fourier Transform
XWT	–	Complex Wavelet Transform

CHAPTER 1

INTRODUCTION

1.1 IMAGE PROCESSING

Image processing is any form of signal processing for which the input an image such as photographs or frames of video; the output of image processing can be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a 2D signal and applying standard signal-processing techniques to it.

As a subfield of digital signal processing, digital image processing has many advantages over analog image processing; it allows a much wider range of algorithms to be applied to the input data, and can avoid problems such as the build-up of noise and signal distortion during processing.

The elements of an image analysis system are shown as below:

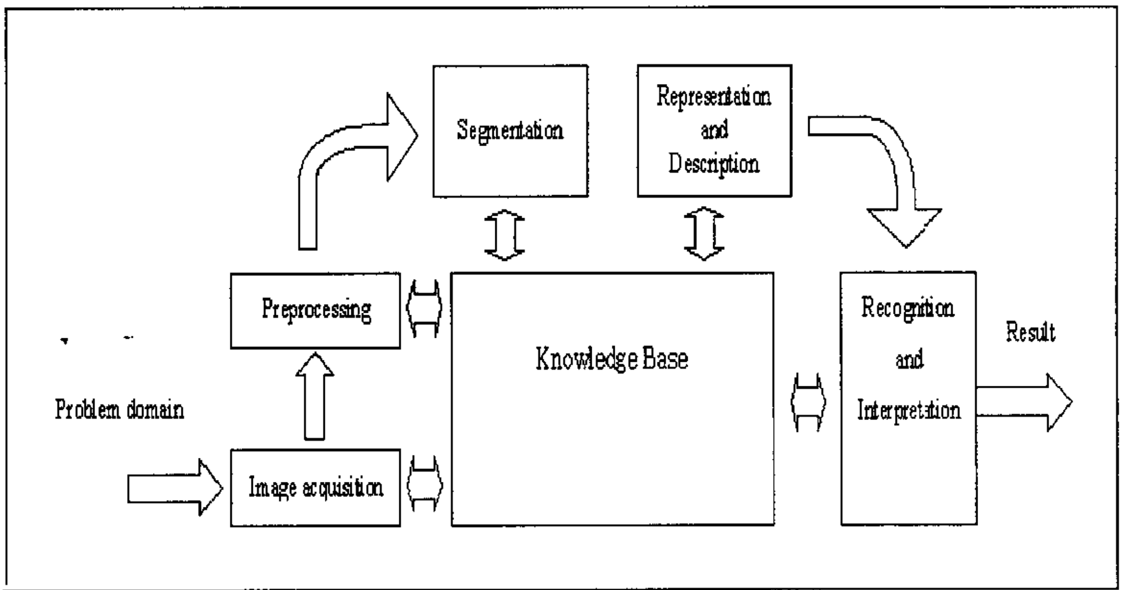

Figure 1.1: Elements of Image Analysis

Image analysis usually starts with a pre-processing stage, which includes operations such as noise reduction. For the actual recognition stage, *segmentation* should be done before it to extract out only the part that has useful information. Image segmentation is a primary and critical component of image analysis. The quality of the final results of an image analysis could depend on the segmentation step. On the other hand, segmentation is one of the most difficult tasks in image processing, especially automatic image segmentation.

Image processing and analysis is an important area in the field of robotics. The operation for autonomous vehicle is based on acquiring data that describe its environment. The motion planning and control of the vehicle requires an intelligent controller to make decisions to allow it to maneuver in an unknown field based on these data.

1.1.1 Applications of image processing

The various applications are:

- Computer Vision
- Feature detection
- Medical Image processing
- Morphological image processing
- Face Detection
- Non - Photorealistic rendering
- Remote Sensing

1.2 IMAGE SEGMENTATION

The aim of computer based image segmentation is to isolate/distinguish/subdivide a digital image into its constituent parts or objects. Image segmentation is often used as an initial transformation for general image analysis and understanding. Examples of image segmentation applications include remote sensing, medical image analysis and diagnosis, computer vision and the image segmentation necessary to enable region indexing in a Content Based Retrieval (CBR) application.

The ultimate aim of an automatic image segmentation system is to mimic the human visual system in order to provide a meaningful image subdivision. This is often unfeasible, not only because of the technical difficulties involved but also because scenes and objects themselves can be hierarchically segmented with no indication of what level of hierarchy is meaningful to an intended user.

Another difficulty of mimicking the human approach to segmentation is that a human can draw on an unconstrained and simply enormous set of object models with which to match scene elements to. The more usual methods of image segmentation rely on more primitive visual aspects of images such edges, color, texture.

The goal of the segmentation process is to define areas within the image that have some properties that make them homogeneous. The definition of those properties should satisfy the general condition that the union of neighboring regions should not be homogeneous if we consider the same set of properties. After segmentation, we can usually establish that the discontinuities in the image correspond to boundaries between regions.

Levels of segmentation:

- ***Complete segmentation:*** result in objects or regions of interest high level knowledge involved
- ***Partial segmentation***: Image is divided into separate regions that are homogeneous with respect to a chosen property such as brightness, color, reflectivity, texture etc.

The methods most commonly used for image segmentation can be categorized into 4 classes

1. ***Edge-based approaches:*** Image edges are detected and then linked into contours that represent the boundaries of image objects. The main advantage of edge-based approaches is their lower computational cost. However, the edge grouping process presents serious difficulties in setting appropriate thresholds and producing connected, one-pixel-wide contours.

2. ***Clustering-based approaches:*** Image pixels are sorted in increasing order as a histogram according to their intensity values. Fuzzy-c-means (FCM) and K-means fall into this method. The main advantage of this approach is that the problem of setting thresholds can be avoided by using iterative processes. Also, the segmented contours are always continuous. But, over segmentation may occur because pixels in the same cluster may not be adjacent.

3. ***Split/merge approaches:*** An input image is first segmented into homogeneous primitive regions using K-means or FCM as a 'Split' step. Then, similar neighboring regions are merged according to a certain decision rule as a 'Merge' step.

4. ***Region-based approaches:*** The goal is the detection of regions that satisfy a certain predefined homogeneity threshold. Region-based approaches are available because the segmented contours are always continuous and one-pixel-wide. The computation time of this approach is short. However, different similarity threshold settings may lead to different segmentation results. Also it can cause over segmentation.

5. ***Watersheds***: The watersheds transformation is studied in this thesis as a particular method of a region-based approach to the segmentation of an image. of using the image directly, the transform uses a gradient image extracted from the original image. The initial stage of any watershed segmentation method is therefore to produce a gradient image from the actual image. The complete transformation incorporates a pre-processing and post-processing stage that deals with embedded problems such as edge ambiguity and the output of a large number of regions.

CHAPTER 2

LITERATURE SURVEY

V. Grau*, A. U. J. Mewes, M. Alcañiz, ***2002*****—The watershed transform has interesting properties that make it useful for many different image segmentation applications: it is simple and intuitive, can be parallelized, and always produces a complete division of the image. However, when applied to medical image analysis, it has important drawbacks (oversegmentation, sensitivity to noise, poor detection of thin or low signal to noise ratio structures). We present an improvement to the watershed transform that enables the introduction of prior information in its calculation. We propose to introduce this information via the use of a previous probability calculation. Furthermore, we introduce a method to combine the watershed transform and atlas registration, through the use of markers.We have applied our new algorithm to two challenging applications: knee cartilage and gray matter/white matter segmentation in MR images. Numerical validation of the results is provided, demonstrating the strength of the algorithm for medical image segmentation.**

C.R. JUNG 1, J. SCHARCANSKI 2003 presents-the watershed transform has been used for image segmentation relying mostly on image gradients. However, background noise tends to produce spurious gradients, that cause over-segmentation and degrade the output of the watershed transform. Also, low-contrast edges produce gradients with small magnitudes, which may cause different regions to be erroneously merged. In this paper, a new technique is presented to improve the robustness of watersheds segmentation, by reducing the undesirable over-segmentation. A redundant wavelet transform is used to denoise the image and enhance the edges in multiple resolutions, and the image gradient is estimated with the wavelet transform. The watershed transform is then applied to the obtained gradient image, and segmented regions that do not satisfy specific criteria are removed.

Andrea Gavlasov´a, Ale˘s Proch´azka, and Martina Mudrov´a

Image segmentation, feature extraction and image components classification form a fundamental problem in many applications of multi-dimensional signal processing. The paper is devoted to the use of Wavelet transform for feature extraction associated with image pixels and their classification in comparison with the watershed transform. A specific attention is paid to the use of Haar transform as a tool for image compression and image pixels feature extraction.

Proposed algorithm is verified for simulated images and applied for a selected MR biomedical image processing in the MATLAB environment.

Nasser Chaji and Hassan Ghassemian 2006 presents a new biologically motivated method is proposed to effectively detect perceptually homogenous region boundaries. This method integrates the measure of spatial variations in texture with the intensity gradients. In the first stage, texture representation is calculated using the nondecimated complex wavelet transform. In the second stage, gradient images are computed for each of the texture features, as well as for grey scale intensity. These gradients are efficiently estimated using a new proposed algorithm based on a hypothesis model of the human visual system. After that, combining these gradient images, a *region gradient* which highlights the region boundaries is obtained. Non maximum suppression and then thresholding with hysteresis is used to detect contour map from the region gradients. Natural and textured images with associated ground truth contour maps are used to evaluate the operation of the proposed method. Experimental results demonstrate that the proposed contour detection method presents more effective performance than conventional approaches.

CHAPTER 3

METHODOLOGY

The watershed transform is a well established tool for the segmentation of images. Vital information characterizing texture can be lost in smoothing operation. In order to improve the generalization of watershed techniques and apply them properly to images containing significant amounts of texture content, the texture content information should be preserved within the algorithm. Texture boundaries have been used for the effective partitioning of natural images using the edge flow technique [18].

However, this technique does not use a measure of texture gradient but compares the texture content at each pixel to its neighbors in order to "flow" its texture content in the maximum gradient direction. Where "texture flows" meet, boundaries are constructed. However, over-segmentation, a major problem with the watershed transform, will not be solved by the use of the texture gradient. We develop a novel marker based solution (basins are flooded from selected sources rather than minima). This method lends itself well to the intended application of image region characterization for content based retrieval.

The summary of the steps involved are depicted in the figure given below:

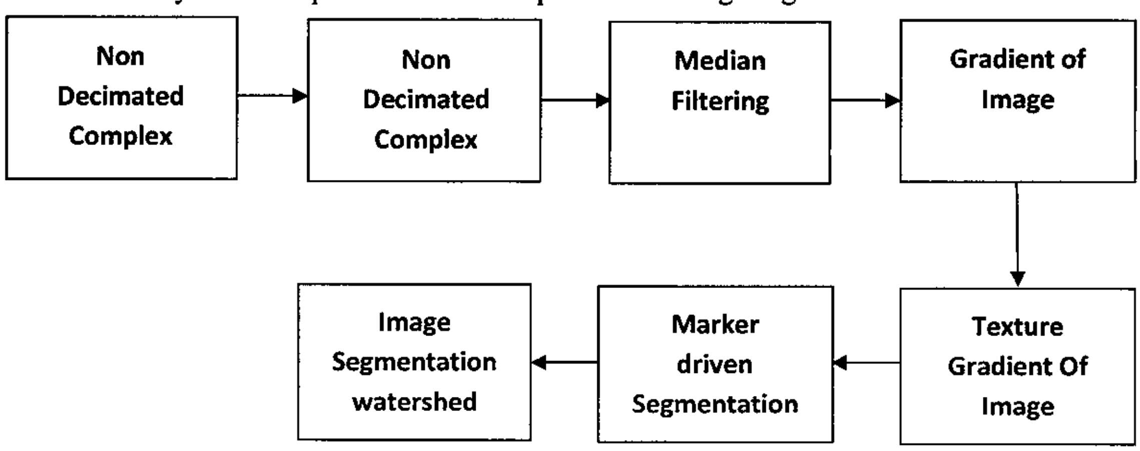

Figure 3.1: Flowchart of image Segmentation

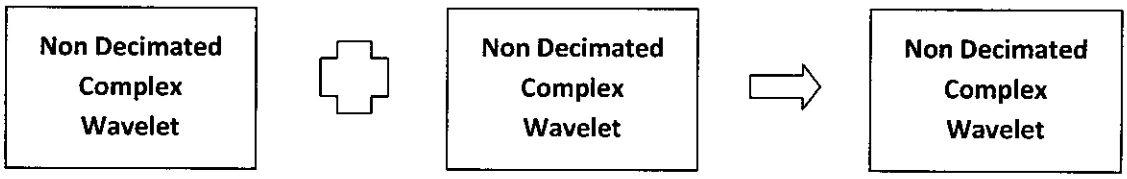

Figure 3.2 Flowchart depicting construction of Non Decimated Complex Wavelet Transform

3.1 WAVELET TRANSFORMS

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale.

Mathematical transformations are applied to signals to obtain a further information from that signal that is not readily available in the raw signal They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. Wavelets were developed independently in the fields of mathematics, quantum physics, electrical engineering, and seismic geology. Interchanges between these fields during the last ten years have led to many new wavelet applications such as image compression, turbulence, human vision, radar, and earthquake prediction.

Different transforms available are:

- Fourier transform
- Short time Fourier transform
- Wavelet transform

Why do we need the frequency information?

Often times, the information that cannot be readily seen in the time-domain can be seen in the frequency domain.

There are two main types of wavelet transform-continuous and discrete[2].The discrete transform is very efficient from the computational point of view but its only drawback is that it is not translation invariant. Translations of the original signal lead to different wavelet coefficients. To overcome this and to get more complete characteristics of the analyzed signal the modified version of the traditional wavelet transform DWT known as the Non-Decimated Wavelet Transform(NDWT)or stationary wavelet transform which has no subsampling step and therefore keeps the same number of coefficients at each level was proposed

3.1.1 CONTINUOUS WAVELET TRANSFORM

The continuous wavelet transform was developed as an alternative approach to the short time Fourier transform to overcome the resolution problem. The wavelet analysis is done in a similar way to the STFT analysis, in the sense that the signal is multiplied with a function, {\it the wavelet}, similar to the window function in the STFT, and the transform is computed separately for different segments of the time-domain signal. However, there are two main differences between the STFT and the CWT:

1. The Fourier transforms of the windowed signals are not taken, and therefore single peak will be seen corresponding to a sinusoid, i.e., negative frequencies are not computed.

2. The width of the window is changed as the transform is computed for every single spectral component, which is probably the most significant characteristic of the wavelet transform. The continuous wavelet transform is defined as follows:

$$CWT_x^{\psi}(\tau, s) = \Psi_x^{\psi}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t)\psi^* \left(\frac{t-\tau}{s}\right) dt$$

Equation 3.1 Continuous Wavelet transforms

As seen in the above equation, the transformed signal is a function of two variables, **tau** and s, the **translation** and **scale** parameters, respectively. **psi(t)** is the transforming function, and it is called **the mother wavelet** .

The term **mother wavelet** gets its name due to two important properties of the wavelet analysis as explained as follows:

The term **wavelet** means a **small wave**. The smallness refers to the condition that this (window) function is of finite length(**compactly supported**). The wave refers to the condition that this function is oscillatory. The term **mother** implies that the functions with different region of support that are used in the transformation process are derived from one main function, or the mother wavelet. In other words, the mother wavelet is a **prototype** for generating the other window functions.

The translation

The term **translation** is used in the same sense as it was used in the STFT; it is related to the location of the window, as the window is shifted through the signal. This term, obviously, corresponds to time information in the transform domain. However, we do not have a frequency parameter, as we had before for the STFT. Instead, we have scale parameter which is defined as $1/frequency$. The term frequency is reserved for the STFT. Scale is described in more detail in the next section.

The Scale

The parameter **scale** in the wavelet analysis is similar to the scale used in maps. As in the case of maps, high scales correspond to a non-detailed global view (of the signal), and low scales correspond to a detailed view. Similarly, in terms of frequency, low frequencies (high scales) correspond to a global information of a signal (that usually spans the entire signal), whereas high frequencies (low scales) correspond to a detailed information of a hidden pattern in the signal (that usually lasts a relatively short time).

MAHATMA GANDHI CENTRAL LIBRARY KCT CBE-6

Scaling, as a mathematical operation, either dilates or compresses a signal. Larger scales correspond to dilated (or stretched out) signals and small scales correspond to compressed signals.

In the definition of the wavelet transform, the scaling term is used in the denominator, scales **s > 1** dilates the signals whereas scales **s < 1** , compresses the signal

3.1.2 WAVELET THEORY

This section describes the main idea of wavelet analysis theory, which can also be considered to be the underlying concept of most of the signal analysis techniques. The FT defined by Fourier use **basis functions** to analyze and reconstruct a function. **Every vector in a vector space can be written as a linear combination of the basis vectors in that vector space** , i.e., by multiplying the vectors by some constant numbers, and then by taking the summation of the products. The analysis of the signal involves the estimation of these constant numbers (transform coefficients, or Fourier coefficients, wavelet coefficients, etc). The synthesis, or the reconstruction, corresponds to computing the linear combination equation.

A **basis** of a vector space **V** is a set of linearly independent vectors, such that any vector **v** in **V** can be written as a linear combination of these basis vectors. There may be more than one basis for a vector space. However, all of them have the same number of vectors, and this number is known as the **dimension** of the vector space.

For example in two-dimensional space, the basis will have two vectors.

$$\mathbf{v} = \sum_{k} \nu^{k} \mathbf{b}_{k}$$

Equation 3.2

Equation 3.2 shows how any vector **v** can be written as a linear combination of the basis vectors **b_k** and the corresponding coefficients **nu^k** .

This concept, given in terms of vectors, can easily be generalized to functions, by replacing the basis vectors **b_k** with basis functions phi_k(t), and the vector **v** with a function f(t). Equation 3.2 then becomes

$$f(t) = \sum_{k} \mu_k \phi_k(t)$$

Equation 3.3

The complex exponential (sines and cosines) functions are the basis functions for the FT. Furthermore, they are orthogonal functions, which provide some desirable properties for reconstruction.

Although the discretized continuous wavelet transform enables the computation of the continuous wavelet transform by computers, it is not a true discrete transform. As a matter of fact, the wavelet series is simply a sampled version of the CWT, and the information it provides is highly redundant as far as the reconstruction of the signal is concerned. This redundancy, on the other hand, requires a significant amount of computation time and resources. The discrete wavelet transform (DWT), on the other hand, provides sufficient information both for analysis and synthesis of the original signal, with a significant reduction in the computation time.

The DWT is considerably easier to implement when compared to the CWT.

3.2 TEXTURE CHARACTERIZATION

In order to produce a texture gradient we first need to characterize the texture content of the image at each pixel. A number of methods have been proposed to do this. One of the most popular techniques is the use of a set of scaled and orientated complex Gabor filters (e.g., [19]). By suitable spanning of the frequency space, each pixel can be characterized in texture content.
However, when considering the differences in texture within an image (e.g., the texture gradient) this often produces suboptimal characterization for segmentation. To produce

an optimal system, the Gabor filters need to be tuned to the texture content of the image. Different schemes for adaptive Gabor filtering have been implemented [20], [21]. These and other schemes use arbitrary techniques that are entirely separate from the texture feature extraction process whilst also being excessively computationally complex.

In order to integrate an adaptive scheme with the texture feature extraction process the Non-Decimated Complex Wavelet Packet Transform (NDXWPT)is developed. The magnitude of the coefficients of each complex subband can be used to characterize the texture content. This is because the basis functions from each subband (very closely) resemble Gabor filters, i.e., they are scale and directionally selective whilst being frequency and spatially localized. Each pixel can therefore be assigned a feature vector according to the magnitudes of the NDXWPT coefficients. A pixel at spatial position has one feature for each NDXWPT subband coefficient magnitude at that position: defined as , where is the subband number. A feature vector is therefore associated with each pixel characterizing the texture content at that position.

3.3 NON-DECIMATED COMPLEX WAVELET PACKET TRANSFORM

In order to develop the nondecimated complex wavelet packet transform (NDXWPT) transform, it is first instructive to develop a nonadaptive version using the NDXWT transform.

3.3.1 NON-DECIMATED COMPLEX WAVELET TRANSFORM

The structure of the nondecimated complex wavelet packet transform is based upon the Non-Decimated Wavelet Transform(NDWT) (defined in [15]) combined with the Complex Wavelet Transform (XWT) (defined in [16]).

.3.2.1 NON-DECIMATED WAVELET TRANSFORM(NDWT)

The nondecimated wavelet transform uses the same structure as the DWT but without any subsampling. This leads to a transform that has every subband having exactly the same number of coefficients as the number of

samples in the original signal. This of course leads to an over complete representation. However, this transform has the advantage of being completely shift invariant, an aspect which has been exploited by applications such as image fusion , where it was described as the Shift Invariant Discrete Wavelet Transform (SIDWT). The structure of the NDWT is shown in figure below. This clearly shows that in order to retain the same effective filtering without the downsampling, the filters at each stage must be upsampled with the right number of zeros. The filters in below figure are best represented in the –transform domain with the filters and representing the low and high pass filters respectively. The filters used at level are $H_0(z^{2i-f})$ and $H_1(z^{2i-f})$. The terms represent the
original filters upsampled with zeros in-between each original filter tap.
More precisely, it applies the transform at each point of the image and saves the detail coefficients and uses the low-frequency coefficients for the next level. By using all coefficients.

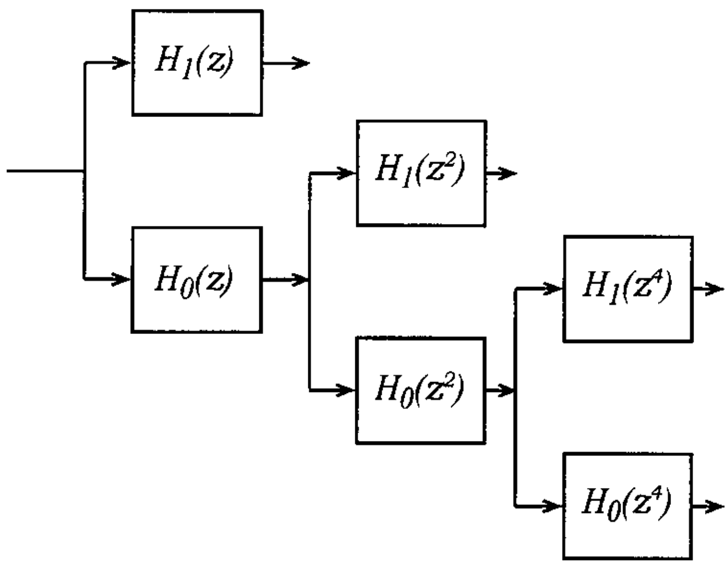

Figure 3.3 Non-decimated wavelet Transform.

at each level, we get very well allocated high-frequency information. With this transform the number of pixels involved in computing a given coefficient grows slower and so the relation between the frequency and spatial information is more precise. From the computational point of view the NDWT has larger storage space requirements and involves more computations. Two main algorithms a trous [6] and Beylkin's algorithm [1] exist for computing the transform .For reviews on the topic see [5], [12], and [13]

3.3.2.2 COMPLEX WAVELET TRANSFORM (XWT)

Complex wavelet transform a form of discrete wavelet transform, which generates *complex* coefficients by using a *dual tree* of wavelet filters to obtain their real and imaginary parts for motion estimation was developed. It provides both shift invariance and good directional selectivity with only modest increases in signal redundancy and computational load. However development of a CWT with perfect reconstruction and good filter characteristics has proven difficult until recently.

Nick Kingsbury[2]has developed a dual-tree algorithm with a real biorthogonal transform, and an approximate shift invariance can be obtained by doubling the sampling rate at each scale, which is achieved by computing two parallel subsampled wavelet trees respectively and it generates the real and imaginary parts of the wavelet coefficients separately. It introduces limited redundancy ($2m$:1 for m-dimensional signals) and allows the transform to provide approximate shift invariance and directionally selective filters (properties lacking in the traditional wavelet transform) while preserving the usual properties of perfect reconstruction and computational efficiency with good well-balanced frequency responses.

In [7,8], the Dual-Tree Complex Wavelet Transform (DT CWT) is introduced. The following properties are exhibited:

- Approximate shift invariance;
- Good directional selectivity in 2-dimensions (2-D) with Gabor-like filters (also true for higher dimensionality, m-D);
- Perfect reconstruction (PR) using short linear-phase filters;
- Limited redundancy, independent of the no. of scales,2 : 1 for 1-D ($2m$: 1 for m-D);
- Efficient order— only N computation twice the simple DWT for 1-D ($2m$ times for m-D).

The dual-tree complex DWT of a signal $x(n)$ is implemented using two critically-sampled DWTs in parallel on the same data. The DTCWT has 6 filters, all of which are directionally selective (−15°, −45°, −75°, +15°, +45°+75°).

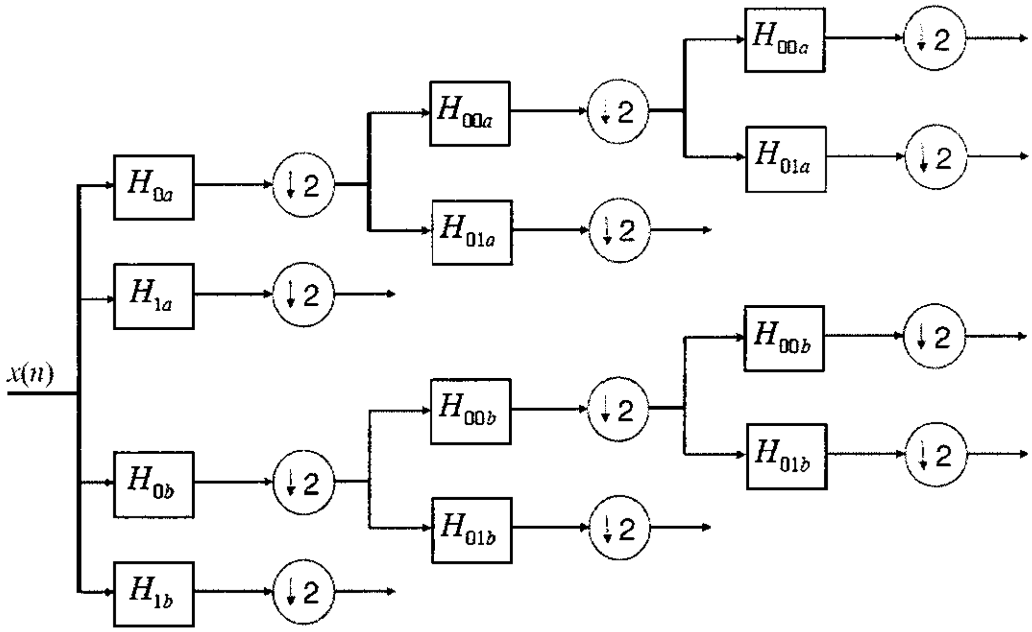

Figure 3.4 Complex Dual Tree

Any finite energy analog signal $x(t)$ can be decomposed in terms of wavelets and scaling functions via

$$x(t) = \sum_{n=-\infty}^{\infty} c(n)\,\phi(t-n) + \sum_{j=0}^{\infty}\sum_{n=-\infty}^{\infty} d(j,n)\,2^{j/2}\,\psi(2^j t - n).$$

The scaling coefficients $c(n)$ and wavelet coefficients $d(j, n)$ are computed via the inner products

$$c(n) = \int_{-\infty}^{\infty} x(t)\,\phi(t-n)\,dt,$$

$$d(j, n) = 2^{j/2} \int_{-\infty}^{\infty} x(t)\,\psi(2^j t - n)\,dt.$$

There exists a very efficient, linear time complexity algorithm to compute the coefficients *c(n)* and *d(j, n)* from a fine-scale representation of the signal (often simply *N* samples) and vice versa based on two octave-band, discrete-time FBs that recursively
Apply a discrete-time low-pass filter *h*0*(n)*, a high-pass filter *h*1*(n)*, and upsampling and downsampling operations. These filters provide a convenient parameterization for designing wavelets and scaling functions with desirable properties, such as compact time support and fast frequency decay (to ensure the analysis is as local as possible in time frequency)and orthogonality to low-order polynomials (vanishing moments)[12].

The key challenge in dual-tree wavelet design is thus the joint design of its two FBs to yield a complex wavelet and scaling function that are as close as possible to analytic. The question of how to design filters *h*0*(n)* and *h*1*(n)* satisfying the perfect reconstruction (PR) conditions so that the wavelet $\psi(t)$ has short support and vanishing moments was answered by Daubechies [12].The two real wavelet transforms use two different sets of filters, with each satisfying the PR conditions. The two sets of filters are jointly designed so that the overall transform is approximately analytic. Let *h*0*(n)*, *h*1*(n)* denote the low-pass/high-pass filter pair for the upper FB, and let *g*0*(n)*, *g*1*(n)* denote the low-pass/high-pass filter pair for the lower FB.

1) FILTER DESIGN FOR THE DUAL-TREE CWT

In particular, the analysis-reconstruction subsystem can introduce three separate types of distortions: aliasing, short-time phase distortion, and short-time frequency distortion [9 10]. The use of quadrature mirror filters (QMF's) [6] allows the aliasing to be removed in the reconstruction stage, and consequently, QMF's have become a fundamental building block in most tree-structured subband coder systems. This leads to the definition of a new class of exactly reconstructing analysis/ reconstruction filters called "conjugate quadrature filters" or CQF's.

In this system, the analysis is performed by the two frequency selective filters, $H_o(e^{j\omega})$ and $H_1(e^{j\omega})$, which are nominally a half-band low-pass and half-band high-pass filter, respectively. In the reconstruction, the bands are filtered by low pass and high pass synthesis filters $(G_0(e^{j\omega}), G_1(e^{j\omega}))$

$$\hat{X}(e^{j\omega}) = (\tfrac{1}{2})[H_0(e^{j\omega})\, G_0(e^{j\omega}) + H_1(e^{j\omega})\, G_1(e^{j\omega})]\, X(e^{j\omega})$$

$$+ (\tfrac{1}{2})[H_0(-e^{j\omega})\, G_0(e^{j\omega}) + H_1(-e^{j\omega})\, G_1(e^{j\omega})]\, X(-e^{j\omega}).$$

----------(1)

component is contained in the first term of (1), while the second term contains the aliasing. In the classic QMFs solution [11], the aliasing is removed by defining the reconstruction filters as

$$G_0(e^{j\omega}) = H_1(-e^{j\omega})$$

$$G_1(e^{j\omega}) = -H_0(-e^{j\omega})$$

----------(2)

In addition, QMF's are defined to be frequency shiftedversions of one another, i.e.,

$$H_1(e^{j\omega}) = H_0(-e^{j\omega})$$

-----------(3)

and are also constrained to have even length. For this class of analysis/reconstruction system, exact reconstruction requires that

$$H_0^2(e^{j\omega}) - H_1^2(e^{j\omega}) = 2.$$

----------(4)

The filters used are hence designed in such a manner that the wavelet used in one branch is approximately the Hilbert Transform of the wavelet in the other branch. The coefficients in one branch can then be considered the real part of a complex wavelet transform with the coefficients in the other branch considered as the imaginary part. The dual-tree CWT used here is of length-8 filters, the table of coefficients of the analysing filters in the first stage. The reconstruction filters are obtained by simply reversing the alternate coefficients of the analysis filters.

h(0)	-0.01059740	g(0)	-0.23037781
h(1)	0.03288301	g(1)	0.71484657
h(2)	0.03084138	g(2)	-0.63088076
h(3)	-0.18703481	g(3)	-0.02798376
h(4)	-0.02798376	g(4)	0.18703481
h(5)	0.63088076	g(5)	0.03084138
h(6)	0.71484657	g(6)	-0.03288301
h(7)	0.23037781	g(7)	-0.01059740

Table 1: Coefficients of Daubechies wavelet transform filter used in the experiment.

2) THE 1-D DUAL-TREE CWT

Because the 1-D wavelet and scaling filters neglect the negative half spectrum, these regions cover only the first quadrant of the frequency cell. However, real images contain non-redundant information in both first and second quadrants; therefore, we need to include a parallel processing path whose equivalent filters Each level of the complete tree produces six complex valued bandpass subimages as well as two lowpass subimages and on which the subsequent filtering stage operates.

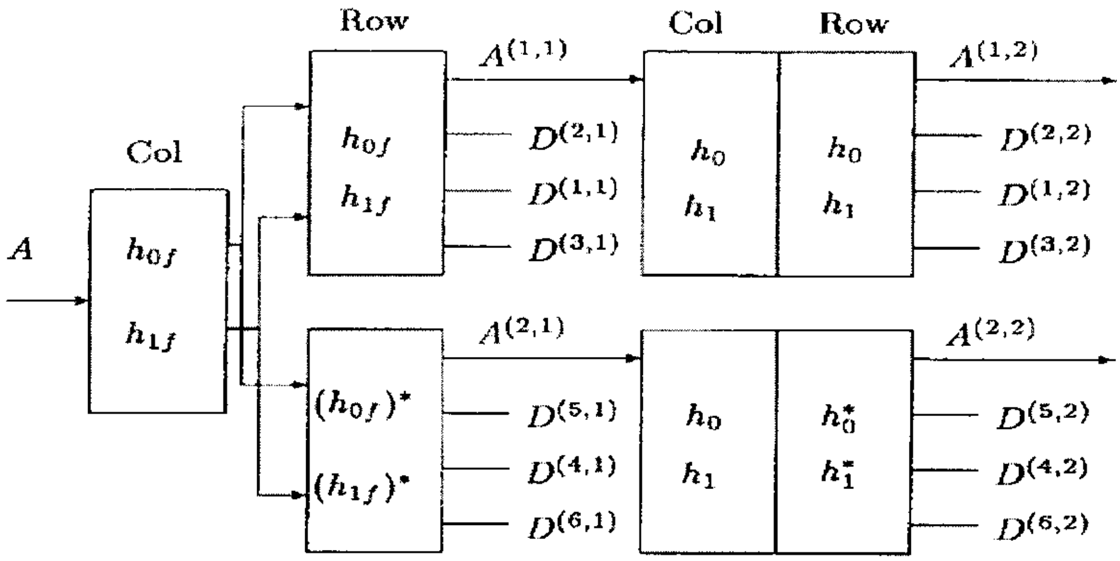

Figure 3.5 2D-Complex Wavelet Transform

3) EXTENSION TO MULTI-DIMENSIONS

To extend the transform to higher-dimensional signals, a filter bank is usually applied separably in all dimensions. The 2-D DWT produces three bandpass subimages at each level, which are corresponding to LH, HH, HL, and oriented at angles of 0° , ±45°, 90°.

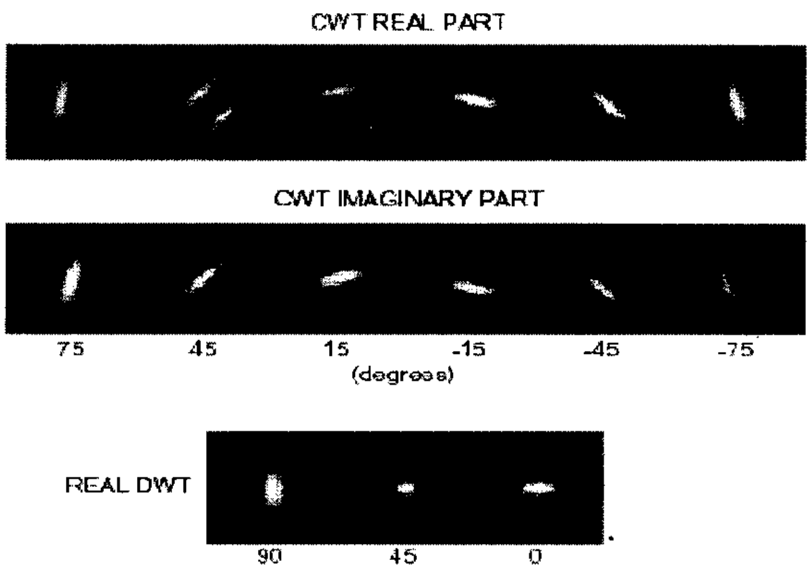

Figure 3.6. 2-D impulse responses of the complex wavelets at level 4 (6 bands at angles from -75° to +75°) and equivalent responses for a real wavelet transform (3 bands)

The 2-D CWT can provide six subimages in two adjacent spectral quadrants at each level, which are oriented at angles of $\pm 15^{\circ}$, $\pm 45^{\circ}$, $\pm 75^{\circ}$. This is shown in fig 3.6

The strong orientation occurs because the complex filters are asymmetry responses. They can separate positive frequencies from negative ones vertically and horizontally. So positive and negative frequencies won't be aliasing. The orientations of details is shown in fig 3.7.

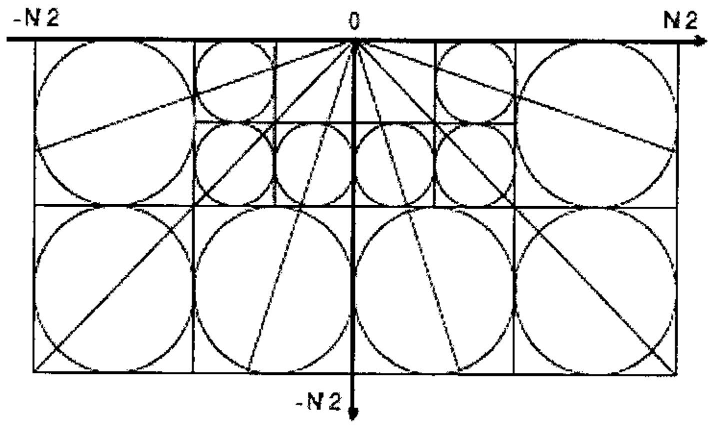

Figure 3.7 Directional selectivity of the frequency space corresponding to the complex wavelet transform

CONVERSION FROM THE XWT TO THE NDXWT

The implementation of the XWT is identical to the implementation of the DWT with complex valued filters. The conversion of the XWT to its non-decimated form (NDXWT) is therefore identical to the conversion of the DWT to its non-decimated form(NDWT), i.e., all subsampling is removed and all filters past the first level are upsampled.

3.3.2.3 WAVELET PACKET TRANSFORM

Wavelet packet transform (WPT) is a generalization of the dyadic wavelet transform (DWT) that offers a rich set of decomposition structures. WPT was first introduced by Coifman *et al.* [16] for dealing with the nonstationarities of the data better than DWT does.

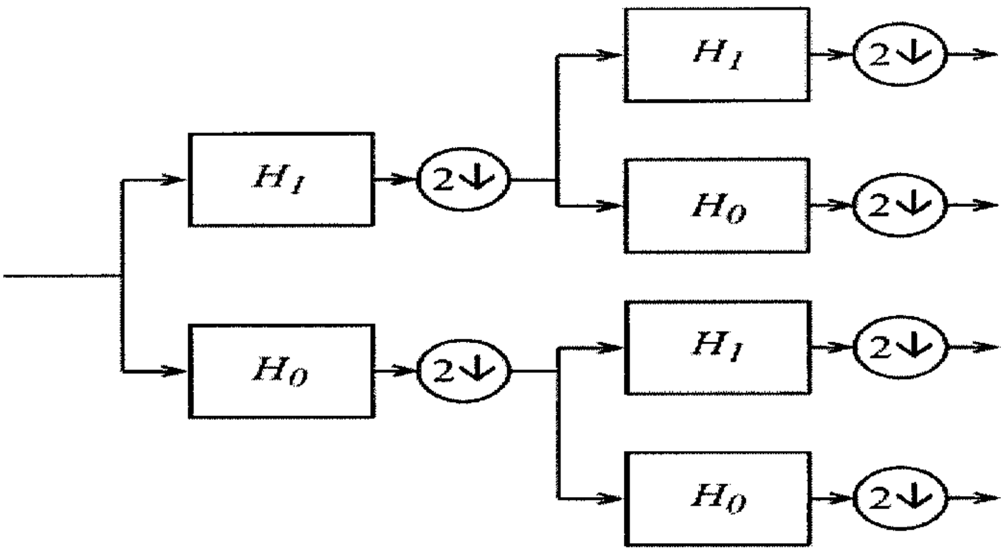

Fig. 3.8 Full wavelet packet binary tree for two levels

The wavelet decomposition can be done in two different ways:

- pyramid-structured wavelet transform,
- tree –structured wavelet transform

The wavelet decomposition of an image, using pyramid-structured wavelet transform, generates a set of subimages, which contains low-frequency components of the original image. This decomposition is suitable for images in which the majority of information is concentrated in low frequency region, i.e. for inspection images primarily with smooth components.

Many researchers have concluded [4] that the most significant information of texture often appears in the middle frequency bands. Hence, further decomposition in the lower frequency region by conventional wavelet transform may not help much Therefore an appropriate way to perform wavelet transform for textured image is to locate dominant frequency bands and then decompose them further. This leads to the concept of tree-structured wavelet transform or wavelet packets.

The hierarchical wavelet transform uses a family of wavelet functions and its associated scaling functions to decompose the original signal (image) into different subbands due to their finite duration which provides both the frequency and spatial locality The key difference between the traditional pyramid algorithm and the wavelet packet algorithm is that the recursive decomposition is no longer applied to the low frequency sub-bands. Instead, it is applied to any of the frequency bands based on some criterion or cost function, leading to quadtree structure decomposition.

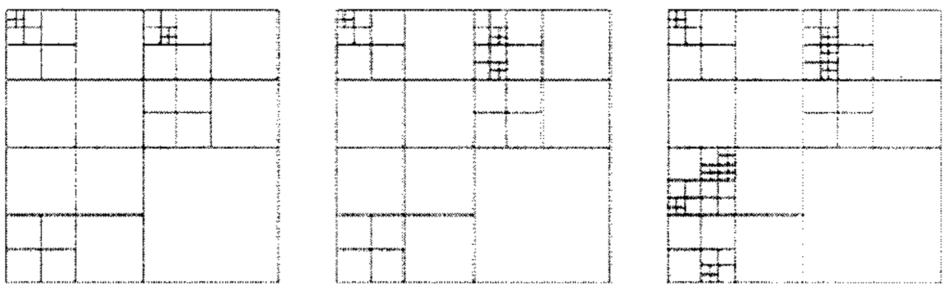

Fig 3.9 Decomposition structure at 0.5bpp of the boat image

If an orthonormal wavelet basis has been chosen, the coefficients computed are independent and possess a distinct feature of the original signal. Wavelet packets can be described by the following collection of basis functions[18]:

$$W_{2n}(2^{p-1}x-l)=\sqrt{2^{1-p}}\sum_{m}h(m-2l)\sqrt{2^{p}}W_{n}(2^{p}x-m)$$

$$W_{2n+1}(2^{p-1}x-l)=\sqrt{2^{1-p}}\sum_{m}g(m-2l)\sqrt{2^{p}}W_{n}(2^{p}x-m)$$

---------(5)

where p is a scale index, l is a translation index, h is a lower-pass filter , g is a high-pass filter with $g(k) = (-1)k\ h(1-k)$. The function $W_0(x)$ can be identified with the scaling function ϕ and $W_1(x)$ with the mother wavelet ψ .The search for the "best" non-redundant representation of the data by any subtree of the WPT is called best-basis selection [5] . Best-basis selection begins by evaluating each subband with a desired metric (e. g. rate or distortion, entropy, variance, energy).

First, we perform a full wavelet packet decomposition for each class texture and use the energy measure as signatures for the coefficients (See Algorithm1).

Algorithm 1: Tree-Structured Wavelet Transform

1) Decompose a given textured image with 2-D two-scale wavelet transform into 4 subimages, which can be viewed as the parent and children nodes in a tree as

2) Calculate the energy of each decomposed image (children node). That is, if the decomposed image is x*(m, n)*,with $1 <= m <= M$ and $1 <= n <= N$, the energy is

$$e = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} |x(m,n)|. \qquad \text{-------------(6)}$$

3) If the energy of a subimage is significantly smaller than others, we stop the decomposition in this region since it contains less information. This step can be achieved by comparing the energy with the largest energy value in the same scale.

4) If the energy of a subimage is significantly larger, we apply the above decomposition procedure to the subimage.

However, the full wavelet packet decomposition will produce many coefficients and so a large feature set will be obtained. As proposed by Chang and Kou [4], the most dominant frequency channels obtained from tree structured wavelet transform are very good for discriminating textures. Hence, we can reduce the feature set by means of choosing the signatures with highest energy value

Practically, the size of the smallest subimages should be used as a stopping criterion for further decomposition. If the decomposed channel has a very narrow size, the location and the energy value of the feature may vary widely from sample to sample so that the feature may not be robust. According to our experience, the size of the smallest subimages should not be less than 16 x 16. Consequently, if the input image size is 256 x 256 (or 64 x 64), a 4-level (or 2-level) tree-structured wavelet transform is appropriate.

WPT exploits interband and intraband dependencies of the transform coefficients. Experimental results show that wavelet packet transform brings consistent improvement over dyadic wavelet transform. Wavelet Packet Transform (WPT) provides good spectral and temporal resolutions in arbitrary regions of the time-frequency plane.

3.4 GRADIENT ESTIMATION

Since an edge is defined by an abrupt change in intensity value, an operator that is sensitive to this change can be considered as an edge detector. The rate of change of the intensity values in an image is large near an edge and small in constant areas. Therefore, a gradient operator may be used in order to highlight the edge pixels.
In two-dimensional images, it is important to consider level changes in many directions. For this reason, the directional sensitive gradient operators are used. The output of any directional sensitive gradient operator contains information about how strong the edge is at that pixel in the same direction of the operator sensitivity.

3.4.1 GRADIENT EXTRACTION FROM THE NDXWT TRANSFORM

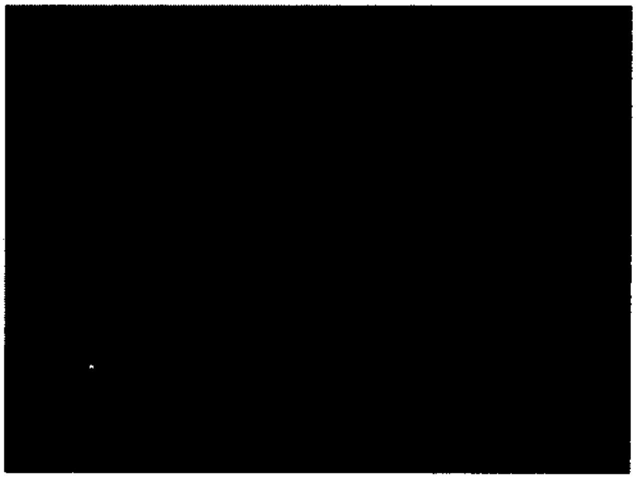

Figure 3.10 Complex Wavelet Subband

The above figure shows the magnitude of a single, orientated, second scale subband from a NDXWT complex wavelet decomposition of the Lena test image, This image shows how the texture content is highlighted by wavelet subbands (see the feather region on Lena's hat). It is therefore reasonable to characterize the texture content at each spatial position (x,y)using the feature vector t(X,Y) (where i is the NDXWT subband number and the vector Ti(x,y) values are defined as the NDXWT subband coefficient magnitude at that position). This is made possible because the nondecimated wavelet subbands are all the same size as the image.

A simple approach to obtaining the texture gradient of an image would then be to calculate the gradient of each subband magnitude and sum them. This would work for purely textured images. However, all texture extraction methods will give high energy values over simple intensity boundaries found in nontextured image regions (see the edge of the top of the hat in Fig. 7). The gradient of the subband magnitudes will give a double edge at such intensity boundaries. The gradient of each subband should therefore aim at step detection rather than edge detection. A simple method to perform this is a separable median filtering on the magnitude image followed by gradient extraction. This has the effect of removing the edges and preserving the steps. The texture content can then be represented by the median filtered versions of the subband magnitudes.

Since an edge is defined by an abrupt change in intensity value, an operator that is sensitive to this change can be considered as an edge detector. The rate of change of the intensity values in an image is large near an edge and small in constant areas. Therefore, a gradient operator may be used in order to highlight the edge pixels.

In two-dimensional images, it is important to consider level changes in many directions. For this reason, the directional sensitive gradient operators are used. The output of any directional sensitive gradient operator contains information about how strong the edge is at that pixel in the same direction of the operator sensitivity.

Here the horizontal gradient is calculated as

$$B(j,k)=A(j,k+1)-A(j,k-1) \quad \text{-----------(7)}$$

The vertical gradient is calculated as

$$B(j,k)=A(j+1,k+1)-A(j-1,k) \quad \text{----------(8)}$$

The gradient of the image is calculated as the square root of the sum of the squares of the horizontal and vertical gradients.

This can be represented by: $MT_i(x,y)=MedianFilter(T_i(x,y))$ for $1<i<n$ ------------(9)

3.4.2 MEDIAN FILTER

The median filter is a non-linear digital filtering technique, often used to remove noise from images or other signals. It is particularly useful to reduce speckle noise and salt and pepper noise. Its edge-preserving nature makes it useful in cases where edge blurring is undesirable

How It Works

The median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. The values in the window are sorted into numerical order; replace the center value in the window with the median of all the pixel values in the window. The kernel is usually square but can be any shape. . (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.) Figure 3.11 illustrates an example calculation.

An example of median filtering of a single 3x3 window of values is shown

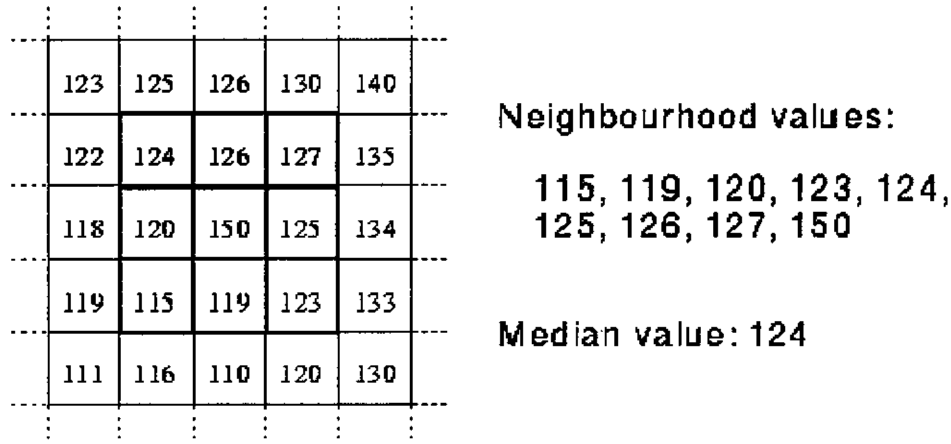

Figure 3.11Calculating the median value of a pixel neighborhood. A 3×3 square neighborhood is used here --- larger neighborhoods will produce more severe smoothing.

The median filter has two main advantages over the mean filter:

- The median is a more robust average than the mean and so a single very unrepresentative pixel in a neighborhood will not affect the median value significantly.
- Since the median value must actually be the value of one of the pixels in the neighborhood, the median filter does not create new unrealistic pixel values when the filter straddles an edge. For this reason the median filter is much better at preserving sharp edges than the mean filter
- In general, the median filter allows a great deal of high spatial frequency detail to pass while remaining very effective at removing noise on images where less than half of the pixels in a smoothing neighborhood have been effected. (As a

consequence of this, median filtering can be less effective at removing noise from images corrupted with Gaussian noise.)

Other advantages

- Gives more smoothing but more distortion
- It removes 'impulse' noise (outlying values, either high or low).
- in the presence of noise it does blur edges in images slightly.

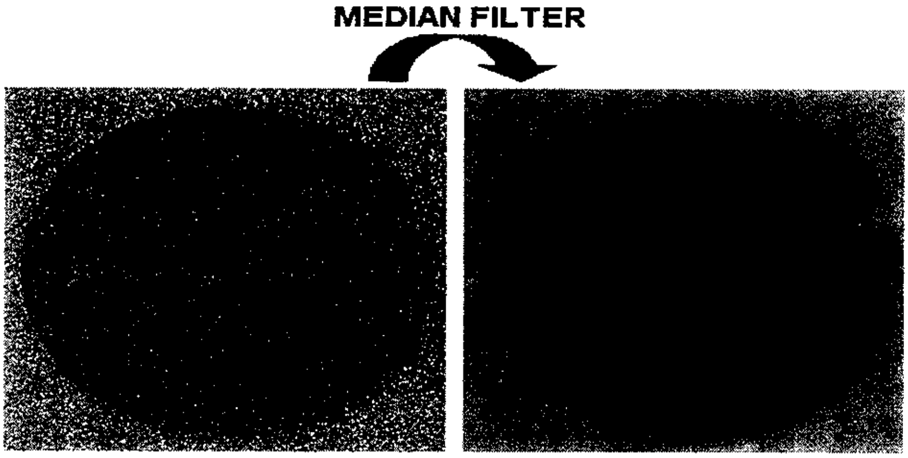

Figure 3.12 Median filtered image

The median filtering should reflect the size of the edges in the complex subbands produced by simple intensity boundaries in order to negate them, i.e.octave scale median filtering. In order to calculate the gradient of the texture content one needs to consider the gradient within the multidimensional feature space. The simplest way to do this is to sum the gradients obtained for each of the individual features. Defining to be the magnitude of the texture gradient we have:

$$TG(x,y) = \sum_{i=1}^{n} \frac{|\nabla(MT_i(x,y))|}{l_2(MT_i)}$$

----------(10)

where is the number of subbands and Δ is approximated using a Gaussian derivative gradient extraction technique [18] [with the scale parameter set to 2.0]. l_2MT_i is the l_2

norm energy of the median filtered subband and is included to normalize the effect of each subband on the gradient.

The TG gradient clearly highlights the edge of the texture regions in the artificial texture images together with the natural image (see edge of the feathers in the hat). Clearly this gradient is suited to the detection of texture boundaries. In order to preserve the ability of the system to detect intensity changes, this gradient is combined with a simple intensity gradient as follows:

$$G(x,y) = \frac{mix \times (|\nabla f|)^{1.2}}{(F(|MT(x,y)|))^3} + (TG(x,y))^3 \qquad \text{----------(11)}$$

where G(x,y) is the final combined gradient and *mix* is a suitably chosen constant for mixing the intensity and texture gradients. ∇ F is just the gradient of the plain intensity image calculated

using the Roberts Operator

F(|MT(x,y)|) is defined as

{ mean(|MT|) if <=mean(|MT|)
{
{ (|MT(x,y) if |MT(x,y)|>=mean(|MT|)

nos - the n umber of complex sub bands produced by the non decimated complex wavelet packet decomposition

The power 1.2 of $(|\nabla f|)^{1.2}$: This is included to increase the dynamic range of the gradient image emphasizing the larger gradient values.

The power of 3 of $(TG(x,y))^3$: This is included to emphasize the larger texture gradient values. This is necessary because the newly defined texture gradient has a smaller dynamic range.

The function F : *The* $F(|MT(x,y)|^3$ factor is included to reduce the effect of the spurious gradients within highly textured regions. The F function is included so the factor $F(|MT(x,y)|^3$ does not relatively elevate the importance of the regions where |MT(x,y)| is small.

3.5 MARKER DRIVEN SEGMENTATION

The problem of over-segmentation of the watershed method was dealt with through the flooding from selected sources (i.e., marker driven segmentation). The other methods were not chosen as they did not apply easily to texture gradients [11] or they tend to produce small residual regions (hierarchical watersheds [8]) and therefore were not suited to an application to region characterization. Most of the marker selection methods suggested by Beucher[5] are application dependant

The aim of marker identification within a content based retrieval application is to pinpoint regions that are homogeneous in terms of texture, color and intensity and of a significant size. To meet these criteria a minimum region, moving threshold and region growing method was adopted as shown in Algorithm II.

This algorithm calculates the mean and standard deviation of the gradient image. Then several thresholded binary images are produced at reasonably spaced thresholds using the mean and standard deviation of gradient image. For each binary thresholded image, the number of closed and connected regions greater than the given minimum size is calculated. The threshold with the maximum number of connected regions is used as the output marker image. This is a similar method to that developed by Deng and Manjunath [21] although this not applied to marker selection.

```
Algorithm II: MinsizeThreshold(minsz,G)
Comment   : minsz        - the minimum acceptable marker size
Comment   :  G           - input Gradient image
      std          =   STANDARDDEVIATION OF G
      median       =   MEDIAN OF G
      threshs[12]  =   { -0.9,-0.6,-0.55,-0.5,-0.4,-0.35, -0.3,-0.2,-0.1,0.0,0.1,0.2}
      for i =1 to 12
            {   thresholdLevel   =  median + threshs[i]  *  std
                thresholdImage   =  GTI(thresholdlLevel,G)
```

```
            markerImage[i]    = GCRLT(minsz)
            regionNumber[i]   = NOR(markerImage[i])
        }
    minIndex= FindMinValue(regionNumbers)
    return(markerImage(minIndex))
```

Comment	:	GTI	-	GetThresholdImage
Comment	:	GCRLT	-	GetConnected RegionsLessThan
Comment	:	NOR	-	Number of Regions

The smaller the value of *minsize may lead to oversegmentation.* However the larger values may lead to separate regions being merged. This is a common problem for segmentation algorithms. However these images show how the presented algorithm is able to properly segment image regions according to intensity and texture changes.

3.6 WATERSHED TRANSFORMATION

WATERSHEDS are one of the classics in the field of topography. Everybody has heard for example about the great divide, this particular line which separates the U.S.A. into two regions. A drop of water falling on one side of this line flows down until it reaches the Atlantic Ocean, whereas a drop falling on the other side flows down to the Pacific Ocean. As we shall see in further detail later, this great divide constitutes a typical example of a watershed line. The two regions it separates are called the catchment basins of the Atlantic and the Pacific Oceans, respectively. The two Oceans are the minima associated with these catchment basins.

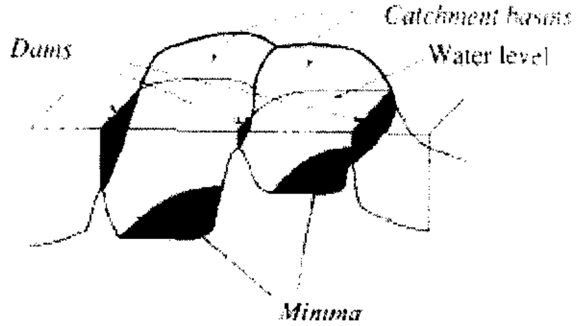

Fig.3.13. Building dams at the places where the water different minima would merge coming from two different minima would merge.

Let us express this immersion process more formally: I being the grayscale image under study, denote h_{min} the smallest value taken by I on its domain DI. Similarly, denote h_{max}, the largest value taken by I on DI. In the following, Th(I) stands for the threshold of I at level h:

$$T_h(I) = \{p \in D_I, I(p) \leq h\}. \quad \text{------------(12)}$$

We also denote C(M) the catchment basin associated with a minimum M and Ch(M) the subset of this catchment basin made of the points having an altitude smaller or equal to h:

$$C_h(M) = \{p \in C(M), I(p) \leq h\} = C(M) \cap T_h(I). \quad \text{-----------(13)}$$

By analogy, we can figure that we have pierced holes in each regional minimum of I, this picture being regarded as a (topographic) surface. We then slowly immerse our surface into a lake. Starting from the minima of lowest altitude, the water will progressively fill up the different catchment basins of I. Now, at each pixel where the water coming from two different minima would merge, we build a "dam" (see Fig. 2). At the end of this immersion procedure, each minimum is completely surrounded by dams, which delimit its associated catchment basin. The whole set of dams which has been built thus provides a tessellation of I in its different catchment basins. These dams correspond to the watersheds of I.

This informal analysis of the immersion process can be established mathematically by the definition of *'geodesic distance'* and *'geodesic influence zone'*. 'Geodesic influence zone' deals with the expansion of the plateau from each local minimum for watersheds transformation.

Definition: Geodesic Distance

The geodesic distance $d_A(x,y)$ between two pixels x and y in A is the infimum (greatest lower bound of given set) length of the paths which join x and y and are totally included in A.

Definition: Geodesic Influence Zone

Suppose A contains a set B consisting of several connected components B_1, B_2,... B_k. The geodesic influence $iz_A(B_i)$ of a connected component B_i of B in A is the locus of the points of A whose geodesic distance to B_i is smaller than their geodesic distance to any other component of B. It can be expressed as

$$iz_A(B_i)=\{ p\in A,\ \forall j \in[1,k]/\{i\},\ d_A,(p,B_i)\} \quad \text{------------(14)}$$

The immersion process and its results along with catchment basins and watersheds can be described in terms of geodesic influence zone.

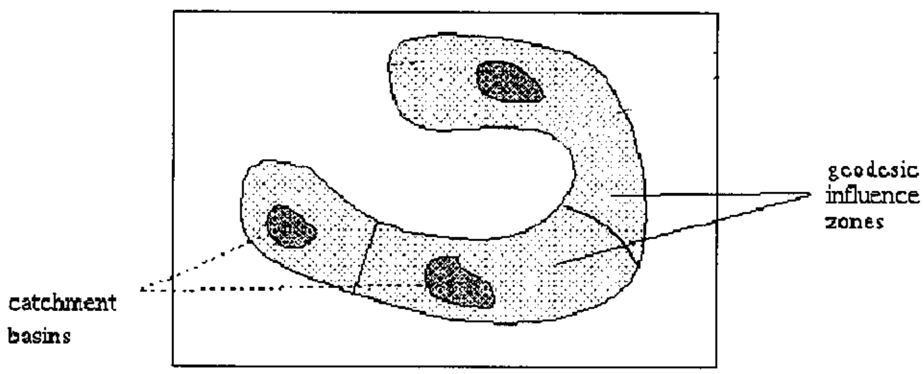

Figure 3.14: Illustration of Geodesic Influence Zone

Definition: Catchment Basins and Watersheds by Immersion

The set of the catchment basins of the gray-scale image I is equal to the set X_{hmax} obtained after the following recursion:

$$1.\ X_{h_{min}} = T_{h_{min}}(I),\ T_h(I) = \{p \in D_I,\ I(p) \leq h\}.$$
$$2.\ \forall h \in [h_{min}, h_{max} - 1],\ X_{h+1} = \min_{h+1} \bigcup iz_{T_{h+1}(I)}(X_h).$$

-------------(15)

The immersion procedure is done in the recursion, and the watersheds of I correspond to the set of the points of D_I which do not belong to any catchment basin.

From these basic concepts and primitive algorithm, Luc and Pierre improved the immersion-based algorithm for watersheds extraction. The algorithm consists of two important
steps, sorting and flooding and the queue structure is used to make computation time shorter. A mathematical dilation process is used to find new local minimum candidates and to find the geodesic influence zone in flooding.

In the topographic representation of a given image I, the numerical value (i.e., the gray tone) of each pixel stands for the elevation at this point.
Roughly speaking, it is based on a sorting of the pixels in the increasing order of their gray values, and on fast breadth-first scannings of the plateaus enabled by a first-in-first-out type data structure. Its adaptation to any kind of underlying grid (4-, 6-, &connectivity . . .) is straightforward, and it can be fairly easily extended to n-dimensional images and even to graphs.

Our implementation, here decomposed is into two steps:

- initial sorting
- flooding step

The Sorting Step

Among the vast number of available sorting techniques [19], one is particularly suited to the present problem. It is a distributive algorithm [19] which resorts to address calculations. This technique was introduced by E. J. Isaac and R. C. Singleton in [20] and is briefly described in [19, pp. 162-1661. The procedure first determines all the exact frequency distribution of each image gray level. The cumulative frequency distribution is then computed. This induces the direct assignment of each pixel to a unique cell in the sorted array.

Let us denote n the number of image pixels and h_{min} and h,, the lowest and largest gray levels, respectively. The present sorting technique has the considerable advantage of requiring only 2n "look and do" operations-one for determining the frequency distribution and the other for the assignment-plus h,, - h_{min} - 1 additions to get the cumulative frequency distribution.

The Flooding Step

Once the pixels have been sorted, we proceed to the progressive flooding of the catchment basins of the image. Suppose the flooding has been done up to a given level h. Every catchment basin already discovered-i.e., every catchment basin whose corresponding minimum has an altitude lower or equal to h-is supposed to have a unique label. Thanks to the initial sorting, we now access the pixels of altitude h + 1 directly and given them a special value, say MASK. Those pixels among them which have an already labeled pixel as one of their neighbors are put into the queue. Starting from these pixels, the queue structure enables to extend the labeled catchment basins inside the mask of pixels having value MASK, by computing geodesic influence zones. After this step, only the minima at level h + 1 have not been reached. Indeed, they are not connected to any of the already labeled catchment basins. Therefore, a second scanning of the pixels at level h + 1 is necessary to detect the pixels which still have value MASK. and to give a new label to the thus discovered catchment basins.

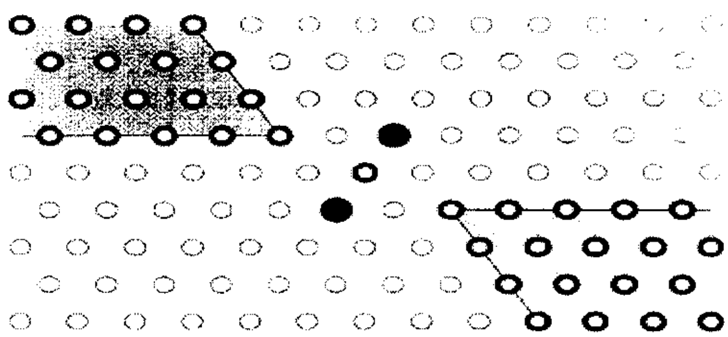

Fig 3.15 According to the hexagonal distance, all the bold pixels (gray areas) are equidistant to the two black ones.

The queue which is used is a first-in-first-out data structure: the pixels which are first put into it are those which can first be extracted. In practice a queue is simply a large enough array of pointers to pixels, on which three operations may be performed:

fifo-add(p) Puts the (pointer to) pixel p into the queue.

fifo-first() Returns the (pointer to) pixel which is at the beginning of the queue, and removes it from the queue.

fifo-empty () Returns true if the queue is empty and false otherwise.

In order to implement such operations, a kind of "circular" queue is one of the most efficient choices: the array representing our FIFO structure is addressed by two indexes, ptr-first and ptr-last. Each time a new element is put into the queue, it is stored at the address toward which ptr-last is pointing. prt-last is then incremented. When the limit of the array is reached, this index loops back to the beginning of the array. Similarly, ptr- first is a pointer toward the first element which can be removed from the structure, and is incremented after each removal. It may also loop back to the start of the array.

Not only does the use of a queue of pixels speed up the computations, it also allows us to solve the accuracy problems.

CHAPTER 4

IMPLEMENTATION RESULTS

4.1 SEGMENTATION : CAMERAMAN IMAGE

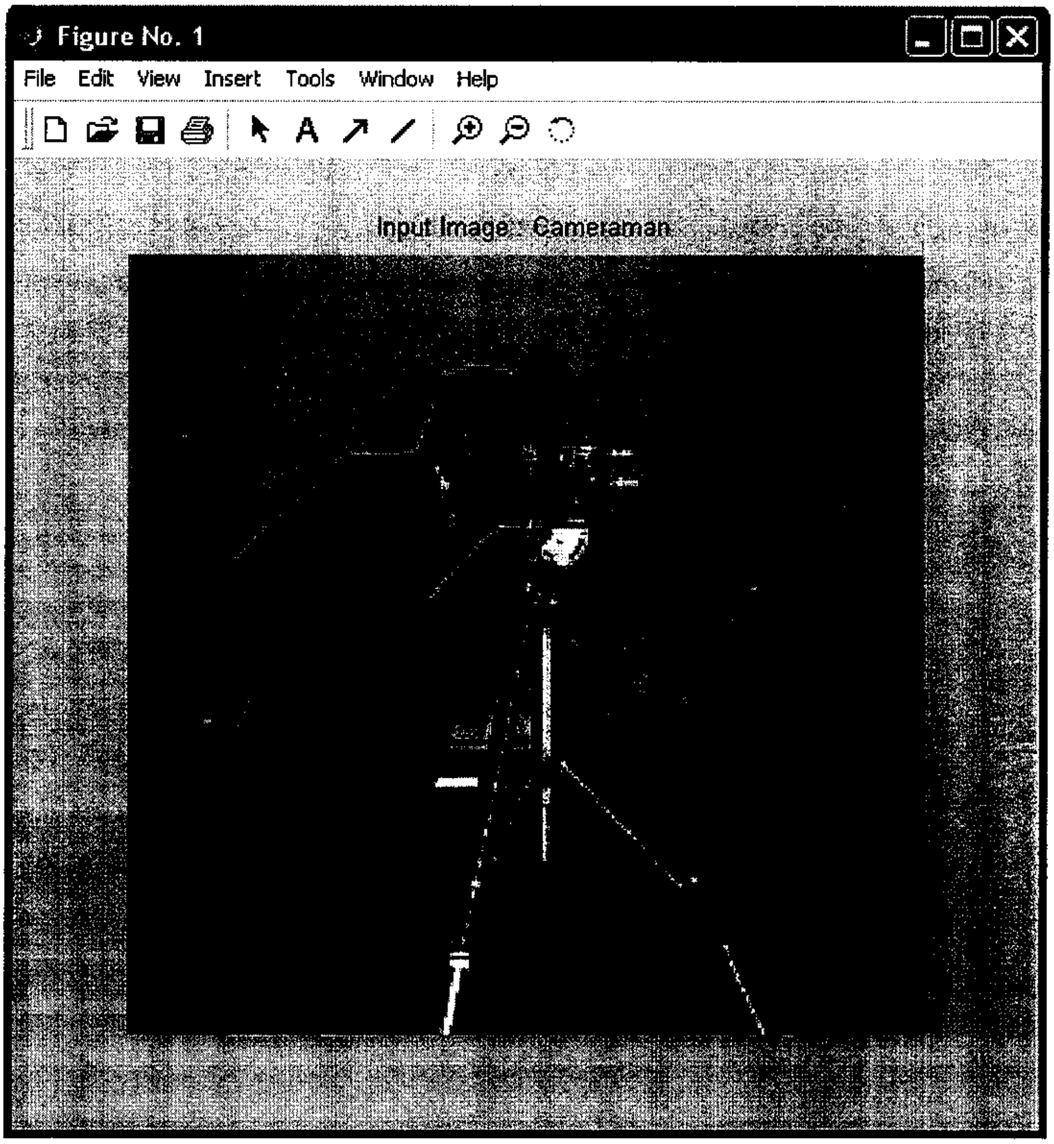

Figure 4.1.1 Input image of Cameraman image for image segmentation

Figure 4.1.2 Texture Gradient image of Cameraman Image

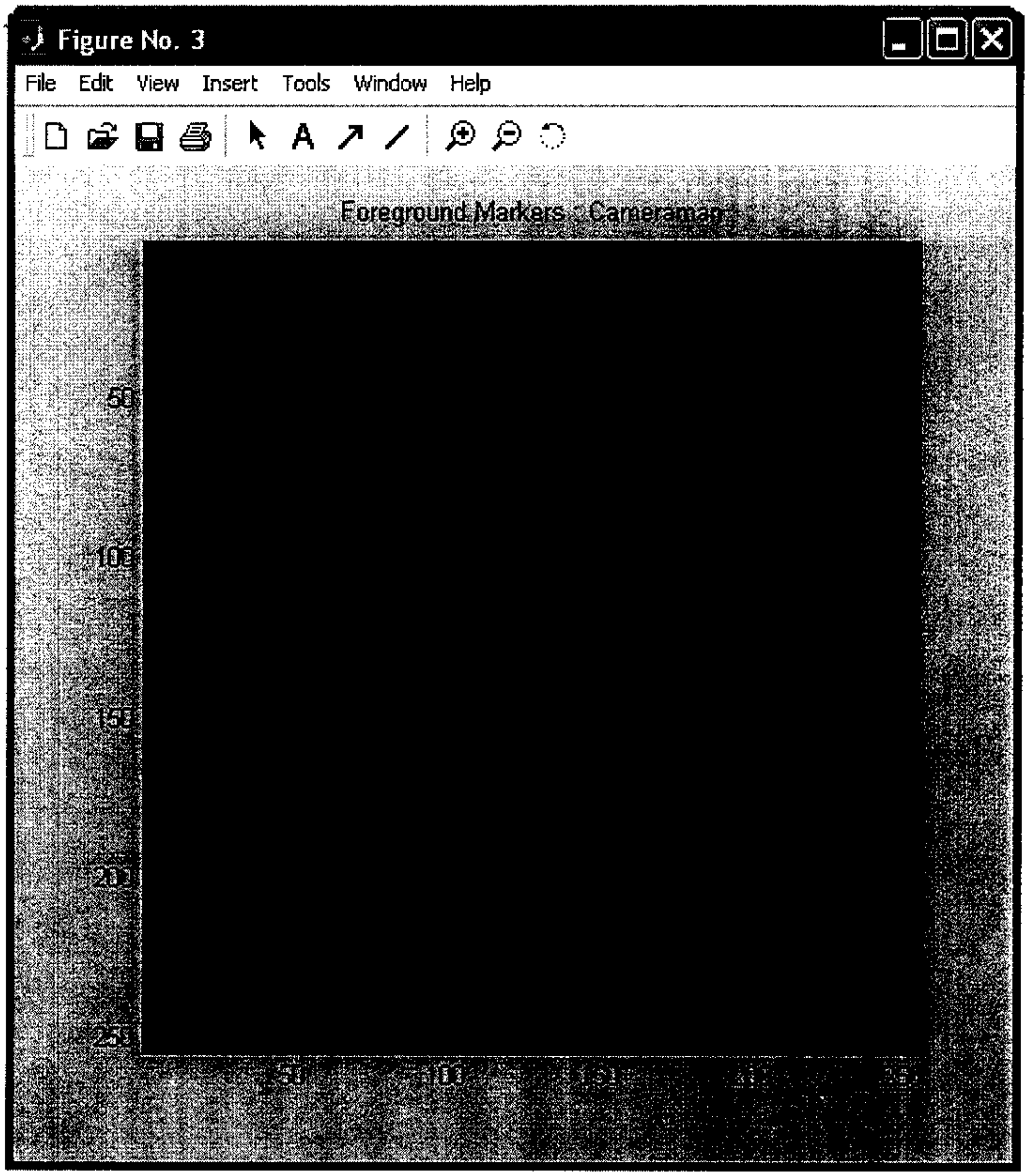

Figure 4.1.3 : Foreground Marker for Cameraman image

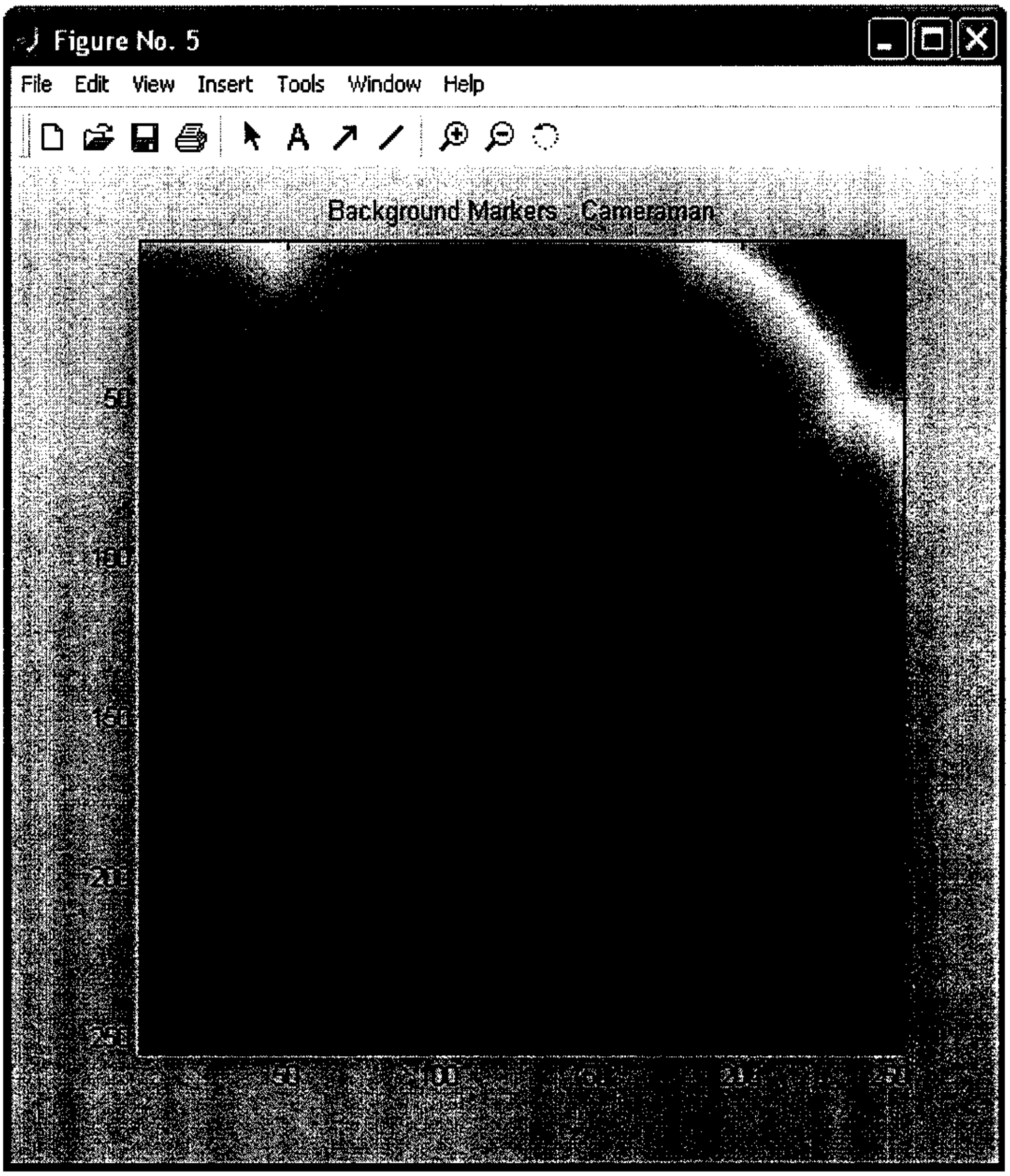

Figure 4.1.4 Background Marker Extracted for Cameraman Image

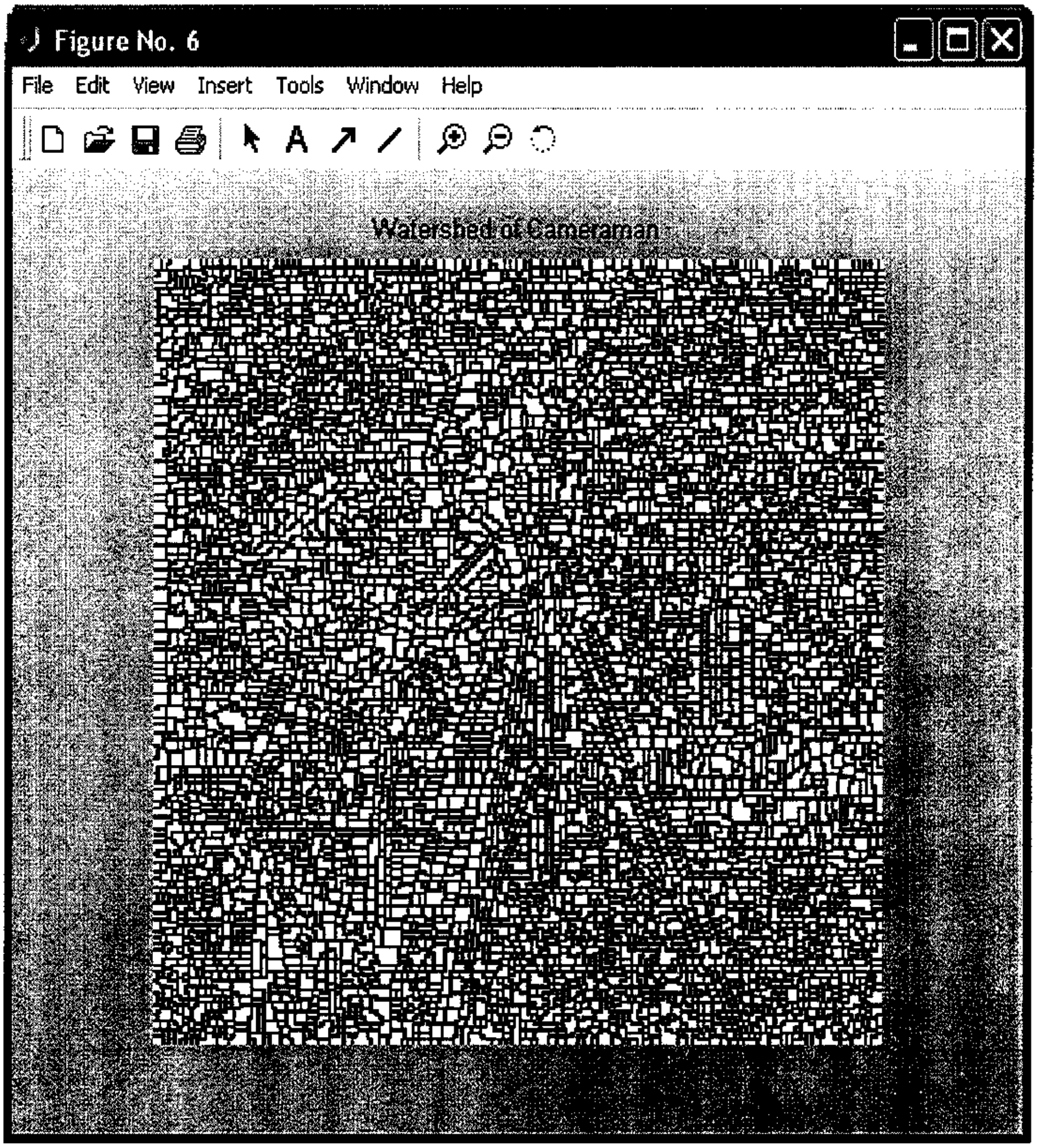

Figure 4.1.5 Watershed extracted image of Cameraman Image

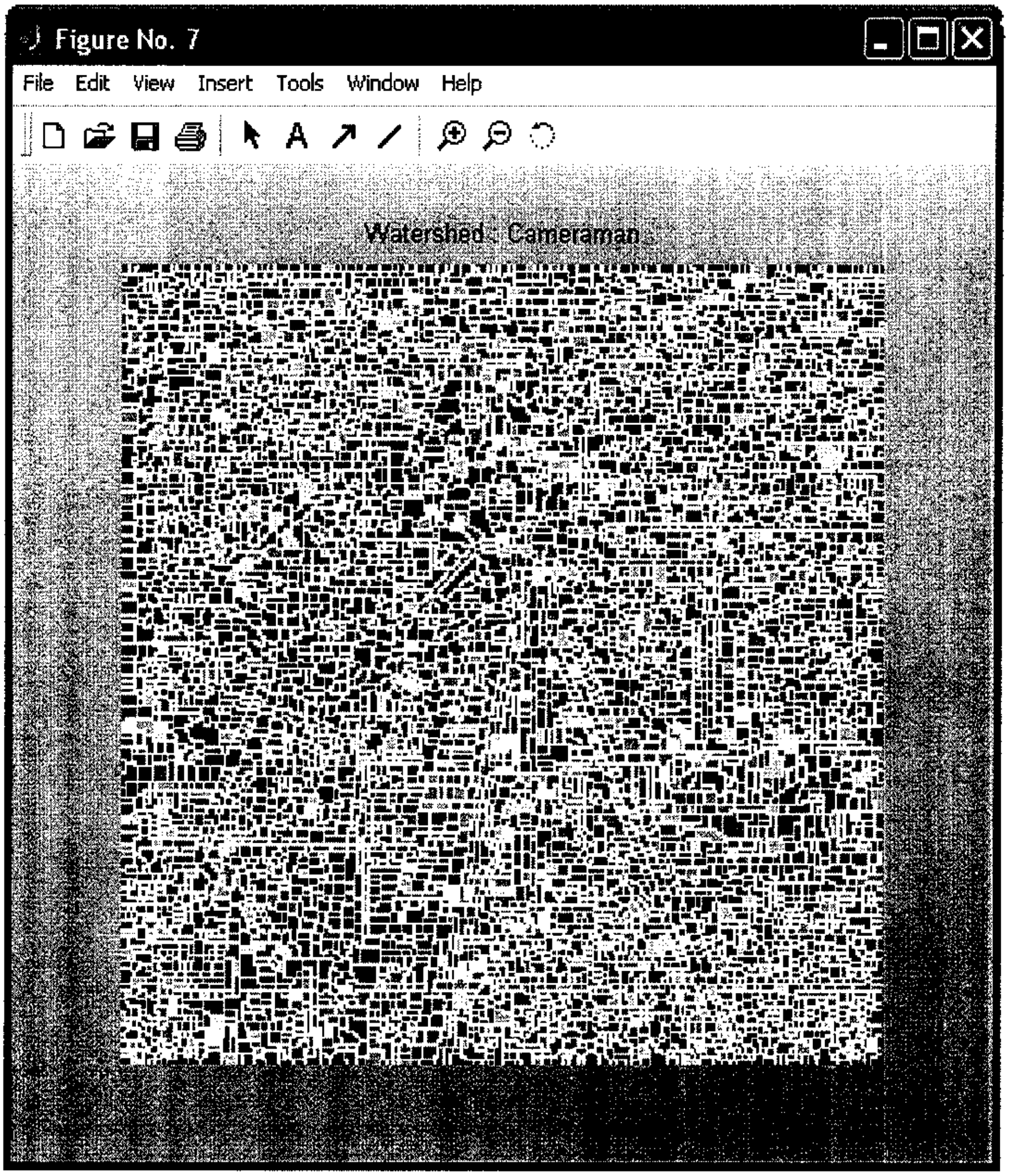

Figure 4.1.6 Watershed extracted image of Cameraman Image in rgb

4.2 SEGMENTATION : LENA IMAGE

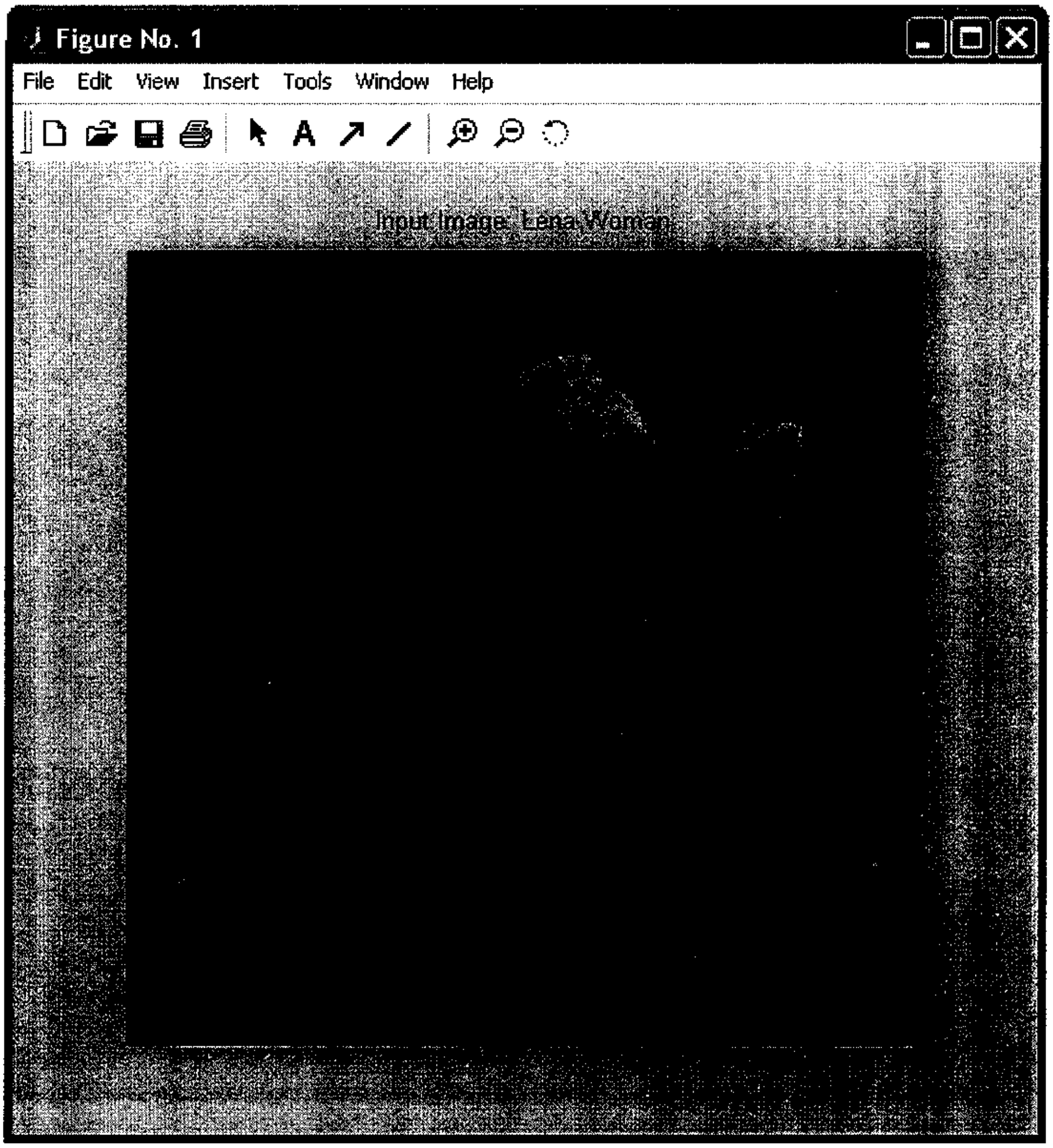

Figure 4.2.1 Input image of Lena image for image segmentation

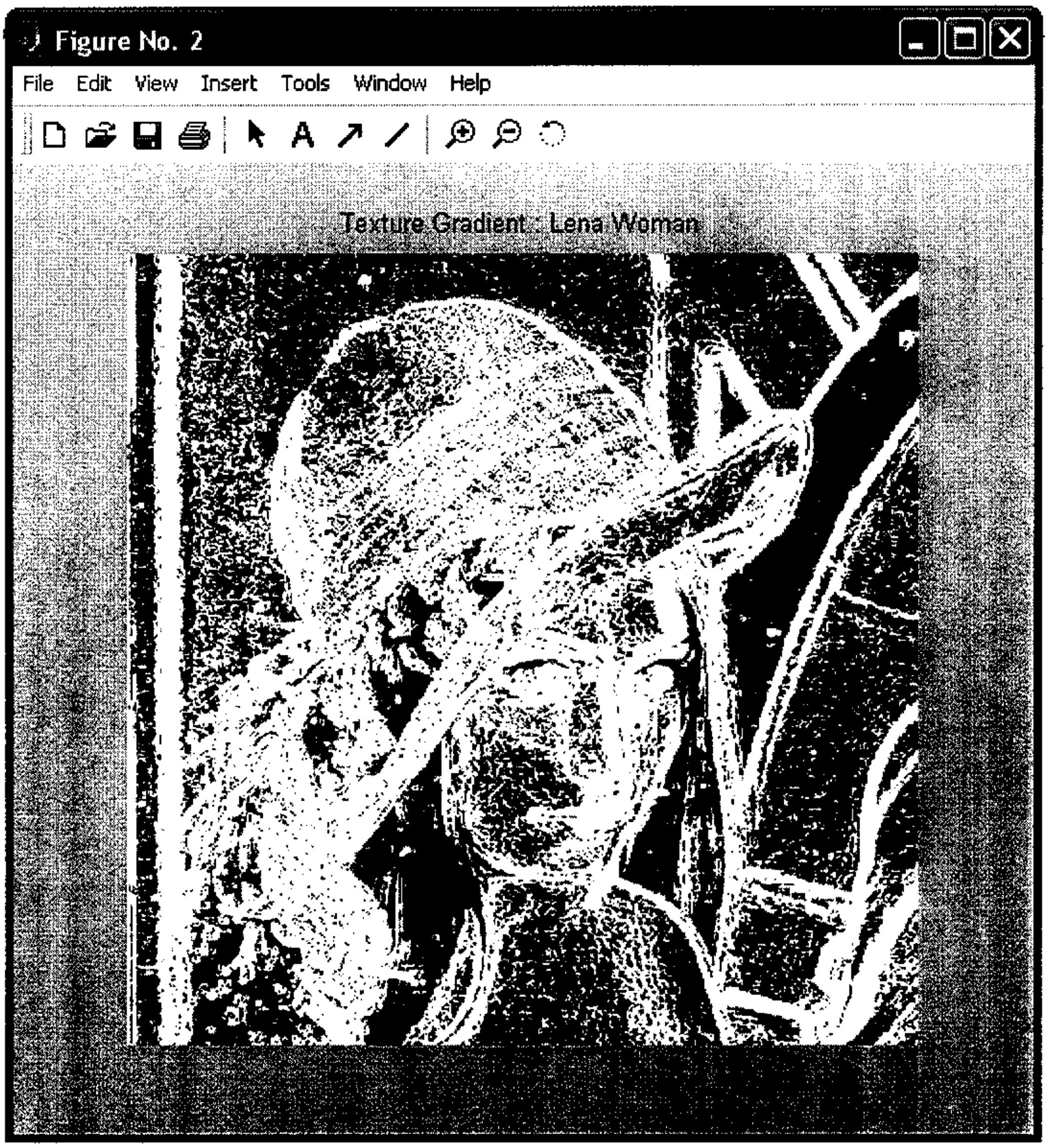

Figure 4.2.2 Texture Gradient image of Lena Image

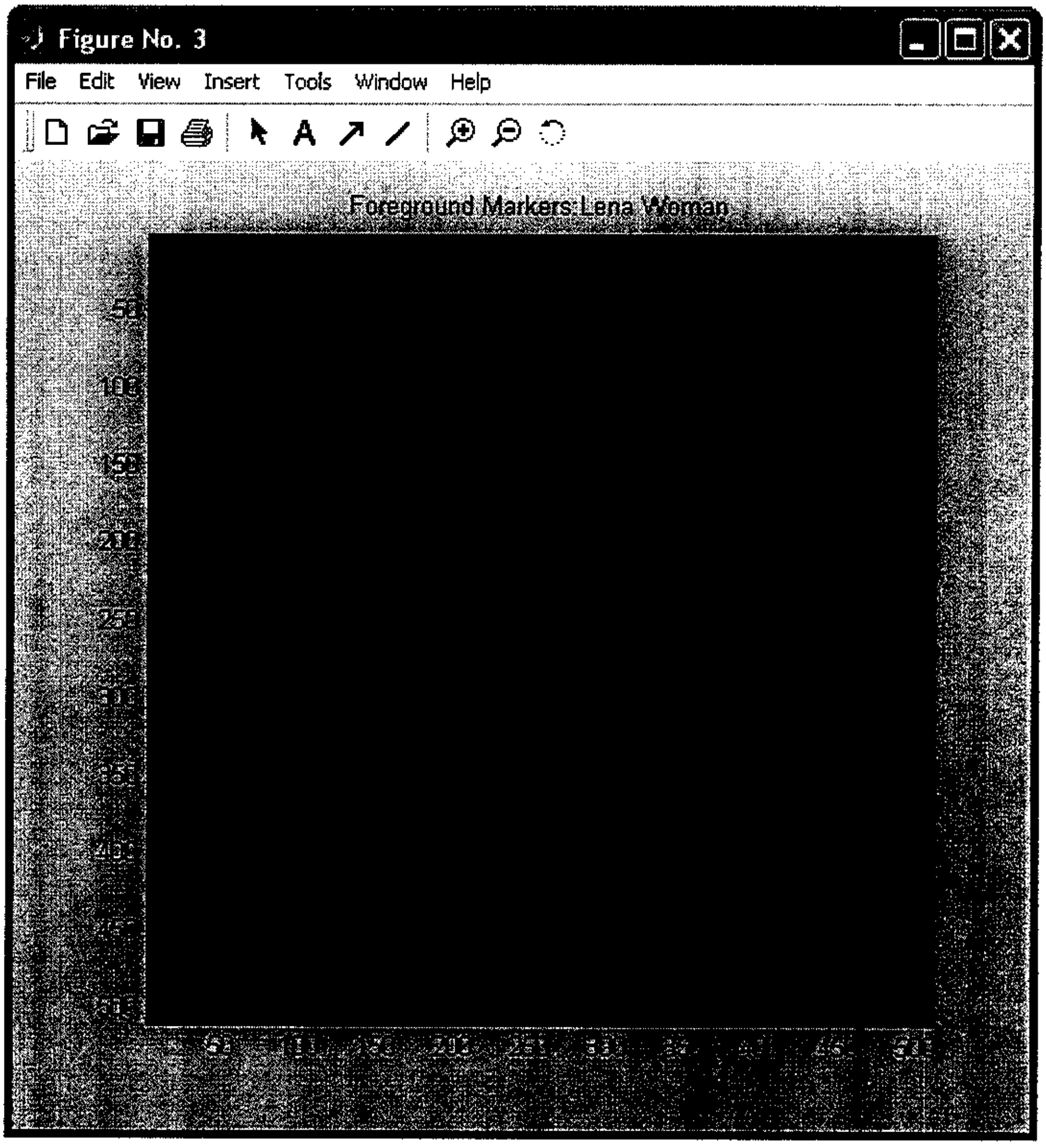

Figure 4.2.3 Foreground Marker for Lena Image

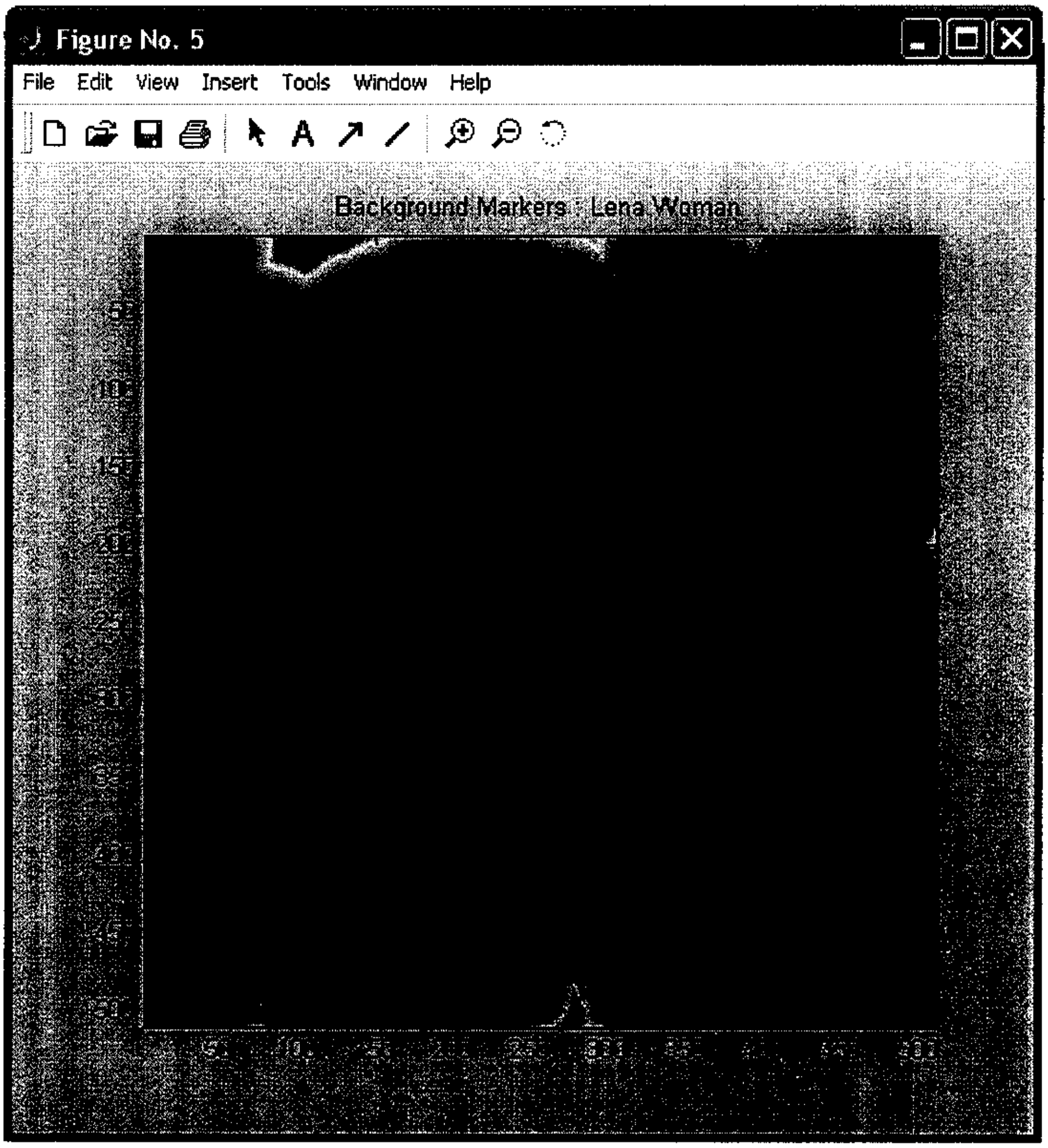

Figure 4.2.4 Background Marker Extracted for Lena Image

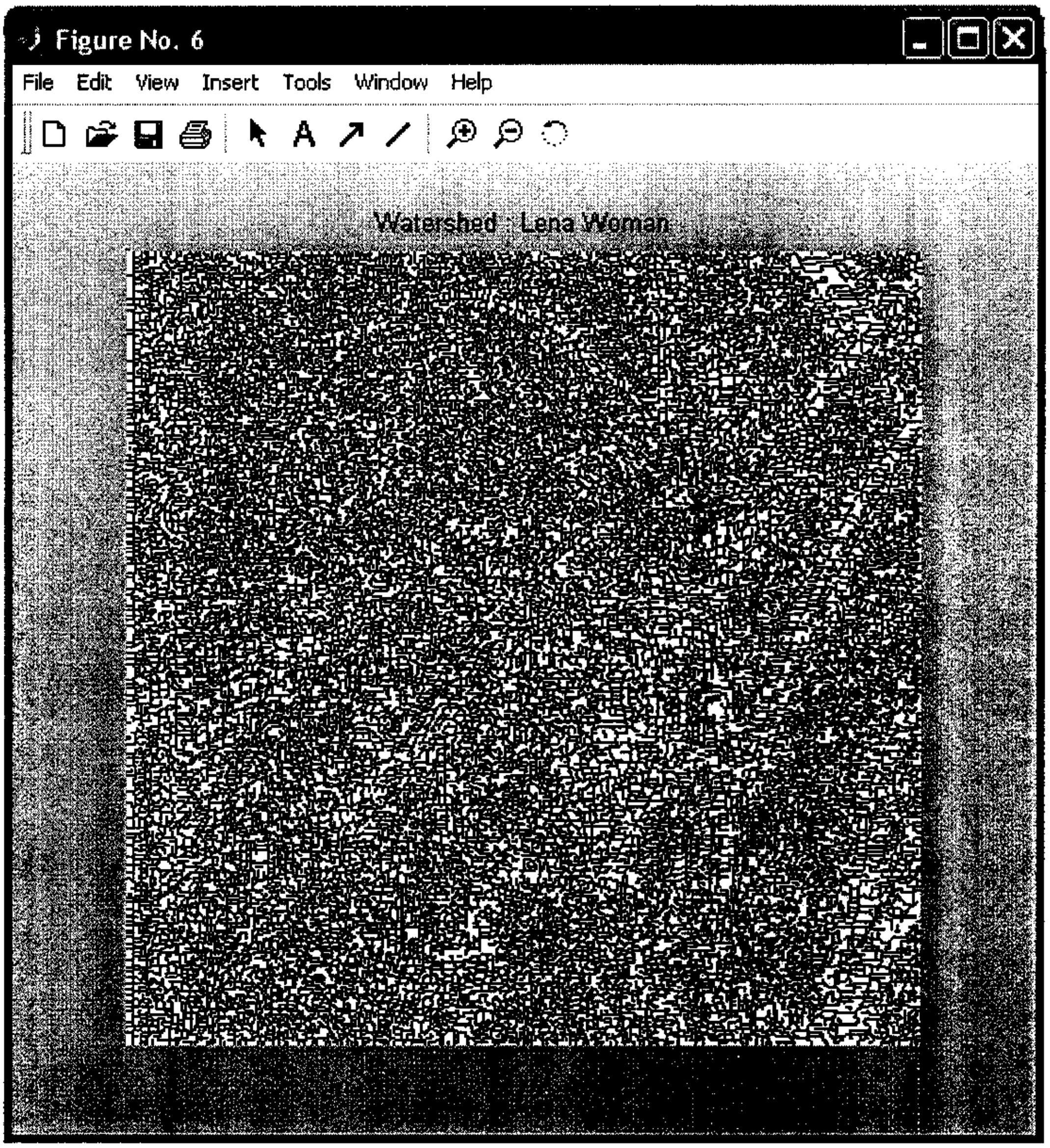

Figure 4.2.5 Watershed extracted image of Lena Image

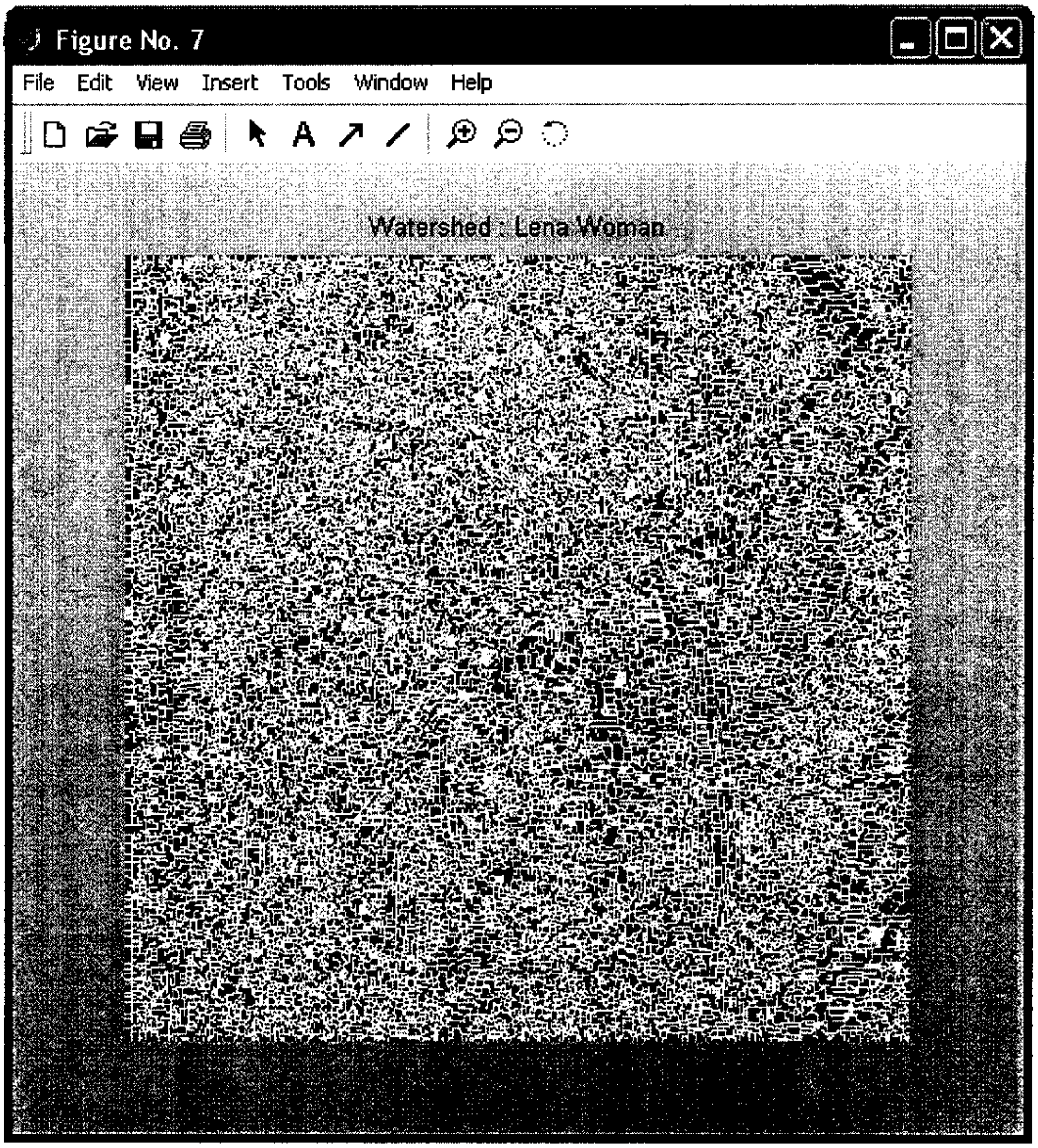

Figure 4.2.6 Watershed extracted image of Lena Image in rgb

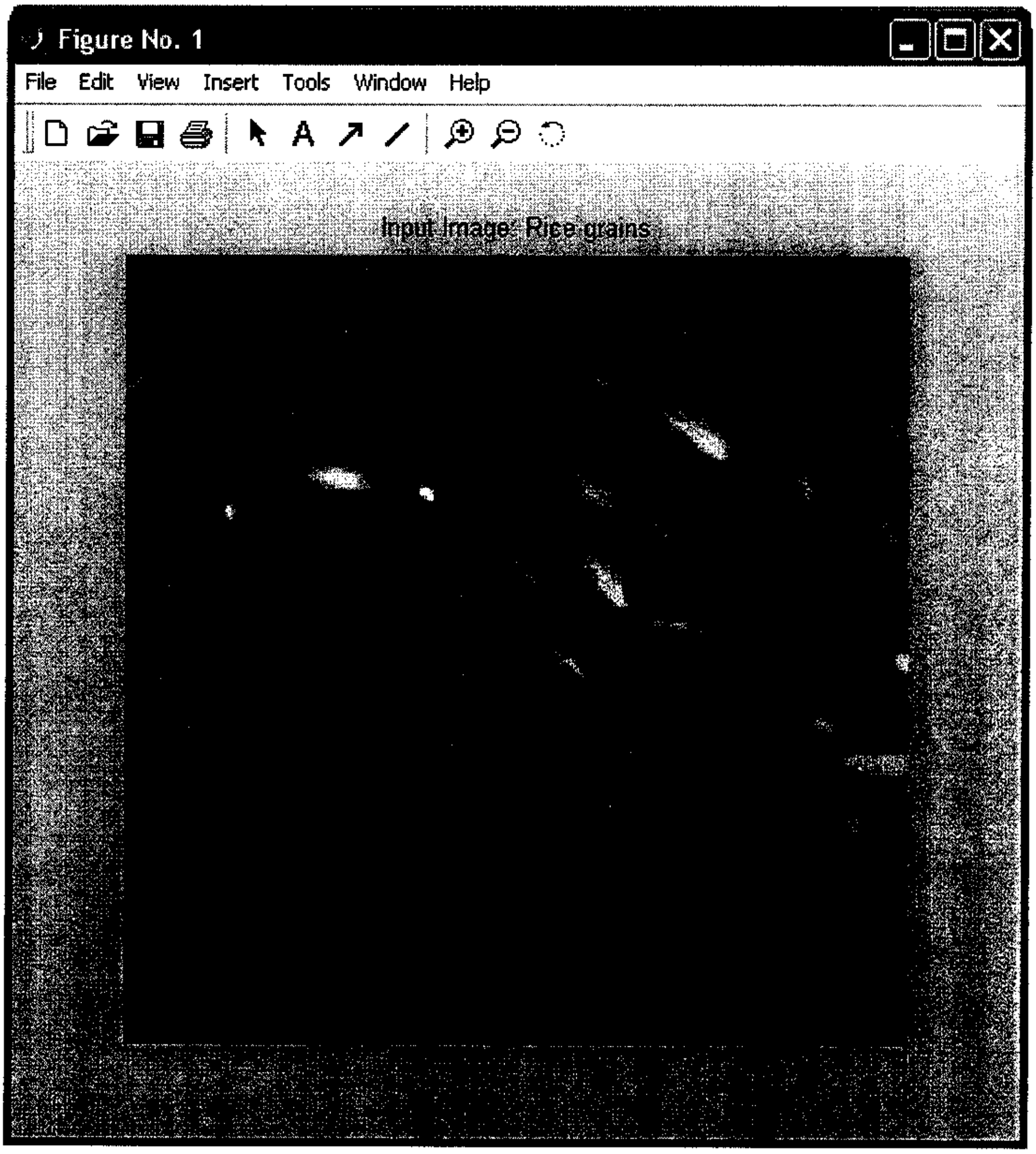

Figure 4.3.1 Input image of Rice Grain for image segmentation

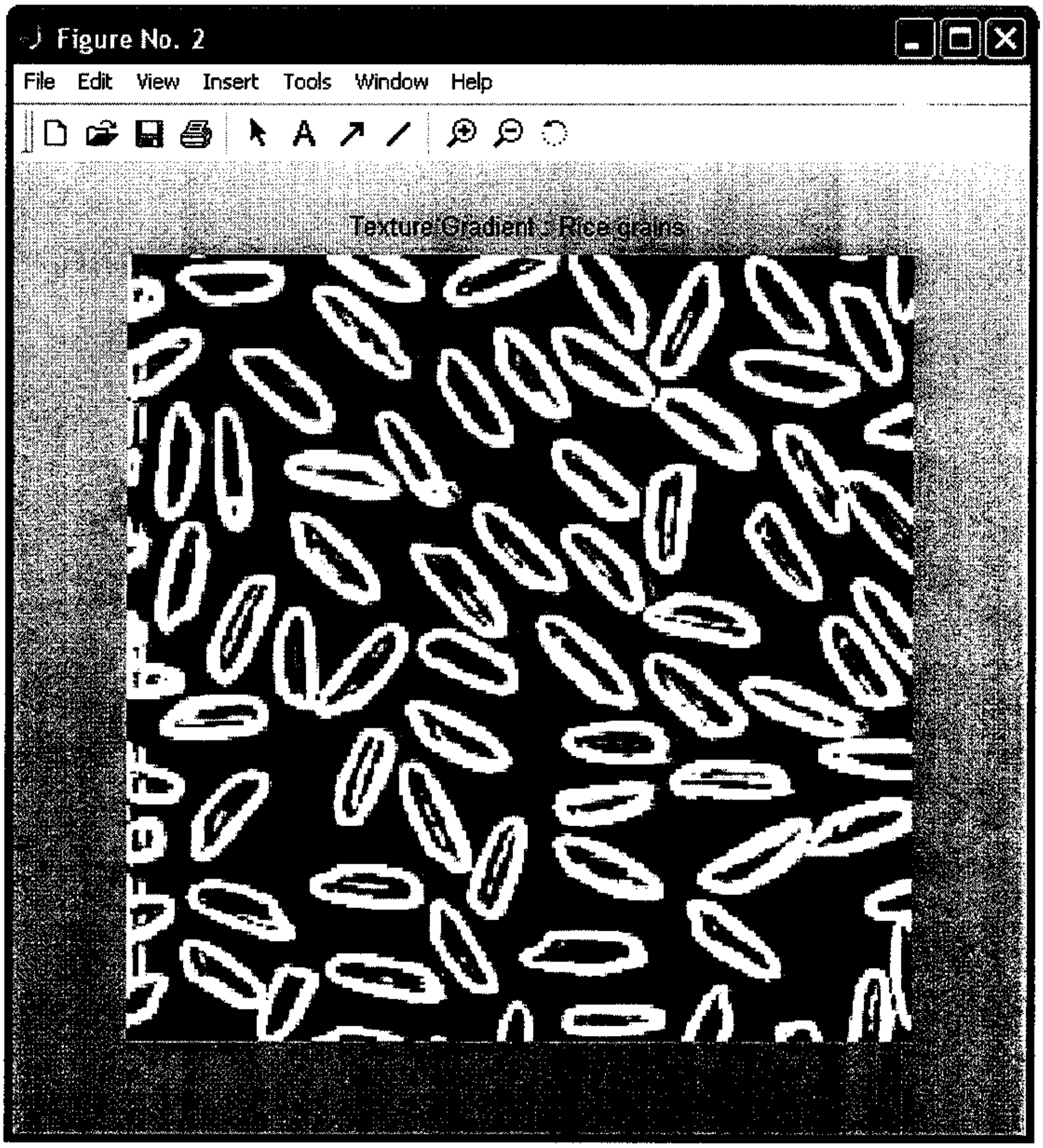

Figure 4.3.2 Texture Gradient image of Rice Grain Image

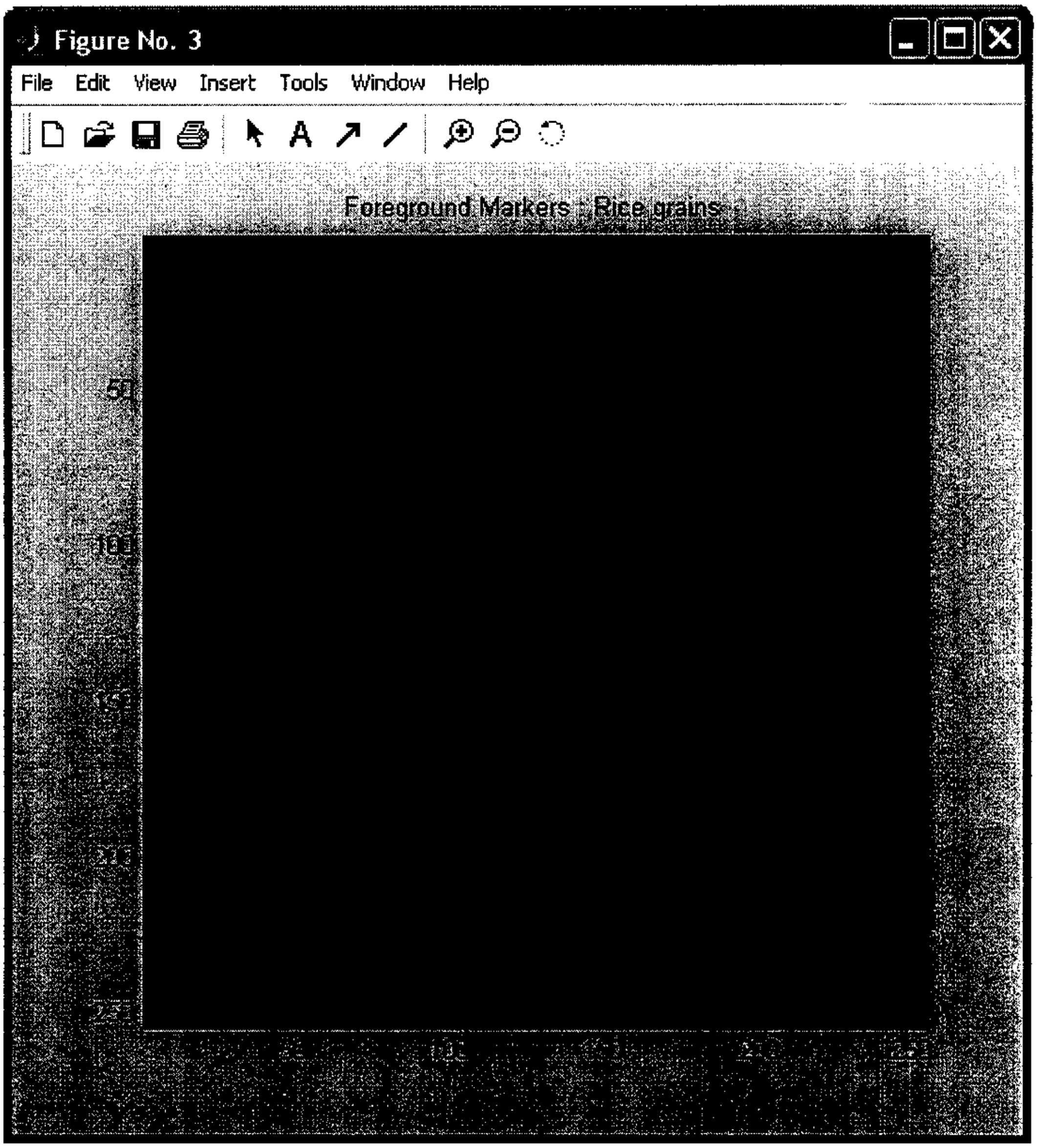

Figure 4.3.3 Foreground Marker for Rice Grain Image

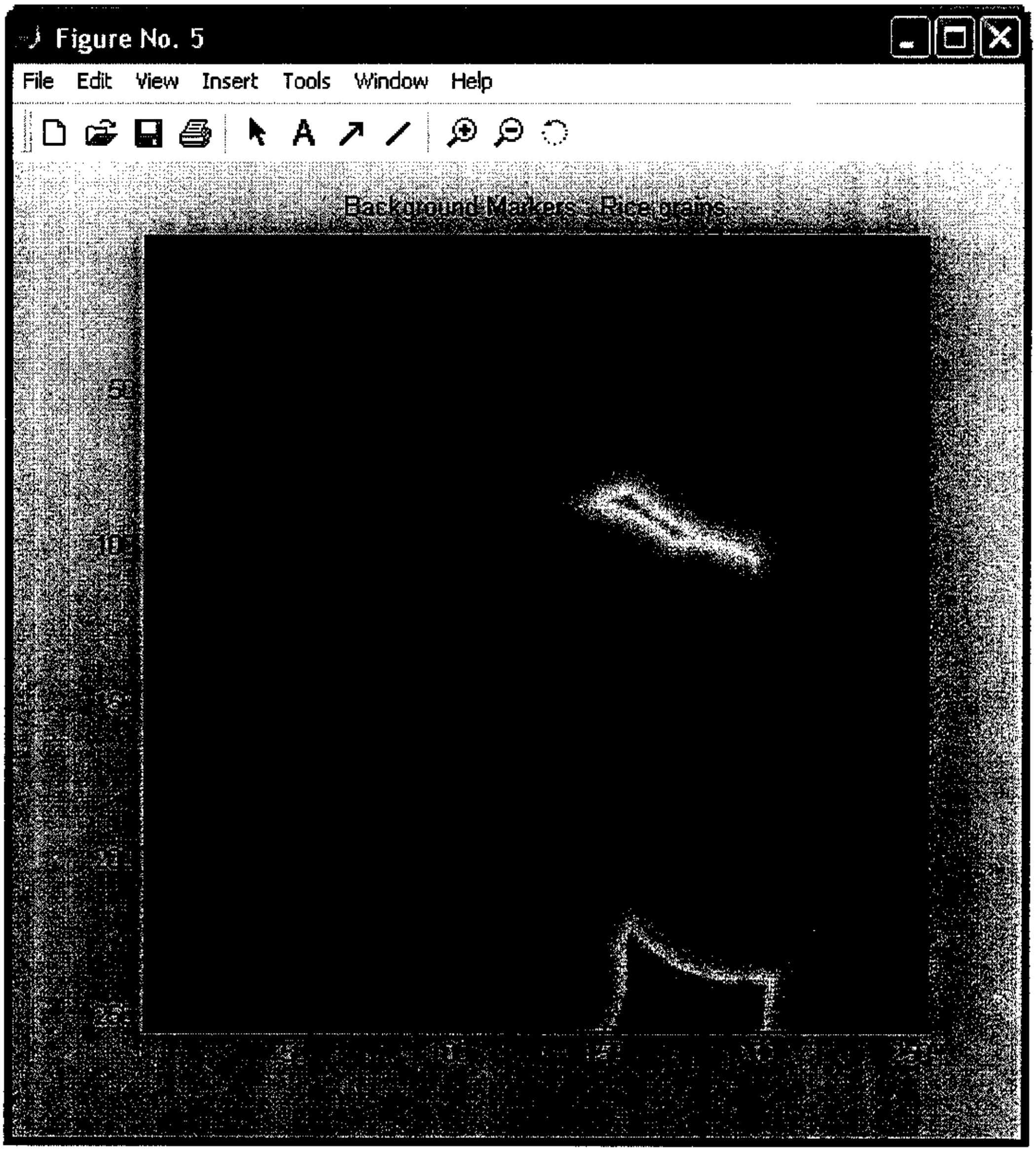

Figure 4.3.4 Background Marker Extracted for Rice Grain Image

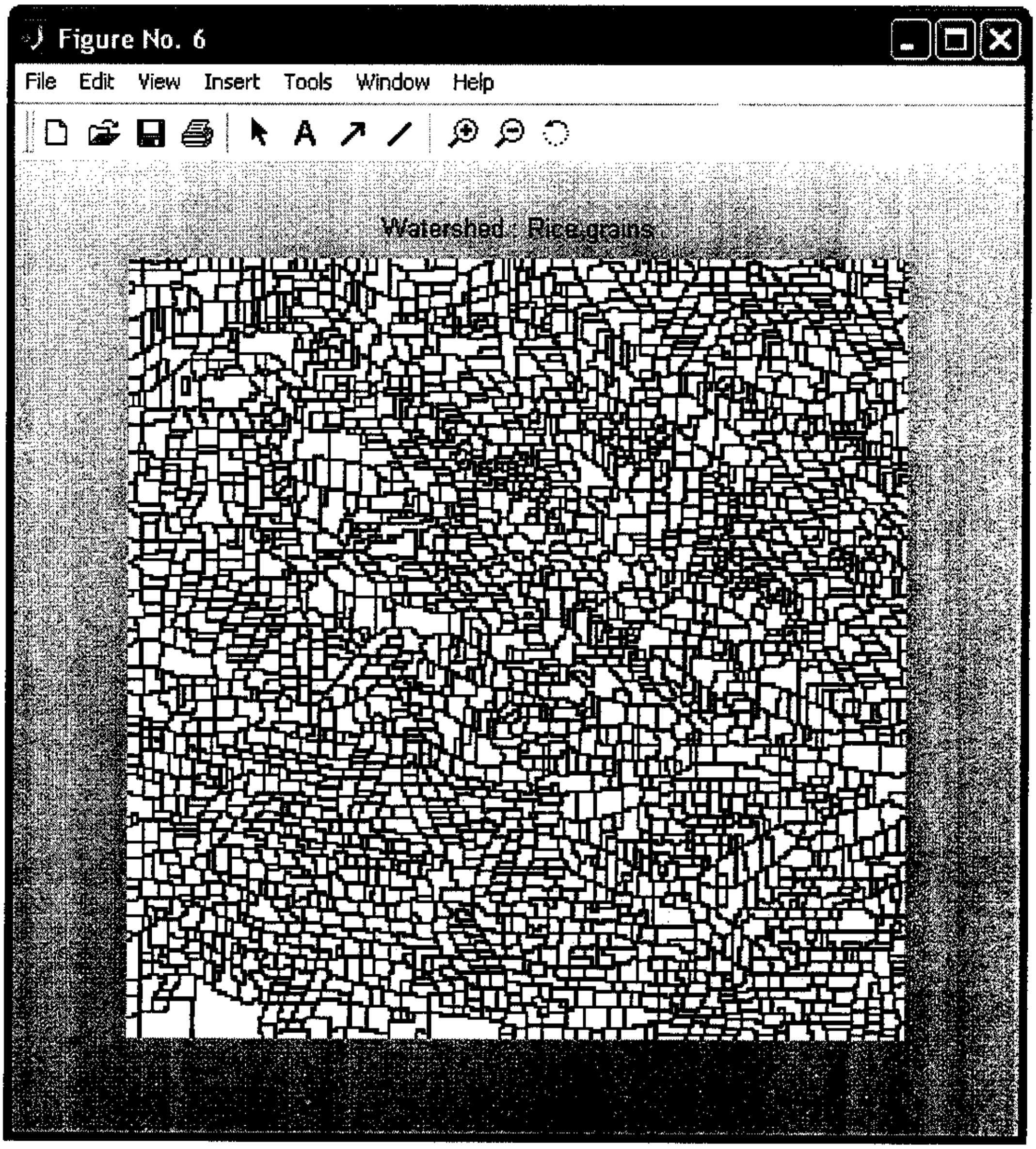

Figure 4.3.5 Watershed extracted image of Rice Grain

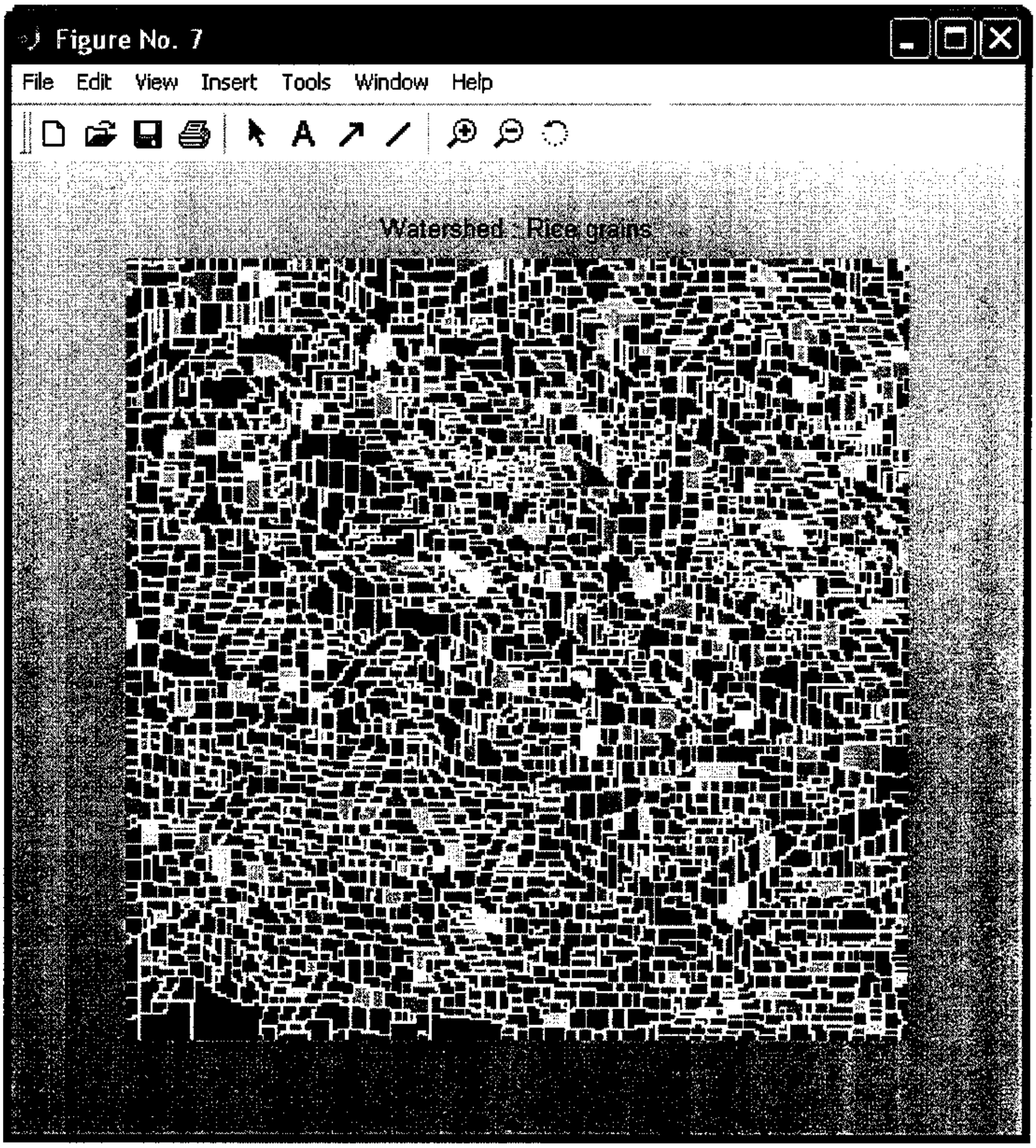

Figure 4.3.6 Watershed extracted image of Rice Grain Image in rgb

CHAPTER 5

CONCLUSION AND FJTURE WORK

5.1 CONCLUSION

In this project, the non decimated complex wavelet packet transform is constructed using Best basis selection algorithm having cost function as energy. The image is then median filtered to perform smoothening and the gradient of the image is extracted using Roberts operator. Then the textured gradient is extracted because it is more efficient for homogeneous region .Then the image is segmented using immersion procedure where the catchment basins are flooded from the minima. To prevent over segmentation the marker driven segmentation algorithm is developed.

5.2 FUTURE WORK

In this project the Roberts Operator is used. Though the computation speed can be increased it could result in less clarity of image. This could be enhanced by using the Guassian derivative gradient. Moreover the entire image segmentation procedure could be applied to more sophisticated applications in medical fields where the proper segmentation of images is very important.

REFERENCES

[1] G. Beylkin. On the representation of operators in bases of compactly supported wavelets. SIAM J. Numer.Anal., 29 (1992), 1716-1740

[2] I. Daubechies. Ten lectures on Wavelets. SIAM,Philadelphia,PA,1992.

[3] H. Guo. Theory and Applications of the Shift- Invariant, Time-Varying and Undecimated Wavelet Transforms.

[4] G. Nason and B. Silverman. The stationary wavelet transform and some statistical applications, in Wavelets and Statistics(Antoniadis and Oppenheim eds.), 281{299, 1995.

[5] M. J. Shensa. The discrete wavelet transform: wedding the μa trous and Mallat algorithms. IEEE Trans. Inform.

[6] Holschneider, M., Kronland-Martinet, R., Morlet, J., and Tchamitchian, Ph.: A real-Time algorithm for signal analysis with the help of the wavelet transform, in Wavelets, Time- Frequency Methods and Phase space, J.M.Combes, A. Grossmann, and Ph. Tchamitchian (eds.), Springer-Verlag, Berlin, pp. 286{297, 1987.

[7] N G Kingsbury, "The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters", *Proc. 8th IEEE DSP Workshop*, Bryce Canyon, Aug1998.

[8] N G Kingsbury: "The dual-tree complex wavelet transform: a new efficient tool for image restoration and enhancement", *Proc. EUSIPCO 98*, Rhodes, Sept 1998.

[9] M. Krasner, "The critical band coder-digital encoding of speech signals based on the perčeptuāl requirements of the auditory system,''in *Proc. 1980 IEEE lnt. Con$ Acoust., Speech, Signal Processing,*Denver, CO, Apr. 1980, pp. 327-331,.

[1 0] T . P. Barnwell, "Subband coder design incorporating recursive Quadrature filters And optimum ADPCM coders," *IEEE Trans. Acoust.,Speech, Signal Processing,* vol. ASSP- 30, pp. 751-765, Oct. 1982:

[11] D. Esteban and C. Galand, "Application of quadrature mirror filters to split band voice coding schemes," in *Proc. 1977 IEEE Int. Con5 Acoust., Speech, Signal Processing* Hartford, CT, May 1977, pp.

[12] N. G. KINGSBURY. The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters . In *Proceedings of the IEEE Digital Signal ProcessingWorkshop*, 1998.

[13] J F A Magarey and N G Kingsbury, "Motion estimation using a complex-valued wavelet transform", *IEEE Trans. Signal Proc.*, 46(4), April 1998.

[14] D B H Tay and N G Kingsbury, "Flexible design of multidimensional perfect reconstruction FIR 2-band filters using transformations of variables", *IEEE Trans. Image Proc.*,2(4), pp 466-480, 1993.

[15] N G Kingsbury: "Shift invariant properties of the Dual-Tree Complex Wavelet Transform

[16] R. Coifman, Y. Meyer, S. Quake, and V. Wickerhauser,"Signal processing and Compression with wave packets", Numerical Algorithms Research Group,New Haven, CT: Yale university, 1990.

[17] T. Chang, and C.C.J. Kuo. "Texture analysis and classification with tree-structured wavelet transform",IEEE Trans. Image Processing, vol. 2 p.429-441,April,1993

[18] R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," IEEE Transactions on Information Theory, vol. 38, no. 2, pp. 713-718,1992.

[19] H.Lorin,sorting and Sort Sytems(The Systems Programming Series).Reading ,MA:Addison–Wesley ,1975

[20] E.J,Issac and R.C. Singleton ,"Sorting by address calculation,"J.ACM,vol.3,pp 169-174 ,1956