**ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF INFORMATICS**

# A DISCRETE FOURIER TRANSFORM BASED SUBBAND DECOMPOSITION APPROACH FOR THE SEGMENTATION OF REMOTELY SENSED IMAGES

**M. Sc. Thesis by**
**Mehmet Enver ERGÜVEN, B. Sc.**

**Department: Advanced Technologies in Engineering**
**Programme: Satellite Communication and Remote Sensing**

**Supervisor: Assoc. Prof. Dr. Işın ERER**

MAY 2006

## FOREWORD

First, I would like to thank my supervisor Assoc. Prof. Dr. Işın ERER who provided me the opportunity to work with her. I appreciate her guidance during this study. Second, I thank Murat Sezgin who is a doctoral student at Informatics Institute of Istanbul Technical University and Okan K. Ersoy who is a Professor of the Purdue University; they provided the prototypes of the algorithms developed in this work from their earlier study on the work.

I also want to express my deepest thanks to my home mates and colleagues one by one who have always been patient with me. Mentioning their names here would be an insufficient attempt to express what they have meant to me.

Finally and most importantly I want to thank my parents for their support, tolerance and patience, and especially my sister Rabia who is the youngest member of the family. Without their existence completion of this work and many other things would have been impossible.

May, 2006. Mehmet Enver ERGÜVEN

**CONTENTS**

**FOREWORD** ..... **ii**
**CONTENTS** ..... **iii**
**ABBREVIATIONS** ..... **v**
**LIST OF TABLES** ..... **vi**
**LIST OF FIGURES** ..... **vii**
**LIST OF SYMBOLS** ..... **ix**
**SUMMARY** ..... **x**
**ÖZET** ..... **xi**
**1. INTRODUCTION** ..... **1**
**2. TEXTURE FEATURE EXTRACTION AND SEGMENTATION USING SUBBAND DECOMPOSITION** ..... **6**
2.1 Feature Extraction Methods ..... 7
2.1.1 Statistical Methods ..... 7
2.1.2 Model Based Methods ..... 7
2.1.3 Transform Based Methods ..... 7
2.2 Subband Decomposition and Filter Bank Approach ..... 8
2.2.1 Pyramid Structured Decomposition ..... 9
2.3 Feature Extraction and Selection ..... 10
2.3.1 Normalizing Feature Vector ..... 11
2.3.2 Calculating the Mean of the Feature Vectors ..... 12
2.4 Segmentation ..... 13
2.5 Classification ..... 13
2.5.1 Classifiers and Decision Rules ..... 14
2.6 Post-processing Operations ..... 16
**3. SUBBAND DECOMPOSITION AND FILTER BANKS** ..... **17**
3.1 1-Dimensional Subband Decomposition ..... 17
3.1.1 Quadrature Mirror Filters (QMF) ..... 17
3.1.2 Decimation and Interpolation in 1-Dimension ..... 17
3.2 1-Dimensional 2-Channel Filter Banks ..... 20
3.3 Perfect Reconstruction Conditions for 2-Channel Filter Banks ..... 22
3.4 2-Dimensional 2-Channel Subband Decomposition ..... 23
**4. WAVELET TRANSFORM BASED SUBBAND DECOMPOSITION** ..... **25**
4.1 Linear Orthogonal Transforms ..... 25
4.2 Fourier Transform ..... 27
4.3 Short Time Fourier Transform ..... 28
4.4 Wavelet Transform ..... 30
4.4.1 Continuous Wavelet Transform ..... 31
4.4.2 Discrete Wavelet Transform ..... 32
4.5 Multiresolution Decomposition and Reconstruction ..... 34
4.6 Discrete Stationary Wavelet Transform ..... 36
4.7 2-Dimensional Discrete Wavelet Transform ..... 37

**5. SUBBAND DECOMPOSITION USING DFT ..... 39**
5.1 Decomposition Algorithm..... 39
5.1.1 Discrete Fourier Transform..... 40
5.1.2 Ideal High-Pass and Ideal Low-Pass Filtering and Decimation ..... 40
5.1.3 Reconstruction in the Synthesis Bank..... 42
5.2 Ideal LP and HP Filters Used to Decompose and Reconstruct a Signal... 43
5.2.1 An Example of a 1-Dimensional 3-Level Subband Decomposition Using DFT…… ..... 45
5.3 The Proposed DFT Transform Based Feature Extraction and Image Segmentation Algorithm ..... 46
5.3.1 Flow Diagram of DFT Subband Decomposition ..... 46
5.3.2 Training Steps ..... 48
5.3.3 Segmentation Steps ..... 48
**6. EXPERIMENTAL RESULTS..... 51**
6.1 Simulation 1: 2-Classes Segmentation Problem ..... 52
6.1.1 Comparison of Decimated DFT and DWT Based Subband Decomposition Methods..... 52
6.1.2 Comparison of Undecimated DFT and SWT Based Subband Decomposition Methods..... 53
6.2 Simulation 2: 4-Classes Segmentation Problem ..... 58
6.2.1 Comparison of Undecimated DFT and SWT Based Subband Decomposition Methods..... 58
6.3 Simulation 3: Detection of 4 Objects on a Background..... 61
6.3.1 Comparison of Decimated DFT and DWT ..... 61
6.3.2 Comparison of Undecimated DFT and SWT Based Subband Decomposition Methods..... 63
6.4 Simulation 4: Detecting Oil Spills on the North Sea ..... 67
6.4.1 Comparison of Decimated DFT, DWT, Undecimated DFT and SWT Based Subband Decomposition Methods..... 67
**7. CONCLUSIONS AND DISCUSSION ..... 72**
**REFERENCES..... 75**
**BIOGRAPHY ..... 78**

**ABBREVIATIONS**

| | | |
|---|---|---|
| ADD | : | Ayrık Dalgacık Dönüşümü |
| AR | : | Autoregressive |
| DCT | : | Discrete Cosine Transform |
| DDD | : | Durağan Dalgacık Dönüşümü |
| DFT | : | Discrete Fourier Transform |
| DST | : | Discrete Sine Transform |
| DWT | : | Discrete Wavelet Transform |
| FB | : | Filter Bank |
| FFT | : | Fast Fourier Transform |
| FIR | : | Finite Impulse Response |
| FT | : | Fourier Transform |
| GIS | : | Geographic Information Systems |
| HP | : | High Pass |
| IFFT | : | Inverse Fast Fourier Transform |
| IIR | : | Infinite Impulse Response |
| LP | : | Low Pass |
| LF | : | Lattice Filter |
| MR | : | Multiresolution |
| MRA | : | Multiresolution Analysis |
| QMF | : | Quadrature Mirror Filter |
| QMFB | : | Quadrature Mirror Filter Bank |
| RS | : | Remote Sensing |
| SAR | : | Synthetic Aperture Radar |
| STFT | : | Short Time Fourier Transform |
| SWT | : | Stationary Wavelet Transform |
| WT | : | Wavelet Transform |
| 1-D | : | One Dimensional |
| 2-D | : | Two Dimensional |

## LIST OF TABLES

**Page No**

**Table 6.1** : Accuracy Rates of Decimated DFT Based Subband Decomposition Method and DWT Based Subband Decomposition Methods…….52
**Table 6.2** : Accuracy Rates of Undecimated DFT and SWT Based Subband Decomposition Methods….....54
**Table 6.3** : Accuracy Rates of Undecimated DFT and SWT Based Subband Decomposition Methods…….....58
**Table 6.4** : Accuracy Rates of Decimated DFT and DWT Based Subband Decomposition Methods…….61
**Table 6.5** : Accuracy Rates of Undecimated DFT and SWT Based Subband Decomposition Methods…….63

**LIST OF FIGURES**

**Page No**


**Figure 2.1** General Model of a Feature Extraction And Segmentation System….6
**Figure 2.2** Pyramid Structured Subband Decomposition and Division of the Spectrum of Signal …….10
**Figure 2.3** An Example of a Feature Space for a 2- Class Segmentation Case, Each Feature Vector Contains 2 Elements…….…….12
**Figure 2.4** Decisions for a Test Pixel in the Segmented Feature Space by Measuring Euclidian Distances to the Means…..……...15
**Figure 3.1** Decimation and Interpolation by the Factor 2…….……18
**Figure 3.2** Analysis and Synthesis Parts of a 2-Channel QMFB…….…….20
**Figure 4.1** Linear Expansion of a Vector onto the Euclidian Space…….26
**Figure 4.2** 2-Dimensional Time-Scale Representation of the Signal……...30
**Figure 4.3** 1-Dimensional 2-Level Full Subband Decomposition…….35
**Figure 4.4** Interpolation and Reconstruction……36
**Figure 4.5** Separable Filtering Achieved Via Row and Column Processing……37
**Figure 4.6** Representation of Subbands of an Image Decomposed in 3-Levels...38
**Figure 5.1** Frequency Responses of Ideal HP and LP Filters. Low-Pass Decomposition Filter: $H_0$, High-Pass Decomposition Filter: $G_0$, Low-Pass Reconstruction Filter : $H_1$, High-Pass Reconstruction Filter: $G_1$ ( from left to right and top to bottom)……43
**Figure 5.2** Impulse Responses of Ideal HP and LP Filters. Low-Pass Decomposition Filter: $H_0$, High-Pass Decomposition Filter: $G_0$, Low-Pass Reconstruction Filter : $H_1$, High-Pass Reconstruction Filter: $G_1$ (from left to right and top to bottom)…….…….44
**Figure 5.3** Daubechies-4 Filter Coefficients. Low-Pass Decomposition Filter, High-Pass Decomposition Filter, Low-Pass Reconstruction Filter, High-Pass Reconstruction Filter (from left to right and top to bottom)……....44
**Figure 5.4** Example: 1-D Decomposition Using DFT Subband Decomposition. $x(n)$, $x_{LP}(n)$, $x_{HP}(n)$, $x_{LP,LP}(n)$, $x_{LP,HP}(n)$, $x_{LP,LP,LP}(n)$, $x_{LP,LP,HP}(n)$ ( from left to right and top to bottom)……...……45
**Figure 5.5** Flow Diagram of DFT Based Subband Decomposition Algorithm…47
**Figure 6.1** Synthesized Image, Decimated DFT and DWT Using: db1, db2, sym2, sym4, sym5, coif1, bior1.1 (from left to right and top to bottom)……53
**Figure 6.2** Undecimated DFT; SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, sym5, sym6, coif1, bior1.1 (from left to right and top to bottom)……....55

**Figure 6.3** Median Filtering, Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, sym5, sym6, coif1, bior1.1 (from left to right and top to bottom)……...56
**Figure 6.4** Morphological Filtering, Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, sym5, sym6, coif1, bior1.1 (from left to right and top to bottom)..…57
**Figure 6.5** Synthesized Image. Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym5 (from left to right and top to bottom)……...59
**Figure 6.6** Median Filtering, Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym5 (from left to right and top to bottom)……...60
**Figure 6.7** Decimated DFT and DWT Using: db1, sym2, sym3, sym4, sym5, sym6, bior1.1 (from left to right and top to bottom)……...62
**Figure 6.8** Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, Sym3, Sym4, Sym5, sym6 (from left to right and top to bottom)……64
**Figure 6.9** Median Filtering, Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom) ……65
**Figure 6.10** Morphological Filtering, Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom)……...66
**Figure 6.11** Real SAR Image, Oil Spills at North Sea……...67
**Figure 6.12** Decimated DFT and DWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom)……68
**Figure 6.13** Median Filtering, Decimated DFT and DWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom)……..……69
**Figure 6.14** Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom)……70
**Figure 6.15** Median Filtering, Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom)……..…….…...71

## LIST OF SYMBOLS

| | | |
|---|---|---|
| $x(n)$ | : | 1-Dimensional Signal |
| $x'(n)$ | : | Sampled 1-Dimensional Signal |
| $H_0(z)$ | : | Analysing Low-Pass Filter |
| $H_1(z)$ | : | Analysing High-Pass Filter |
| $G_0(z)$ | : | Synthesizing Low-Pass Filter |
| $G_1(z)$ | : | Synthesizing High-Pass Filter |
| $<.>$ | : | Inner Product Operator |
| $(.)^T$ | : | Transposition Operation |
| $I(i,j)$ | : | 2-Dimensional Image Signal |
| $\{\psi_i\}$ | : | Basis Vector Set |
| $\{\widetilde{\psi}_i\}$ | : | Dual Basis Vector Set |
| $Z$ | : | Set of Integer Numbers |
| $B$ | : | Basis Vector Space |
| $\{g_n(x)\}$ | : | Set of Discriminant Functions |
| $w_i$ | : | Class $i$ |
| $p(w_i)$ | : | A Priori Probability |
| $m_i$ | : | Mean of the Distribution |
| $C_i$ | : | Covariance Matrix. |
| $\Delta_f$ | : | Bandwidth |
| $\Delta_t$ | : | Duration |
| $R$ | : | Set of Real Numbers |
| $\varphi(t)$ | : | Scaling Function |
| $\psi(t)$ | : | Wavelet Function |
| $fv$ | : | Feature Vector |

## SUMMARY

## A DISCRETE FOURIER TRANSFORM BASED SUBBAND DECOMPOSITION APPROACH FOR THE SEGMENTATION OF REMOTELY SENSED IMAGES

Segmentation is the partitioning of the image into separate homogeneous regions which have common properties. Feature extraction is an important phase in segmentation. Subband decomposition is an efficient way for analysing spatial-spectral content of the signal and provides local energy distributions as features.

Recently, Discrete Wavelet Transform (DWT) approach, which can be seen as a multiresolution filter bank, has been widely used in subband decomposition. But there are two main limitations of this method when overall performance of a segmentation problem is considered. First one is the dependency on the wavelets used as basis of the transform and the spectral signature of the images those are being studied. The other limitation is the loss of the transient information in the sub-images during the decomposition. This is the poor edge localization quality of the segmentation result. Moreover down sampling yields variable filter lengths of the analysing low-pass and high-pass bands and requires efficient recalculation of the filter coefficients. The second limitation has partly lost its importance with the use of the over-complete Stationary Wavelet Transform (SWT).

We propose a Discrete Fourier Transform (DFT) based subband decomposition which uses a simple yet effective zero-phase, non-overlapping, 2-channel filter bank. These filters are ideal low-pass and high-pass filters designed in the frequency domain and their lengths are same as the length of the input image. These filters are applied to the signal in the Fourier domain.

Experiments are carried out using both undecimated and decimated versions of the DFT based method. These are compared with SWT and DWT. In case of WT based methods, different wavelets are tested.

It was shown that undecimated decompositions, although they cause some computational complexity in the image processing, always perform better than decimated decompositions. Another important result is that the DFT based subband decomposition methods are better than some wavelets. Generally DFT based subband decomposition methods provide satisfactory results for the segmentation of the remotely sensed images and may be a good candidate of the DWT which has gained high popularity in last few decades.

# ÖZET

## UZAKTAN ALGILAMA GÖRÜNTÜLERİNİN BÖLÜTLENMESİNDE AYRIK FOURIER DÖNÜŞÜMÜ KULLANAN ALT BANT AYRIŞTIRMAYA DAYALI BİR YAKLAŞIM


Bölütleme (segmentation) görüntünün ortak özellikler taşıyan homojen bölgelere ayrılmasıdır. Öznitelik çıkarma (feature extraction) bölütlemede önemli bir aşamadır. Alt bantlara ayrıştırma (subband decomposition) , bir sinyalin uzaysal-spektral içeriğinin çözümlenmesinde etkili bir yöntemdir ve bu yolla elde edilebilecek yerel enerji dağılımları öznitelik olarak kullanılabilir.

Son yıllarda alt bantlara ayrıştırmada çok çözünürlüklü (multiresolution) süzgeç bankası (filter bank) yaklaşımı olarak da görülebilecek Ayrık Dalgacık Dönüşümü (ADD) yoğun olarak kullanılmaktadır. Ama toplam bölütleme başarısı dikkate alındığında bu yöntemin iki önemli sınırlaması olduğu görülmektedir. Birincisi dönüşümde baz olarak kullanılan dalgacıklara ve çalışılan görüntünün spektral özelliklerine olan bağımlılıktır. İkincisi, alt bantlardaki uzamsal değişim bilgilerinin alt örneklemeden dolayı kaybolmasıdır. Alt örnekleme bölütlemede kenar kestirim kalitesinin düşmesine neden olmaktadır. Daha da önemlisi alt örneklemeden dolayı her ayrıştırma aşamasında sinyalin boyutu değişmekte ve dolayısıyla alçak-geçiren ve yüksek-geçiren filtrelerin boyutları değişmektedir ki bu da filtre katsayılarının hızlı ve verimli bir biçimde yeniden hesaplanmasını gerektirir bu da tekrarlanması gereken hesaplamalar ve dolayısıyle işlem yükü getirmektedir. İkinci kısıtlama Durağan Dalgacık Dönüşümü (SWT) ile bir ölçüde aşılabilmektedir.

Bu çalışmada Ayrık Fourier Dönüşümüne (AFD) dayalı basit ama etkili, sıfır-fazlı, spektrumu örtüşmeyen, iki-kanallı bir süzgeç bankası kullanan alt bant ayrıştırma yöntemi önerilmiştir. Boyutları her zaman giriş sinyalinin boyutuyla aynı tutulan bu filtreler frekans bölgesinde tasarlanmış ideal filtrelerdir. Filtreleme işlemi frekans bölgesinde gerçekleştirilmektedir.

Uygulamalarda AFD'ye dayalı yöntemin hem alt-örneklenmiş hem de örneklenmemiş versiyonları denenmiştir. Bunlar, Ayrık Dalgacık Dönüşümü ve Durağan Dalgacık Dönüşümü ile karşılaştırılmıştır. Dalgacık Dönüşümünde farklı dalgacıklar kullanılarak deneyler gerçekleştirilmiştir.

Alt örneklenmemiş ayrıştırmalar, görüntü işlemede hesaplama yükü getiriyor olsalar da, her zaman, alt örneklenmiş ayrıştırmalardan daha iyi sonuç vermişlerdir. Bir başka önemli bir sonuç da AFD'ye dayalı yöntemin ADD'nin kullandığı db6, sym6 gibi bazı dalgacıklardan daha iyi sonuç verdiğidir. Genel olarak AFD'ye dayalı yöntemlerin bölütleme için yeterli sonuçlar verdiği gözlenmiş ve son yıllarda uzaktan algılama görüntülerinin bölütlenmesinde popülerlik kazanan ADD'ye bir alternatif olabileceği görülmüştür.

# 1. INTRODUCTION

Image can be defined as a two-dimensional signal in scientific and technical context. It should be mentioned that, it has different contextual meanings in distinct disciplines such as optics, mathematics, computer science, finance, arts, philosophy, theology, and social psychology. An image of a physical object is taken with the help of natural or artificial sensors such as human eye, camera, radar etc. The sensors are located at a distance apart from the object. The distance may vary from nano, micro to macro scales but it is right to say that all images are remotely sensed.

Segmentation of an image is the logical reordering and representation of the image. Through segmentation the image is partitioned into separate pieces of regions which have some common properties.

Texture can be defined through the region in an image, in which local statistics or other local properties are constant or slowly varying [1, 2]. Segmentation of the texture content in digital images has received considerable attention during the past decades and numerous approaches have been presented [3, 4, 5]. Segmentation of images significantly reduces the amount of data. Moreover it is a very important phase before edge detection, classification and compression of images. More obvious edges and separate segments can be observed after segmentation process. It may be regarded as a pre or post processing operation before and after any other image processing and representation. It is expected that this new representation of the image which is more suitable for human and machine perception will provide some advantages for further image processing and both for the experts and the end users of the images.

Remote Sensing images are especially appropriate for characterization by textures. SAR images are single band images which contain MR textures due to the non-

stationarity of the microwave reflection of the ground and biomedical images also have textural properties. In last few decades the use of remote sensing is very much emphasized, because remotely sensed images are a valuable source of spatial data for a variety of applications such as urban mapping, natural resources management and environmental monitoring and for classification of geology, water temperature, soil characteristics, soil moisture, water pollution, flood damage estimation, groundwater location, vegetative diseases. RADAR sensors provide all-time and all-weather surveying tool making them ideal candidates for land cover mapping. Moreover they have been playing an important role in the remote sensing of environmental disasters.

The fundamental assumption for most filtering approaches is based on the statements that the energy distribution identifies the texture and the local energy contained in a band can be used as a feature to discriminate the texture. Utilizing these facts, several filter bank approaches have been proposed for subband decomposition and by these subbands, it has become possible to reach local energy distributions of signals. Subband decomposition can be applied by various transformation techniques. We believe that the most important element is the filter or the filter bank used to decompose the signal. At this point we ask how can a filter or filter bank be efficiently designed that will give the best approximation of the signal.

In particular we will focus on the efficient decomposition of the signal using simple low-pass and high-pass filtering in the frequency domain and decimation which is carried out in spatial domain. Discrete Fourier Transform (DFT)/Real Discrete Fourier Transform (RDFT) based subband decomposition and fast algorithms were studied by O. K. Ersoy [6, 7, 8] for the first time. Later, the DFT based subband decomposition method successfully applied to image fusion [6] and speckle reduction [7]. In [6, 7], image fusion and speckle reduction using DFT based subband decomposition were also compared to WT based decomposition methods. In this thesis we will apply this method to the segmentation problem and compare the results to the Wavelet Transform (WT) based methods. The design of such efficient filters is a challenging task to accomplish [8]. Although subband decomposition has been widely used in segmentation applications [9], critically sub sampled or full rate DFT based subband decompositions have not been widely applied to image segmentation problem and have not been compared to critically sampled Discrete Wavelet Transform (DWT) and redundant Stationary Wavelet Transform (SWT)

based subband decomposition methods which use different mother wavelets. Comparative studies tend to be unsatisfactory and don't appear in the literature. Only a few comparative studies have been conducted so far, to justify their effectiveness [10].

It was shown that although decimation is very important in some signal processing applications such as signal coding, it is not useful in many other applications such as segmentation, classification, noise removal and object recognition. Therefore, we question the role of the multiresolution analysis which is commonly played in subband decomposition and especially that of DWT in image segmentation.

A signal can be represented as a weighted sum of certain base functions. WT is computed by the expansion of the signal into a family of functions which are dilations and translations of a unique function and completely characterizes the signal and represents it in an optimum way. WT has been attracting attention in diverse areas such as signal processing and image processing.

DWT which is a useful technique for time-frequency analysis is often used in segmentation applications. The DWT can also be seen as a filter bank which consists of low-pass and high-pass filters. DWT has two distinct features. First it is maximally decimated and second it provides perfect reconstruction condition. DWT corresponds to Multiresolution Analysis (MRA) approximation expressions. This method permits the analysis of the signal in many frequency bands at many scales. In practice, MRA is carried out using 2-Channel filter banks composed of a low-pass and a high-pass filter and each bank is then sampled at a half rate. Mallat applied a critically decimated DWT with a dyadic subband structure [11]. The wavelet coefficients are sampled based on Nyquist criteria, this is critical sampling rate which is equal to a decimation factor that is equal to the number of channels of the filter bank. This representation is accordingly non-redundant and the total number of samples is equal to the size of the original image.

DWT is an effective tool for multiresolution texture analysis. However, down sampling during decomposition does not properly characterize the shift invariance properties of the texture and therefore critically sampled filter banks typically imply inaccurate texture edge localization. This major inconvenience of this representation, that it does not conserve the invariance by translation, is reported in [12, 13].

Discrete Wavelet Frame (DWF), an approach that does not down sample the signal during decomposition, may be used to avoid these circumstances [14]. Thus multiresolution analysis is not desirable for segmentation and estimation/detection problems. In order to preserve the rapid changes, the down sampling operation must be suppressed. SWT preserves the edges and many other textural properties in a SAR image.

The main drawback is the complexity in design of these filters and the dependency of the performance on the type of the filter and the characteristic of the application and more specifically the spectral signature of the image at hand. Moreover, it is reported that the asymmetry of the impulse response of these filters causes bad localization of edges during segmentation [10].

Another challenging issue is the use of a very broad class of filters, namely Quadrature Mirror Filters (QMF) incorporating both Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) filters in subband decomposition. A work on designing IIR and 8-tap, 16-tap, 32-tap FIR filters can be found in [15]. It is natural to question whether the performance of subband decomposition could be enhanced by developing a simpler QMF. We use ideal Low-Pass (LP) and High-Pass (HP) FIR filters, lengths of which are adaptively adjusted to the length of the signal in the frequency domain [6, 7]. The reason why these structures are chosen is the expectation of less complexity and simple design structure and their robustness.

The proposed maximally decimated DFT based filter bank is further extended to an undecimated one. So the subbands have the same size as the original image. And the length of the ideal LP and HP filters designed in the frequency domain and applied to the signal in the frequency domain will be fixed. Like with the wavelet filters, these filters will be tested both at critical sampling rate and full rate.

The work in this thesis, following comparative study in [10], is trying to demonstrate that for certain tasks such as image segmentation, classification, edge detection WT based methods are not necessary but sufficient tools.

In this thesis, contrary to [10], the filters used in DFT based method are ideal, their lengths are not fixed to a constant, wavelets used in WT based methods are different. The classifier, texture sets used and the overall system setup are also different.

We explore whether it is possible to design simple yet effective filters and use them with well-known DFT to be able to appropriately decompose signals, extract features from them and replace this newly developed technique with commonly used DWT.

The work is organized as follows:

In section 2 we give a brief description of a general segmentation scheme and detail its phases. The energy content of the subbands is expected to characterize the signal best and is used in feature vector. There are many classification schemes described in the literature we will use nearest mean classifier which uses Euclidian Distance as a similarity metric. Section 3 introduces the concept of subband decomposition using QMF. 2-Channel QMF which will be used in this thesis are explained. Decimation and interpolation are described here and then conditions and criteria that should be met for perfect reconstruction of the signal are also briefly given. Orthogonal Linear Transforms used in this thesis are detailed in Section 4. First we briefly give general concept of linear expansion of a signal and then the fundamental frequency analysis tool FT and a modification of FT the STFT is introduced as a time-frequency analysis method. Finally WT methods are explained in detail. In section 5 we present our the proposed algorithm for subband decomposition [6, 7], feature extraction, image segmentation, and give other components of the full system. In section 6 numerical experiments are carried out and the results are shown in figures along with accuracy rates listed in tables to demonstrate the potential of the new DFT based method. We accomplish our task in four experiments as follows, first we simulate a 2-class texture classification problem and try to segment the regions in a synthetic texture. Second we apply same methods to another artificial image which is composed of four different textures that are aggregated to form a new image. In third experiment, a problem which can be regarded as target detection, is simulated. In this image four objects that belong to the same texture are replaced on a uniform background. Finally, the experiments are carried out on a real SAR image of an oil spillage accident which is a remotely sensed monochrome image with gray-levels in the range {0–255}. In each of the experiments the overall system setup is kept same except for changing the transformation method and wavelets in WT.

## 2. TEXTURE FEATURE EXTRACTION AND SEGMENTATION USING SUBBAND DECOMPOSITION

Many images contain regions characterized not so much by a unique value of brightness, but by a variation in brightness that is often called texture. Remote Sensing images are especially appropriate for characterization by textures. The addition of texture measures to spectral features has been widely studied to enhance classification accuracy. Textures can be used as the basis for discriminating various structural regions [16]. Many definitions for the term texture can be found in the literature, texture can be defined as a region in an image if a set of local statistics or other local properties of that region are constant or slowly varying [1, 2].

We will apply Discrete Fourier transform (DFT) and Wavelet transform (WT) to decompose image to be able to extract useful information from spatial distribution of subband images. Several studies have shown that it is possible to enhance the segmentation results by combining the textural and the spectral information [17]. In this study, the technique used takes the spatial as well as the spectral information into account. A general model for such a system is illustrated in Figure 2.1. Before applying the new method to the real SAR image, we will run it on textures. Textures used are selected from Brodatz album and are arranged manually on a computer using MATLAB platform.

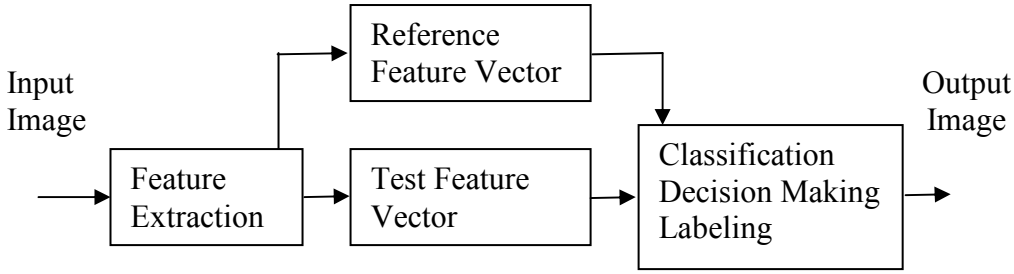


**Figure 2.1**: General Model of a Feature Extraction and Segmentation System

## 2.1 Feature Extraction Methods

Extraction and selection of features is critical in detection, recognition, identification, segmentation and classification problems. Prior to segmentation and classification stages it is necessary to extract features in order to be able to appropriately characterize the classes, the logical groups. In Figure 2.1, a general model for feature extraction and segmentation is illustrated. Tuceryan and Jain [18] identify five major categories of features for texture identification; statistical, geometrical, structural, model-based, and signal processing features. We use transform based methods at feature extraction phase.

### 2.1.1 Statistical Methods

From the second half of 70s until the mid 80s statistical methods were very popular. Statistical methods more generally used first order and second order statistics such as mean value, variance, energy, entropy and angular second moment of the spatial distribution of the gray level intensities of pixels in an image. The elements of co-occurrence matrices are quantitative measures of the occurrence of gray levels in an image. Co-occurrence matrices were run on the spatial images. Later other methods, such as transform based methods also utilized co-occurrence matrice**s** [19].

### 2.1.2 Model Based Methods

Model based methods assume a signal as a synthesized process at the output of a model. The model for example, can take white noise as the input, and parameters of the model are used as the features to characterize the image. The models proposed are Autoregressive (AR) models; Lattice Filter (LF) based models etc. In the AR model case the coefficients of the AR filter are taken, while in LF model the reflection coefficients are used in feature vector. These methods are also called signal processing methods in the signal processing society [9, 20].

### 2.1.3 Transform Based Methods

The transform based methods first transform the image into a different domain in which it is expected to obtain more satisfactory information about the signal. Fourier transform for example can successfully locate frequencies of sinusoids in the

frequency domain where it is difficult to estimate the frequencies of sinusoids contained in a signal in the original time or spatial domain. FT, Discrete Sine Transform (DST), Discrete Cosine Transform (DCT), WT are most commonly used transformations. The data in the transform domain can then be manipulated. For example, the mean, variance and energy of the signal may be used in the feature vector.

## 2.2 Subband Decomposition and Filter Bank Approach

Subband coding of speech was introduced by Crochiere [21], and this technique was extended to multidimensional case [9, 22], and found applications in image and video processing [23]. In this method, time domain or spatial domain signal, is applied to the input of filter bank to produce low-pass and high-pass signals. Specifically if the filter bank is composed of just two filters, the filters are high-pass and low-pass. These complementary filters together cover the whole range of signal spectrum. If the reconstruction of the signal is among the tasks, then a synthesis stage follows. Image fusion, for example, requires reconstruction processes. In this thesis, we do not have to reconstruct the signal; however we have detailed the conditions and criteria that must be met in order to reconstruct the signal perfectly from its subbands. Once the analyzing low-pass filter defined, the analyzing high-pass and synthesizing filters, which are mirror-image of their analyzing stage counterparts, can be easily determined.

The wavelet transform is a new technique for decomposing signals and found many applications in texture analysis [24]. Short Time Fourier Transform (STFT), unlike FT which gives no information about time of occurrence of frequency components, by mapping 1-D time signal to 2-D time-frequency domain, successfully locates the time of changes in the frequency domain. WT can be seen as an extension of STFT with better time-frequency localization. WT produces approximation and detail coefficients of a signal. Approximations are less changing parts while details are rapidly changing parts. In other words wavelet transform resembles high-pass and low-pass filtering, since low frequency components are slowly changing parts and high frequency components are rapidly changing parts of a signal. In DWT to obtain and analyze the signal at different levels and resolutions, filtering techniques are used. The resolution of the signal can be changed by sub-sampling or decimation. To

obtain the approximation, a scaling function which can be seen as a LP filter is used and half of the frequencies, those are higher than the mid-frequency point, are eliminated, this implies the reduction of the information and the resolution. Therefore half of the samples are redundant and need not be contained as far as reconstruction is concerned. The same is true for the LP part of the signal, therefore the total number of samples is kept same as the original signal and the redundancy is avoided. 1-Dimensional case is extended to 2-Dimensional case by first applying the 1-D filters to the rows and second to the columns of the image. Two approaches exist for 2-D case, separable and non-separable 2-D filters. A separable filter is the multiplication of two 1-D filters and most of the previous work is based on separable filtering, as they will be used in this work [25]. Non-separable filters are 2-D filters that are directly processed on image and have directionality property, which is not limited to vertical and horizontal directions of separable filters [9, 26].

### 2.2.1 Pyramid Structured Decomposition

The pyramid structured decomposition recursively decomposes sub-signals in the low frequency channel. In this approach it is assumed that the most of the energy of the signal is concentrated in the low end of the frequency spectrum, emitting the high frequency information at second and third stages of decomposition [11]. In this work we will neglect high frequency components in further decomposition steps, however other decomposition schemes that are provided for selecting best subbands are found in the literature. Depending on the type of the signal to be processed among all bands some may contain majority of the information and only essential bands should be further processed at each stage. Furthermore signal-adaptive concepts have been developed for the selection of appropriate branches of the tree and the selection is made on some predefined rules. Tree structure is an alternative to pyramid structure. Although it will not be used in this thesis, at this point it is important to mention the study of Chang and Kuo who insist that for some signals the most important features are contained in intermediate frequencies, thus that the octave band decomposition is not optimal for such signals [11, 27]. They propose a decomposition scheme which explores the branches having the higher energy which possibly contain the maximum information. At each step, the energies of all subbands are calculated and compared with a threshold to see whether it passes this level, if any band falls under this level

then decomposition on that branch is stopped. In Figure 2.2 pyramid structure decomposition is shown.

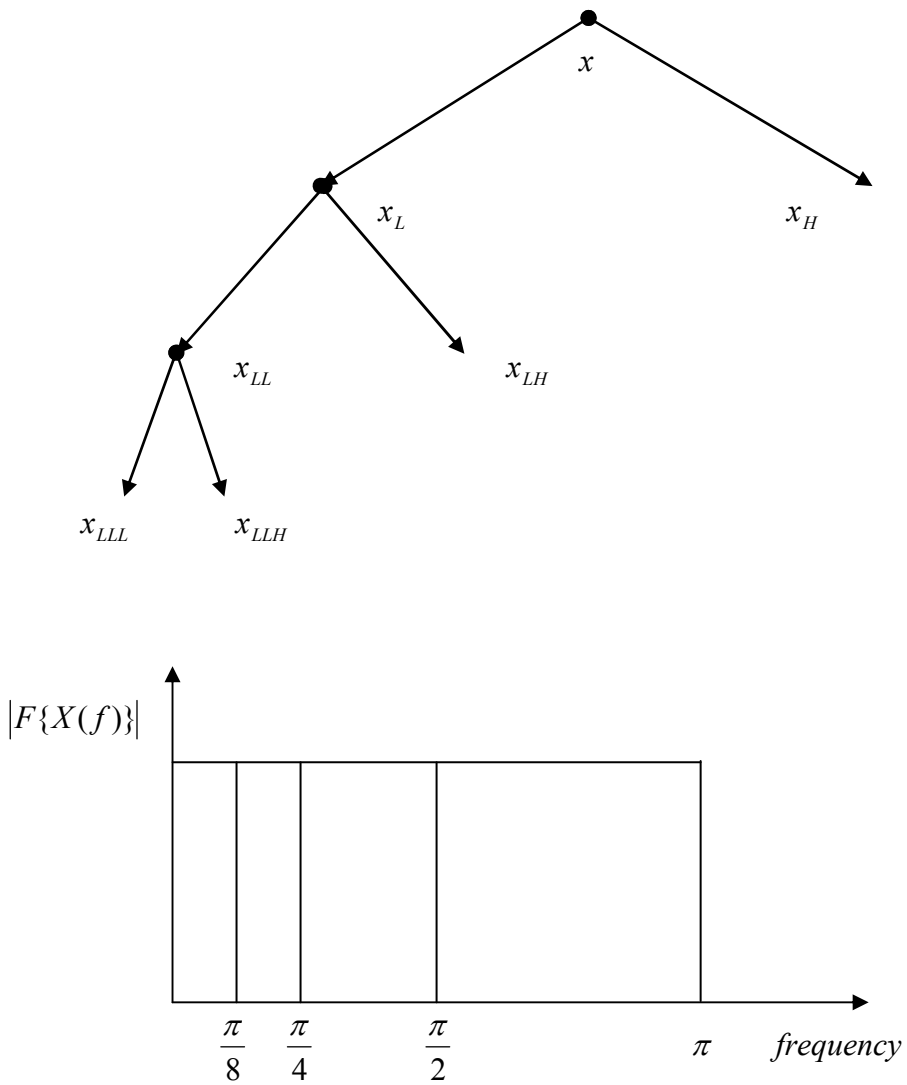


**Figure 2.2:** Pyramid Structured Decomposition and Division of the Spectrum of Signal

## 2.3 Feature Extraction and Selection

In this thesis the spatial-spectral features are used. The image is a 2-D gray-level intensity matrix indicating the spatial distribution of objects on a background. The decomposition scheme, whether be wavelet or DFT based, transforms the image into a different domain which is mainly characterized by the spectral changes, and divides the spectrum in different parts depending on some predefined rules, and back

transforms the useful ones to their original spatial domain. This co-existence of the operations provides spatial-spectral features.

1-Level decomposition produces 4 subbands, an approximation component, and three detail components. Three detail components are horizontal, vertical and diagonal details. The l1-norms of each of four bands are computed and the feature vector is filled in by these entries. Energies of subbands will be used as features. Thus the feature vector is composed of only 4 entries. The definition of l1-norm is given as:

$$e_I = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left|I(i,j)\right| \tag{1.1}$$

After 1-Level decomposition, to further explore the spectrum one needs more decomposition levels. In 2-Level decomposition the approximation component of the first level is decomposed and detail coefficients are remained untouched. In 2-level decomposition the l1-norms of detail coefficients of first stage and l1-norms of the four new subbands produced from the approximation of the first stage are computed and stored. Then the feature vector of 2-Level decomposition has 7 elements.

Similarly in 3-level decomposition we used the approximation component of the second stage and the feature vector is composed of l1-norms of 10 subbands. The feature vectors are composed as follows:

$$fv_{1,level} = [e_{a1}, e_{h1}, e_{v1}, e_{d1}] \tag{1.2}$$

$$fv_{2,level} = [e_{a2}, e_{h2}, e_{v2}, e_{d2}, e_{h1}, e_{v1}, e_{d1}] \tag{1.3}$$

$$fv_{3,level} = [e_{a3}, e_{h3}, e_{v3}, e_{d3}, e_{h2}, e_{v2}, e_{d2}, e_{h1}, e_{v1}, e_{d1}] \tag{1.4}$$

### 2.3.1 Normalizing Feature Vector

Normalization of the feature vector is needed to get more accurate results. By normalizing the feature vectors, the feature space is projected to a bounded space whose elements can take on values between 0 and 1.

$$fv_{\max} = \max\{fv\} \tag{1.5}$$

$$fv_{\min} = \min\{fv\} \tag{1.6}$$

$$fv_{normailzed} = \frac{fv - fv_{\min}}{fv_{\max} - fv_{\min}} \tag{1.7}$$

### 2.3.2 Calculating the Mean of the Feature Vectors

In order to test a pixel of the image to specify its class, a comparison with reference features is needed. These references, as explained above, are acquired from the training stage. The number of training samples is more than one, therefore mean of these values have to be computed as follows,

$$fv_{class_i,mean} = \frac{[fv_{class_i,sample_1} + fv_{class_i,sample_2} + \ldots + fv_{class_i,sample_N}]}{N} \tag{1.8}$$

The mean values for all classes are evaluated in the same manner as above. A typical two dimensional feature space of two classes is shown in Figure 2.3

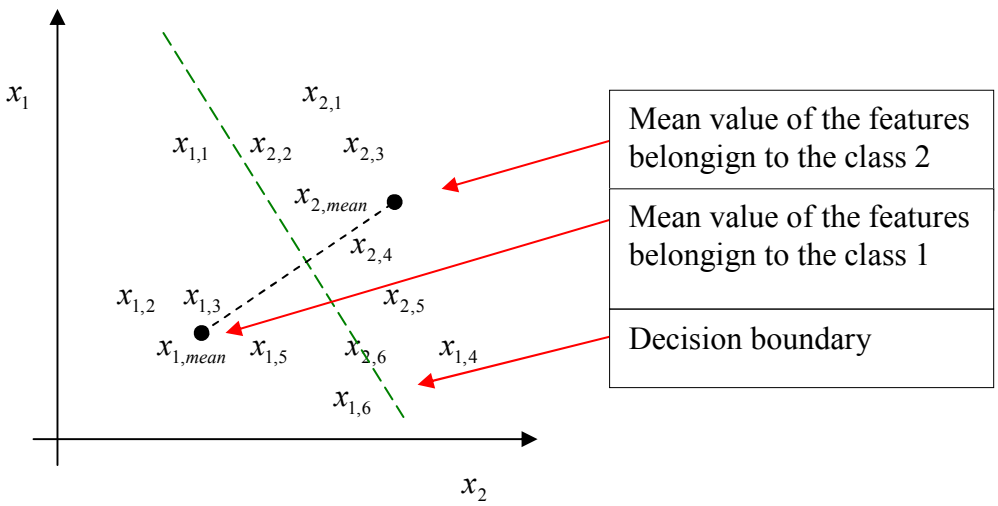


**Figure 2.3:** An Example of a Feature Space for a 2- Class Segmentation Case, Each Feature Vector Contains 2 Elements.

## 2.4 Segmentation

Segmentation is the process of determining the borders of homogeneous regions in an image. The most important task is to isolate regions that have common properties and distinguish them from the neighboring regions. Since the borders or edges are important for discriminating homogeneous regions, edge detection algorithms are widely used in segmentation procedures. And local statistics are also used to discriminate the borders of homogeneous sections. Segmentation is the partitioning the image data into a set of disjoint regions with uniform property. Segmentation segments an image into disjoined regions corresponding to objects, or parts of objects that differ from their surroundings, and thus enables further classification to be performed based on the information provided by the clusters rather than individual pixels [16]. Generally classification process comes after the segmentation.

## 2.5 Classification

Classification is one of the most widely used information extraction techniques in remote sensing. The objective of classification is to assign all pixels in the image to particular classes. It turns the remote sensing data into meaningful categories representing surface conditions or classes. The resulting classified image is called thematic map. A thematic map is an informational representation of an image, which shows the spatial distribution of particular classes. Thematic maps are often used as input for information systems such as Geographic Information System (GIS) [16].

Classification is the labeling of the sections. The labeling is logical naming of the heterogeneous regions. Decision making on the grouping of regions is a process accomplished manually or automatically. The manual handling of the selection of classes or at least determining the number of classes is called supervised classification. The analyst identifies pixels of known cover types, and then a computer algorithm is used to group all other pixels into one of these groups. Unsupervised classification on the other hand uses no prior information about the classes, since there might exist hidden classes in the image that are not expected or previously observed even by the experts. For that case, class labels are unknown and

pattern classes must be inferred from available data via unsupervised learning algorithms. In this thesis we will use supervised classification methods.

### 2.5.1 Classifiers and Decision Rules

There are many classification methods used in classification problems. The objective of the classification is to find the decision rules, which partition the feature space into the volumes called decision regions, each one corresponding to a given class. If a feature vector is located in a particular region it will be assigned to the class associated with that region. The decision regions are separated by surfaces called the decision boundaries [28]. So, partitioning the feature space implies establishing decision boundaries. Among a set of discriminant functions $\{g_1(x), g_2(x), ..., g_m(x)\}$ the one having lager value than others determines the class $i$ which $x$ belongs to [29].

The quantitative distance measure used in this thesis is the Euclidian distance. Nearest Mean classifier, using Euclidian distance, decides which class an individual pixel belongs to. After this step decision making, labeling is done.

Bayes decision theory is a fundamental statistical approach for classification. The Bayes classification problem can be seen as an optimization problem in which one desires to construct a classifier, which minimizes the average probability of error called the error rate, or a loss function called the overall risk [28].

When the distribution of features in feature space is Gaussian or normal, the distribution function becomes [16]

$$g_k(x) = \ln p(w_i) - \frac{1}{2}\ln|C_i| - \frac{1}{2}\left[(x - m_i)^T C_i^{-1}(x - m_i)\right] \tag{1.9}$$

Where $w_i$ indicates the class $i$, $x$ is the feature vector , $p(w_i)$ is a priori probability defined as the probability that a random observed feature vector $x$ belongs to class $w_i$, $m_i$ is the mean of the distribution and $C_i$ is the covariance matrix.

If one takes the classes as equiprobable with a common covariance C for simplification, the discriminant function will reduce to [29]

$$g_k(x) = (x - m_i)^T C^{-1}(x - m_i) \tag{1.10}$$

If further one makes the assumption that the classes are equiprobable with the same covariance matrix that is equal to the identity matrix, the discriminant function, which is Minimum Distance to Means Classifier, reduces to

$$g_k(x) = (x - m_i)^T (x - m_i) \tag{1.11}$$

In this thesis we will use this classifier for its simplicity, and its contributions to increase the efficiency of the algorithms by reducing the process time. The decision boundary for this classifier is linear and in fact it is the perpendicular bisector of the line between the two class values. Its location does not depend on the distributions of the classes [29**]**.

Euclidian distance measure is used to decide which class an individual test pixel belongs to. As shown in the Figure 2.4, a line which is perpendicular to the line interconnecting the two means is the boundary between classes. A test pixel is labeled as the member of a class if it is nearer to the mean of that class than the mean of the other class or classes, whether it actually be or not be member of that class.

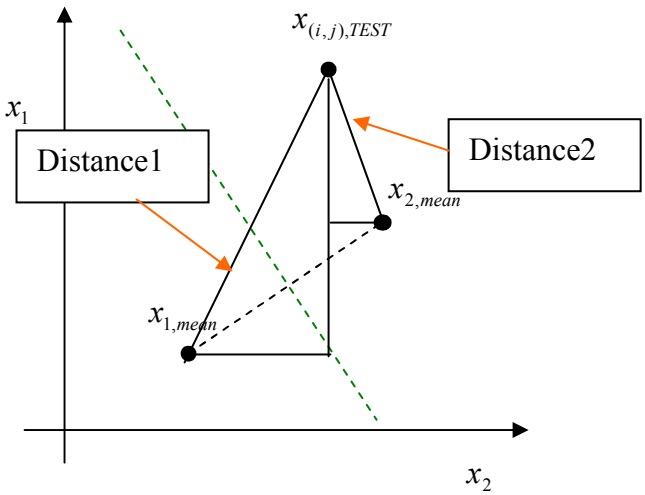


**Figure 2.4:** Decisions for a Test Pixel in the Segmented Feature Space by Measuring Euclidian Distances to the Means

## 2.6 Post-processing Operations

Preprocessing operations such as contrast stretching, histogram equalization, gray level slicing, bit-plane slicing can increase the segmentation and classification accuracy when applied before these operations. In order to get better results we apply few post-processing operations to the image to increase the segmentation and classification accuracy. These operations may include mean and median filtering and mathematical morphology.

Mean Filters are linear averaging filters. Spatial mean filtering is a convolution operation applied to a pixel group to obtain a new value for the center pixel. This blurs the image and corresponds to low-pass filtering. Uniform random noise is removed via this type of filtering but mean filtering is not practical when binary noise is present in the image. To remove binary noise and speckle noise a non-linear type averaging filter, median filter must be used. It is good at removing extreme values whereas preserving edges. The pixels are rearranged in ascending order and the value of the pixel that is in the middle of the mask is replaced with the value of the one in the middle of the rearranged order [30].

Mathematical morphology includes filling small holes in objects, separating adjacent or slightly overlapping objects, and joining broken boundaries into continuous segments, morphological operations have been successfully used in remote sensing applications [31].

Structuring elements such as disk, ball, square etc. are used with the four most basic operations: dilation, erosion, opening and closing. Opening and closing are combinations of dilation and erosion [32].

# 3. SUBBAND DECOMPOSITION AND FILTER BANKS

## 3.1 1-Dimensional Subband Decomposition

In this section 1-Dimensional 2-Channel filter banks and the perfect reconstruction conditions for these will be reviewed. The derivations used in this section heavily depend on the previous work by Vaidyanathan. The generalization to M-band case is also studied in [26, 33, 34]. The equations derived in sub sections can be found in [9, 26, 33, 34].

### 3.1.1 Quadrature Mirror Filters (QMF)

Let $h_0(n)$ be an FIR filter with real coefficients. The filter $h_1(n)$ is defined as

$$h_1(n) = (-1)^n h_0(n) \tag{3.1}$$

And its frequency domain equivalent is

$$\mathrm{H}_1(\mathrm{e}^{\mathrm{jw}}) = \mathrm{H}_0(\mathrm{e}^{\mathrm{j(w-\pi)}}) \tag{3.2}$$

Substituting $\omega$ with $\frac{\pi}{2} - \omega$ and noting that the magnitude is an even function of $\omega$, the following relation is obtained,

$$\left|H_1(e^{j(\pi/2-\omega)})\right| = \left|H_0(e^{j(\pi/2+\omega)})\right| \tag{3.3}$$

From above it can be seen that $H_0(z)$ and $H_1(z)$ have the mirror image property about the point $\omega = \frac{\pi}{2}$. Therefore they are called Quadrature Mirror Filters (QMF) and are used to eliminate the aliasing effect in two channel subband decomposition.

### 3.1.2 Decimation and Interpolation in 1-Dimension

Decimation is the process of reducing the sampling rate of a sequence. It is also called down sampling. The full-band signal is passed through an antialiasing filter

and an intermediate signal $x'(n)$ is obtained then sub sampled. The down sampling process by $M$ is shown below. In Figure 3.1 decimation and interpolation for the factor 2 are shown.

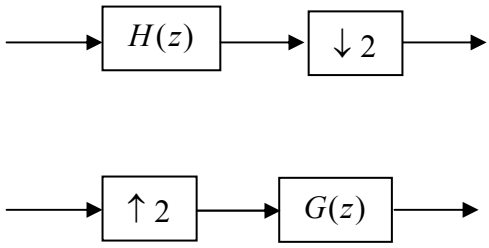


**Figure 3.1:** Decimation and Interpolation by the Factor 2

$$x'(n) = \begin{cases} x(n) & n = 0, \pm M, \pm 2M, ... \\ 0 & otherwise \end{cases} \tag{3.4}$$

$$y(n) = x'(Mn) = x(Mn) \tag{3.5}$$

The intermediate signal $x'(n)$ can be expressed in terms of the impulse train

$$x'(n) = \left[ \sum_{r=-\infty}^{\infty} \delta(n - rM) \right] x(n) \tag{3.6}$$

In Discrete Fourier Series, the expansion of $x'(n)$ is as

$$x'(n) = \frac{1}{M} \sum_{k=0}^{M-1} x(n) e^{j2\pi nk/M} \tag{3.7}$$

And the Discrete Fourier Transform of the intermediate signal is

$$X'(z) = \frac{1}{M} \sum_{k=0}^{M-1} X(z W_M{}^k) \tag{3.8}$$

Where $W_M = e^{-j2\pi/M}$. The frequency response of the $x'(n)$, is obtained by replacing $z$ with $e^{jw}$

$$X'(e^{jw}) = \frac{1}{M} \sum_{k=0}^{M-1} X(e^{j(w-2\pi k)/M}) \tag{3.9}$$

Equation (3.9) shows that the Discrete Fourier Transform of the intermediate signal $x'(n)$ is the sum of $M$ replicas of the frequency response of the original signal $x(n)$ spaced at $\frac{2\pi}{M}$.

After selecting the values at the critical sampling rate, zeros between samples are eliminated and therefore time scale is compressed by the factor $M$.

$$Y(z) = \sum_{n=-\infty}^{\infty} x'(Mn)z^{-n} = \sum_{k=-\infty}^{\infty} x'(k)z^{-k/M} \tag{3.10}$$

$Y(z)$ can be written in terms of $X'(z)$

$$\mathrm{Y(z) = X'(z^{1/M})} \tag{3.11}$$

And in terms of $X(z)$

$$Y(z) = \frac{1}{M}\sum_{k=0}^{M-1} X(z^{1/M}W^{k}) \tag{3.12}$$

On the unit circle the Frequency response is obtained by putting $z = e^{jw}$

$$Y(e^{jw}) = \frac{1}{M}\sum_{k=0}^{M-1} X(e^{j(w-2\pi k)/M}) \tag{3.13}$$

Compression in time results in an expansion in Frequency domain. Without proper selection of the sampling filter, down sampling may also lead to a loss of information during reconstruction.

Interpolation is the process of increasing the sampling rate of a signal by an integer factor $M$. This process is achieved by the combination of up sampler and low pass filter. This operation is defined by

$$y(n) = \begin{cases} x(n/M) & n = 0, \pm M, \pm 2M, \ldots \\ 0 & otherwise \end{cases} \tag{3.14}$$

This operation inserts $M-1$ zeros between samples.

$$Y(z) = \sum_{n=-\infty}^{\infty} y(n)z^{-n} = \sum_{n=-\infty}^{\infty} x(n/M)z^{-n} = \sum_{k=-\infty}^{\infty} x(n)z^{-Mk} \tag{3.15}$$

Up sampling in terms of $X(z)$ can be written as,

$$Y(z) = X(zM) \tag{3.16}$$

Two effects of interpolation are compression in frequency domain and generation of the additional high frequency signals which is a consequence of inserting zeros between samples which causes additional high frequency components [9, 26].

## 3.2 1-Dimensional 2-Channel Filter Banks

In the Figure 3.2, a QMFB which is composed of the analysis and synthesis parts is shown. The analysis part decomposes the signal into separate bands and down samples by the factor 2. At the synthesis stage, the two subbands are up sampled, filtered and summed up to reconstruct the input signal. Here in the Figure 3.2 the filters $H_0(z)$ and $H_1(z)$ at the analysis stage are low-pass and high-pass filters respectively. These are anti-aliasing filters used to split the signal into two equal subbands. The anti-aliasing filtered signals $v_0$ and $v_1$ are down sampled by two to give the outputs $p_0$ and $p_1$ of the analyzing filter. At the synthesis stage these signals are up sampled by two by inserting zeros between samples to give out signals $q_0$ and $q_1$ where $q_0$ and $q_1$ are filtered through interpolation filters $G_0(z)$ and $G_1(z)$ respectively. The outputs of the interpolating filters are summed up to give a perfect or approximate of the original signal $x$.

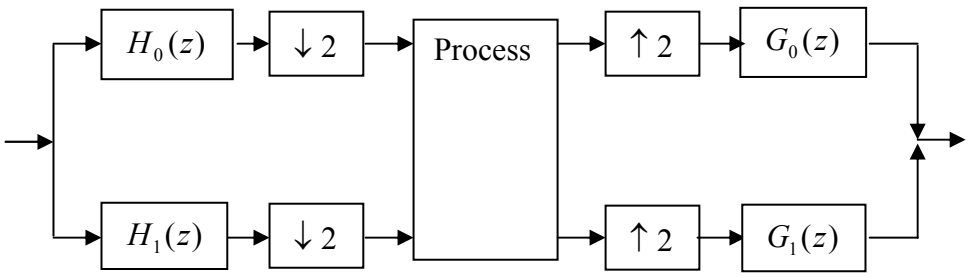


**Figure 3.2:** Analysis and Synthesis Parts of a 2-Channel QMFB

The equations for decimation and interpolation processes in the z-domain are given in equations (3.17) and (3.18). From decimation equation:

$$P_0(z) = \frac{1}{2}(V_0(z^{1/2}) + V_0(-z^{1/2})) \tag{3.17}$$

$$P_1(z) = \frac{1}{2}(V_1(z^{1/2}) + V_1(-z^{1/2})) \tag{3.18}$$

and from the interpolation equation:

$$Q_0(z) = P_0(z^2) = \frac{1}{2}(V_0(z) + V_0(-z)) \tag{3.19}$$

$$Q_1(z) = P_1(z^2) = \frac{1}{2}(V_1(z) + V_1(-z)) \tag{3.20}$$

The output of the synthesis filter can be obtained as in the following equations,

$$Y_0(z) = Q_0(z)G_0(z) = \frac{1}{2}(V_0(z)G_0(z) + V_0(-z)G_0(z)) \tag{3.21}$$

$$Y_1(z) = Q_1(z)G_1(z) = \frac{1}{2}(V_1(z)G_1(z) + V_1(-z)G_1(z)) \tag{3.22}$$

By expressing the intermediate signals $V_0(z)$ and $V_1(z)$ in terms of the input signal and the analyzing filters $X_0(z)H_0(z)$ and $X_1(z)H_1(z)$ respectively, one can rewrite the above relationships as

$$Y_0(z) = \frac{1}{2}(X(z)H_0(z)G_0(z) + X(-z)H_0(-z)G_0(z)) \tag{3.23}$$

$$Y_1(z) = \frac{1}{2}(X(z)H_1(z)G_1(z) + X(-z)H_1(-z)G_1(z)) \tag{3.24}$$

These two signals are used to reconstruct the signal $\hat{X}(z)$

$$\hat{X}(z) = \frac{1}{2}\left[H_0(z)G_0(z) + H_1(z)G_1(z)\right]X(z)$$
$$+\frac{1}{2}\left[H_0(-z)G_0(z) + H_1(-z)G_1(z)\right]X(-z) \tag{3.25}$$

The right side of the equation is composed of two parts. First part describes the transmission of the signal $X(z)$ through the system and the second describes the aliasing component of the output of the filter bank.

$$\hat{X}(z) = T(z)X(z) + S(z)X(-z) \tag{3.26}$$

### 3.3 Perfect Reconstruction Conditions for 2-Channel Filter Banks

Perfect reconstruction of the signal requires two conditions; firstly, transmission component of the signal must be delayed and constant $c$ times of the input signal and secondly no aliasing effect should remain. These are satisfied by the following two equations respectively,

1- $T(z) = cz^{-n_0}$ Where c is a constant

2- $S(z) = 0$ for all z

If these two conditions satisfied then the output signal will be of the form

$$\hat{x}(n) = cx(n - n_0) \tag{3.27}$$

Perfect reconstruction for a 2-channel QMFB is achieved via canceling the aliasing effect completely, which requires a simple design of the synthesis filters according to the task as follows,

$$G_0(z) = -H_1(-z) \tag{3.28}$$

$$G_1(z) = H_0(-z) \tag{3.29}$$

For this choice of filters the aliasing effect is cancelled but non-constant amplitude and the non-linear phase of the $T(z)$ cause distortions over the approximated signal

$$S(z) = 0 \tag{3.30}$$

$$T(z) = \frac{1}{2}\left[H_0(z)H_1(-z) + H_1(z)H_0(-z)\right] \tag{3.31}$$

One of the several choices for canceling amplitude and phase distortions, is FIR Para unitary solution which relates $H_0(z)$ and $H_1(z)$, where both are N-tap FIR filters and power complementary pairs, given as,

$$H_1(z) = z^{-(N-1)}H_0(-z^{-1}) \tag{3.32}$$

Inserting this into the transfer equation $T(z)$ yields a pure delay multiplied by a constant as follows,

$$T(z) = \frac{1}{2}z^{-(N-1)}\left[H_0(z)H_0(z^{-1}) + H_1(z)H_1(z^{-1})\right] \tag{3.33}$$

$$T(z) = cz^{-(N-1)} \tag{3.34}$$

### 3.4 2-Dimensional 2-Channel Subband Decomposition

2-Dimensional filter bank implementation is utilized especially in image processing. 2-D filters can either be separable or non-separable. Each has specific properties and filter type should be chosen according to the application. There are various applications, for example in image processing, in which non-separable 2-D filters are preferable. On the other hand separable filters are constructed by multiplying two 1-D filters. And separable filters provide several advantages over the non-separable ones. The most important properties of separable filters are given as

1- No need to process along both dimensions at the same time. Each dimension can be processed at a time, for example a 2-D image can be filtered by a 1-D filter first filtering along rows and then later filtering the output along columns.

2- Separable filters are much easier since they do not require 2-D kernels

3- Operations using separable filters are faster, since $mn$ operations are needed in separable case while $m^2n$ operations are needed in a filtering operation using a non-separable filter

For the sake of simplicity we will restrict ourselves to separable 2-D filters, and process on the rows and columns of the input signal successively. In Equation (3.35) 1-D filters $m_1$ and $m_2$ are multiplied to form 2-D filter $M$ and a numerical example is provided.

$$M(k_1,k_2) = m_1(k_1)m_2(k_2) \tag{3.35}$$

$$m_1(k_1) = \frac{1}{4}\{1\,2\,4\,2\,1\} \tag{3.36}$$

$$m_2(k_2) = \frac{1}{4}\{1\,2\,4\,2\,1\} \tag{3.37}$$

$$M(k_1,k_2) = \frac{1}{16}\begin{bmatrix} 1 & 2 & 4 & 2 & 1 \\ 2 & 4 & 8 & 4 & 1 \\ 4 & 8 & 16 & 8 & 4 \\ 2 & 4 & 8 & 4 & 2 \\ 1 & 2 & 4 & 2 & 1 \end{bmatrix} \tag{3.38}$$

# 4. WAVELET TRANSFORM BASED SUBBAND DECOMPOSITION

In this section Wavelet Transform (WT) and its derivatives are explained. We begin with frequency analysis, explain FT briefly and introduce STFT which provides joint time-frequency analysis. In DWT, the concept of Multiresolution Analysis (MRA) is explained. In the following we start by linear series expansion of a signal, since all transformation methods reviewed here represent a signal as a weighted linear summation.

## 4.1 Linear Orthogonal Transforms

A signal in a vector space S can be written as a linear combination of basis vectors of space B. This representation of the signal in another way may be suitable for a specific problem type.

$$x = \sum_i a_i \psi_i \tag{4.1}$$

Where, the set of elementary signals $\{\psi_i\}\ i \in \mathrm{Z}$ is complete for the vector space S, if all signals $x \in S$ can be expanded as in (4.1). The dimension of the vector space S equals the number of the elements of the basis vector space B [35].

If the vectors $\psi_i$ are linearly independent, in that case there exists a dual set $\{\widetilde{\psi}_i\}\ i \in \mathrm{Z}$, such that the coefficients $a_i$ can be computed as

$$a_i = \sum_n \widetilde{\psi}_{i,n} x_n \tag{4.2}$$

If the set $\{\psi_i\}\ i \in \mathrm{Z}$ is orthonormal and complete, then B is an orthonormal basis, and the basis and its dual are the same,

$$\psi_i = \tilde{\psi}_i \tag{4.3}$$

The inner product, using Kronecker's symbol, is defined as

$$\langle \psi_i, \psi_j \rangle = \delta_{i,j} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{4.4}$$

In orthogonal case the equation for $x$ can be written as:

$$x = \sum_i \langle \psi_i, x \rangle \psi_i \tag{4.5}$$

If the basis vectors in the complete set are linearly independent but not orthonormal, then the set is biorthogonal. In that case the relations are given as:

$$\langle \psi_i, \tilde{\psi}_i \rangle = \delta_{i,j} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases} \tag{4.6}$$

$$x = \sum_i \langle \tilde{\psi}_i, x \rangle \psi_i = \sum_i \langle \psi_i, x \rangle \tilde{\psi}_i \tag{4.7}$$

If a redundancy in the complete basis vectors set exists, that is if the vectors are linearly dependent, then the representation has redundancy and is called a frame. Figure 4.1 shows a linear expansion of the vector $x$ onto the 2-D Euclidian space

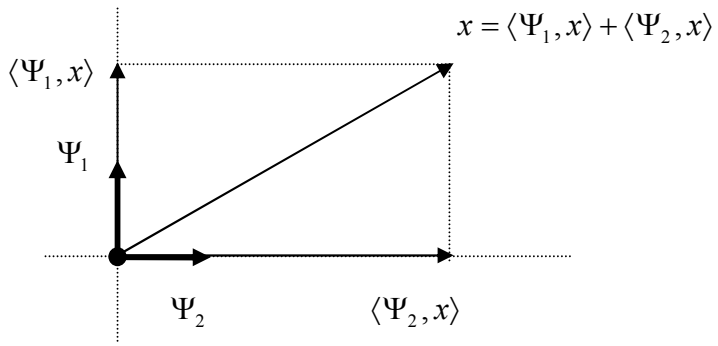


**Figure 4.1:** Linear Expansion of a Vector onto the Euclidian Space

## 4.2 Fourier Transform

The frequency representation of a continuous signal $x$ is the multiplication and integration of the signal by a sinusoidal kernel throughout the entire domain. Fourier Transform pairs for the continuous time signal $x$ and its transform $X$ are defined as:

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt \tag{4.8}$$

$$x(t) = \int_{-\infty}^{\infty} X(\omega) e^{j\omega t} d\omega \tag{4.9}$$

Where $j = \sqrt{-1}$ and the base function $e^{j\omega t} = \cos \omega t + j \sin \omega t$ is infinite in time and frequency. Fourier Transform separates the waveform into a sum of sinusoidal functions having different frequencies.

Signals which are discrete in time are transformed into the Fourier Domain as:

$$X(e^{j2\pi f}) = \sum_{n=-\infty}^{\infty} x[n] e^{-j2\pi n f} \tag{4.10}$$

$$x[n] = \int_{-\pi}^{\pi} X(e^{j2\pi f}) e^{j2\pi n f} df \tag{4.11}$$

Signal $X$ is continuous, periodic with $2\pi$, and is computed on the entire frequency domain, while $x$ is discrete in time. The frequency $f$ of $X$ is a normalization of the actual frequency to the sampling frequency. Since $X$ is continuous it is impossible to use its actual analog values in digital systems. The widespread use of digital systems requires discrete formulations of the expressions for both domains. As a result Discrete Fourier Transform equations are defined as:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi n k / N} \qquad k = 0,1,\ldots,N-1 \tag{4.12}$$

$$x[n] = \sum_{k=0}^{N-1} X[k] e^{j2\pi n k / N} \qquad n = 0,1,\ldots,N-1 \tag{4.13}$$

In Equations (4.12) and (4.13) the signal $x$ and its DFT $X$ is periodic with period N.

The basis functions used in Fourier Transform are infinite in time; conversely they are located well at frequency domain. Due to this property of the sinusoids information about the time is lost during Fourier Transformation. Since FT does not convey information about time to the transformed domain, the need to new techniques has arisen and a new version of Fourier Transform with time dependency, namely Short Time Fourier Transform is introduced.

## 4.3 Short Time Fourier Transform

Short Time Fourier Transform (STFT) is applied to the signals which do have rapidly changing properties with respect to time or space, in such cases the signal is segmented by finite duration windows. The length of the window depends on the length of time intervals during which the spectrum of the signal can be accepted as stationary. After windowing operation each windowed section is transformed into Fourier domain. Partitioning the signal into stationary parts is achieved through sliding the window. It should be noted that window is a rectangular one with amplitude 1, the width and amplitude of the window is constant.

$$STFT(\omega,k) = \int_{-\infty}^{\infty} f(t)\omega(t-k)e^{-j\omega t}dt \tag{4.14}$$

SHFT is invertible if the windowed function is of finite energy; the inverse is given as,

$$f(t) = \frac{1}{E}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} STFT(\tau,k)\omega(t-\tau)e^{-j2\pi tk}d\tau dk \tag{4.15}$$

Where $E = \int_{-\infty}^{\infty} |w(t)|^2 dt$. From another point of view the basis functions are,

$$w_{t,w}(\tau) = h(\tau - t)e^{j2\pi f\tau} \tag{4.16}$$

Which is same as that of Fourier Transform except for that, it is translated in time and modulated in frequency.

STFT provides time-frequency representation of a signal. Although this representation is very practical in use, it is subject to Heisenberg's Uncertainty Principle, a phenomenon known well from classical mechanics [36]. It dictates that both time and frequency information of a signal cannot be exactly known. The width of the window determines the resolution. Narrow window provides good time resolution and poor frequency resolution, while wide window acts opposite. The type of window to be utilized is strongly related to the application and the signals at hand. On the other hand to obtain good frequency resolution one needs a narrow band window. Although it is possible to adjust window size to trade-off between time and frequency resolution, there is a fundamental limitation on the freedom of design for a fixed window length. For a window $w(t)$ and its Fourier Transform $W(w)$, both satisfying $\int t|\omega(t)|^2 dt = 0$ $\int f|W(f)|^2 df = 0$, the spreads in time and frequency, using Root-Mean Square as a measure, are defined as:

$$\Delta_t^2 = \frac{\int t^2 |\omega(t)|^2 dt}{\int |\omega(t)|^2 dt} \tag{4.17}$$

$$\Delta_f^2 = \frac{\int f^2 |W(f)|^2 df}{\int |W(f)|^2 df} \tag{4.18}$$

The effective time duration and bandwidth of signals satisfy the following condition,

$$\Delta_t \Delta_f \geq \frac{1}{4\pi} \tag{4.19}$$

This means that, if a signal has bandwidth $\Delta_f$ then its duration must be $\Delta_t \geq \frac{1}{4\pi\Delta_f}$.

Two pulses in time can be distinguished if they are more then $\Delta_t$ apart, and two impulses at frequency domain can be distinguished if they are more then $\Delta_f$ apart. Due to this trade-off between time and frequency, resolution in time and in frequency can not be arbitrarily small, since their product is lower bounded.

## 4.4 Wavelet Transform

Following the above analyses another transformation method called Wavelet Transform gained high popularity among signal processing society in last two decades. Wavelet transform uses a family of special functions called wavelets. Unlike base functions used in FT and STFT which are sinusoidals having infinite duration and constant frequency, Wavelets are limited in time. Moreover a wavelet's width does not remain constant and is adjusted by a factor called scale. Scaling can be regarded as changing the frequency.

Instead of fixing the time-frequency resolutions $\Delta_t$ and $\Delta_f$, one can let both resolutions vary in time-frequency plane in order to obtain a multiresolution analysis [26]. This variation can be carried out without violating the Heisenberg inequality [36]. In this case, time resolution must increase as frequency increases and the frequency resolution must increase as frequency decreases. This can be obtained by fixing the ratio of $\Delta_f$ over $f$ to be equal to a constant $c$

$$\frac{\Delta f}{f} = c \tag{4.20}$$

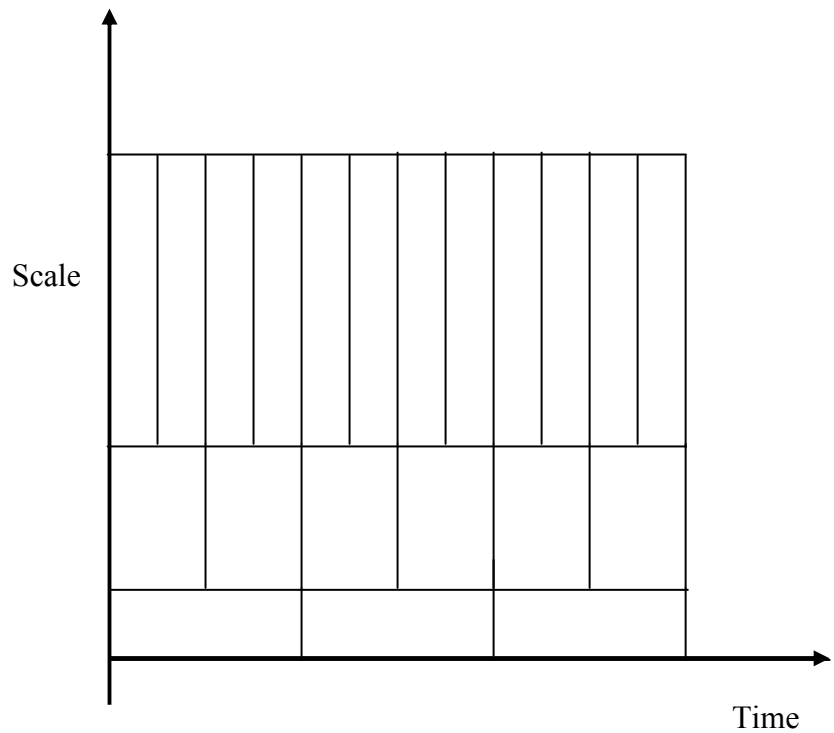


**Figure 4.2:** 2-Dimensional Time-Scale Representation of the Signal

### 4.4.1 Continuous Wavelet Transform

The forward and backward wavelet transform for 1-Dimensional signals is given in the equation below,

$$C(a,b) = \int_{t=-\infty}^{\infty} x(t)\Psi_{a,b}(t)dt \tag{4.21}$$

Where translations and dilations of the mother wavelet $\Psi(t)$ is as follows

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi(\frac{t-b}{a}) \qquad a = R^{+} - \{0\} \quad b = R \tag{4.22}$$

The synthesis equation of the signal $x$ is

$$x(t) = \frac{1}{K_{\Psi}} \int_{R^{+}} \int_{R} C(a,b)\frac{1}{\sqrt{a}}\Psi(\frac{t-b}{a})\frac{dadb}{a^{2}} \tag{4.23}$$

for the perfect reconstruction $K_{\Psi}$ which is the interval $0 < K_{\Psi} < \infty$ must satisfy following condition,

$$K_{\Psi} = \int_{-\infty}^{\infty} \frac{|\Psi(t)|^{2}}{|t|} dt \tag{4.24}$$

In Continuous Wavelet Transform (CWT) case mother wavelets' position is continuous over the time and also scaling is handled in continuous manner. Through Wavelet Analysis a signal is decomposed on a family of analyzing functions. As far as reconstruction is concerned CWT is highly redundant, since its base space is over complete [35].

### 4.4.2 Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) equation is identical to that of Continuous Wavelet Transform (CWT) except for the definition of $a$ and $b$. To obtain a time orthonormal basis, the scale and time parameters, $a$ and $b$ must be appropriately discretized. Natural way is to discretize them in a logarithmic manner, as for example, the adaptation of hearing of human beings is similar, and hence the subband coding of sound is successfully handled by the wavelet transformation. In discrete case, the resultant wavelet analysis equation is,

$$C(a,b) = \int_{t=-\infty}^{\infty} x(t)\frac{1}{\sqrt{a}}\Psi(\frac{t-b}{a})dt \qquad a = 2^{j},\ b = k2^{j}, (j,k) \in Z^{2} \qquad (4.25)$$

Synthesis of the signal is as follows

$$x(t) = \sum_{j}\ \sum_{k} C(j,k)\Psi_{j,k}(t) \qquad (4.26)$$

When the scale is changed in powers of 2, that is if the computations are done octave by octave, the transform is named dyadic wavelet transform

$$x(t) = \sum_{j}\ \sum_{k} C(j,k)2^{\frac{-j}{2}}\Psi(2^{-j}t-k) \qquad (4.27)$$

For this choice of scaling, the mother wavelet from which a family of orthogonal bases are derived as below and a true orthonormal basis will be obtained only for very special choices of $\psi$

$$\Psi_{j,k}(t) = 2^{\frac{-j}{2}}\Psi(2^{-j}t-k) \qquad (4.27)$$

the factor $2^{-j/2}$ normalizes each wavelet to maintain a continuous norm independent of scale $j$. Through this, lower resolution wavelets can be calculated from higher resolution coefficients. This approach is called DWT and is strongly related to Multiresolution Analysis (MRA) [37].

The concept of MRA is better understood by a function called scaling function.

$$\Phi_{j,k}(t) = 2^{\frac{-j}{2}} \Phi(2^{-j}t - k) \quad \text{For} \quad \text{k=1, 2, …} \tag{4.28}$$

Any continuous function $x(t)$ can be represented at a given resolution or scale $j_0$, by a sequence of coefficients given in $x_{j_0}(t) = \sum x_{j_0}(k)\Phi_{j_0,k}(t)$. In other words, the sequence $x_{j_0}(k)$ is the set of samples of the continuous function $x(t)$ at resolution $j_0$. The higher values of $j_0$ corresponds to higher resolution [38, 39].

The lower resolution scaling function filter can be expressed by a weighted sum of shifted versions of the same scaling function at the next higher resolution $\varphi(2t)$, as follows,

$$\Phi(t) = \sqrt{2}\sum_k h_0(k)\Phi(2t - k) \tag{4.28}$$

The set of coefficients $h(k)$'s are called the scaling function coefficients for the scaling filter. This equation is also called,

1-the refinement equation

2-MRA equation

3-dilation equation

The mother wavelet can be constructed using scaling function, design of which depends on $h(k)$'s, the scaling function is related to the mother wavelet as below

$$\Psi(t) = \sqrt{2}\sum_k h_1(k)\Phi(2t - k) \tag{4.29}$$

Where the relationship between real scalar series $h_0$ and $h_1$ is given in (4.30) as defined in [38].

$$h_0(k) = (-1)^k h_1(1 - k) \tag{4.30}$$

$h_0(n)$ is related to the original waveform $x(t)$. For a given scalar series $h(t)$ it is difficult to solve for the scaling function. Scaling function is a smoothing function in behavior where wavelet function acts as a detailing one.

It is shown in [38] and [40] that any continuous function can be represented by the following expansion, defined in terms of a given scaling function and its wavelet derivatives,

$$x(t) = \sum_{k=-\infty}^{\infty} c_{j0}(k)\Phi_{jo,k}(t) + \sum_{j=j0}^{\infty}\sum_{k=-\infty}^{\infty} d_j(k)\Psi_{j,k}(t) \tag{4.31}$$

this set of coefficients in the wavelet expansion is called the DWT of the function $x(t)$. In this expansion, the first summation gives a function that is a low resolution or coarse approximation of $x(t)$ at scale $j_0$. For each increasing $j$ in the second function, a finer resolution function is added, which increases details.

## 4.5 Multiresolution Decomposition and Reconstruction

It is shown that the scaling and wavelet coefficients at scale $j-1$ are related to the scaling coefficients at scale $j$ by the following relations [38, 40].

$$\begin{aligned} c_{j-1}(k) &= \sum_m h_0(m-2k)c_j(m) \\ d_{j-1}(k) &= \sum_m h_1(m-2k)c_j(m) \end{aligned} \quad for \quad j = J, J-1, ..., j_0+1 \tag{4.31}$$

The scaling coefficients at higher scale, using filter coefficients $h_0$ and $h_1$, can be used to calculate the wavelet and scaling coefficients at lower scales. Since in practice, a discrete signal, in its original resolution can be accepted as the first approximation, using the above relations in equation (4.31) one can completely determine the DWT coefficients of the signal at all desired levels, with only given filter coefficients $h_0$ and the original signal. The process described by the equation (4.31) takes place as follows; first by assuming that $c_J(k) = x(k)$, and starting from $j = J$ the coefficients $c_{j-1}$ are obtained filtering $c_j$ with the FIR filter $h_0$ and then decimating the output by keeping only every other sample of the output. The details

$d_{j-1}$ are obtained in a similar fashion [38, 41]. The filters $h_0$ and $h_1$ are low-pass and high-pass filters respectively. The procedure explained above gives the DWT coefficients set, for a signal $x(t)$ this collection is $\{d_{J-1}(k), d_{J-2}(k), ..., d_{j_{0+1}}(k), d_{j_0}(k), c_{j_0}(k)\}$. Total length of the elements of this set equals the length of the original signal. This is achieved via decimation by 2 at each stage. The above process is visualized in the Figure 4.3.

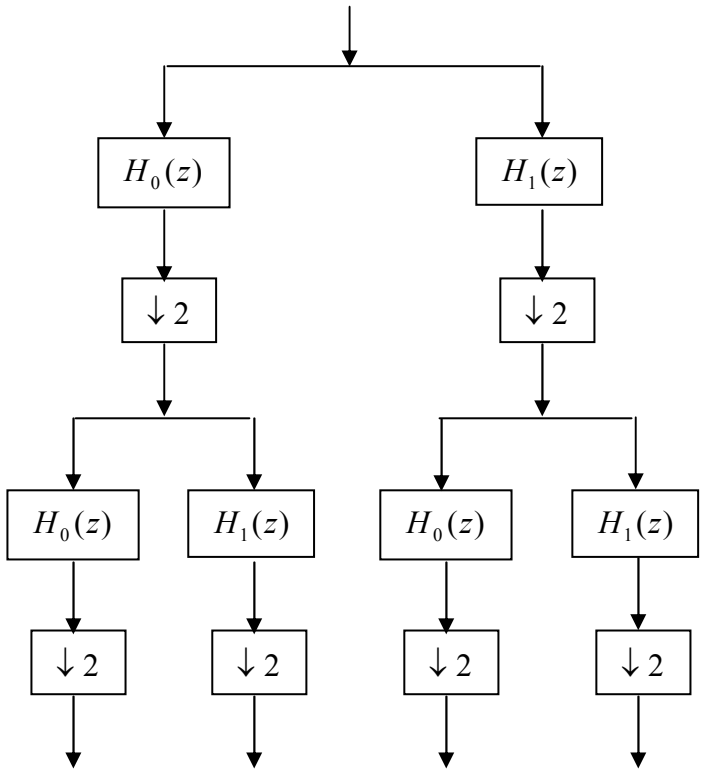


**Figure 4.3:** 1-Dimensional 2-Level Full Subband Decomposition

The higher resolution scaling coefficients are related to the lower resolution scaling and wavelet coefficients as described below

$$c_{j+1}(k) = \sum_m c_j(k) g0(k-2m) + \sum_m d_j(k) g1(k-2m)$$
$$for \quad j = j_0, j_{0+1}, ..., J-1 \tag{4.32}$$

this equation indicates how the approximation and detail coefficients at resolution $j_0$ can be used to reconstruct the approximation of the original signal at the maximum

achievable resolution J. In the reconstruction process, the FIR filters $g_0$ and $g_1$ are flipped versions of $h_0$ and $h_1$ respectively, and more importantly, before filtering, zeros are placed between every two consecutive samples. After expansion in time which corresponds to interpolation, filtering occurs and a finer approximation is obtained.

$$D_j(t) = \sum_k C(j,k)\Psi_{j,k}(t) \tag{4.33}$$

$$f(t) = \sum_j D_j(t) \tag{4.34}$$

$$A_J = \sum_{j>J} D_j(t) \tag{4.35}$$

$$f(t) = A_J + \sum_{j \le J} D_j \tag{4.36}$$

$$A_j = A_{j+1} + D_{j+1} \tag{4.37}$$

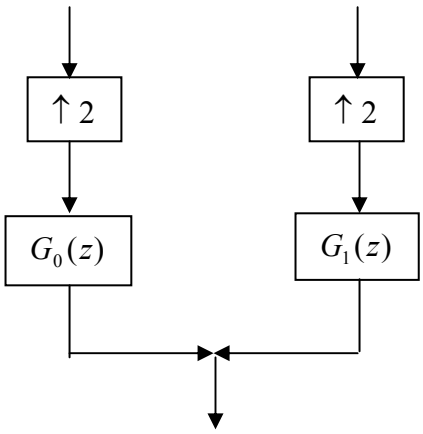


**Figure 4.4:** Interpolation and Reconstruction

## 4.6 Discrete Stationary Wavelet Transform

We know that the classical DWT suffers a drawback: the DWT is not a time invariant transform. This means that, even with periodic signal extension, the DWT of a translated version of a signal X is not, in general, the translated version of the DWT of X. To overcome the lost of the translation invariance which is a desirable property the Stationary Wavelet Transform (SWT) is defined. This property is useful

for several applications such as edge points detection. The main application of the SWT is de-noising. For more information, see [6, 42].

## 4.7 2-Dimensional Discrete Wavelet Transform

For the analysis and the processing of 2-D images, 2-D scaling functions and wavelets must be constructed. For the sake of simplicity we will restrict ourselves to the separable case as we did before for 2-D filter bank structures. Since the scaling and wavelet functions can be obtained from low-pass and high-pass filters respectively, 2-D scaling and wavelet functions are then defined as in equations (4.38) and (4.39). The rows and columns of the image are processed one by one to decompose the signal as in Figure 4.5.

$$\tilde{h}_0[k_1,k_2] = h_0[k_1]h_0[k_2] \tag{4.38}$$

$$\tilde{h}_1[k_1,k_2] = h_1[k_1]h_1[k_2] \tag{4.39}$$

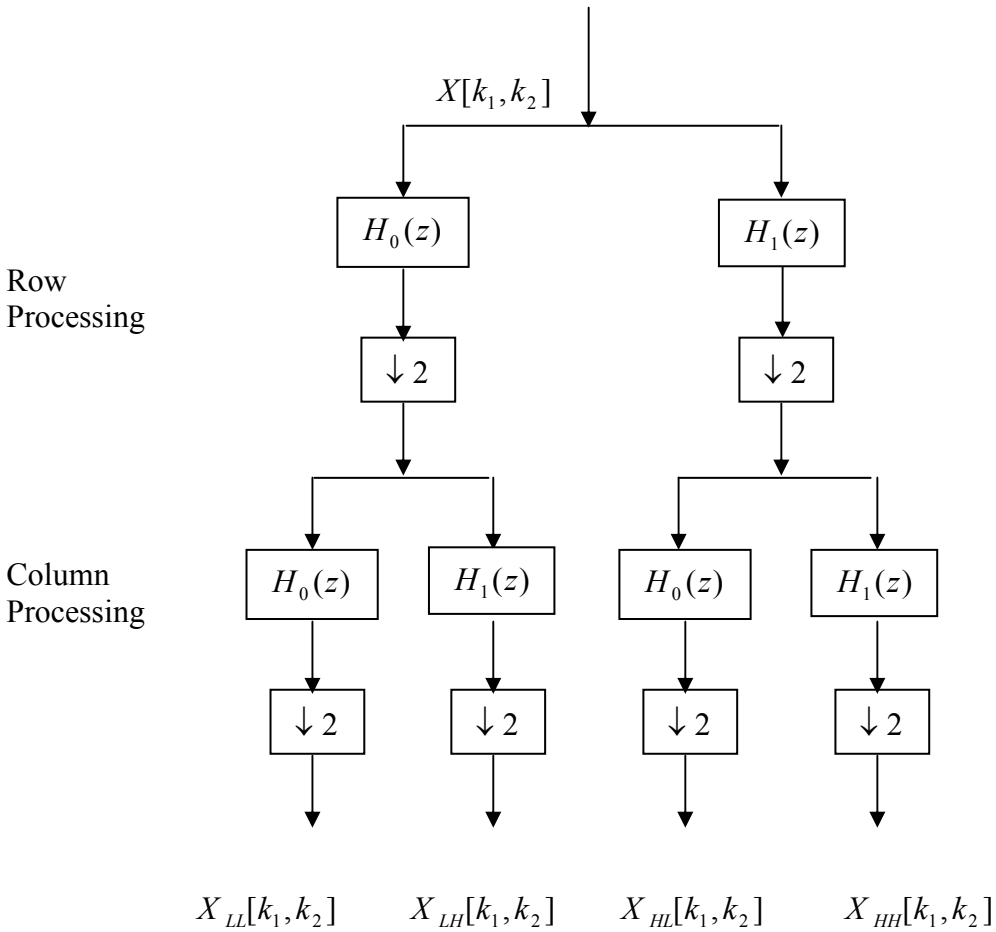


**Figure 4.5:** Separable Filtering Achieved Via Row and Column Processing

At each stage during the decomposition process 4 sub-images, approximation; horizontal, vertical and diagonal details are obtained. These subbands are shown in two dimensional domains as in Figure 4.6.

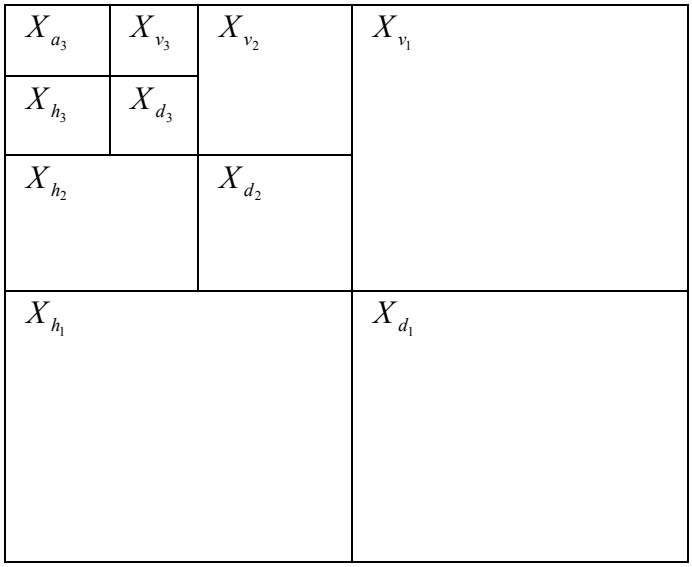


**Figure 4.6:** Representation of Subbands of an Image Decomposed in 3-Levels

# 5. SUBBAND DECOMPOSITION USING DFT

Frequency domain representation using DFT provides new perspectives on the signal. We can divide the spectrum into two regions, the high frequency region and the low frequency region. After separating the image into a blur approximation component and four high frequency components which are details, the sub images are back transformed to the spatial domain, and feature vectors that are used in segmentation are built from spatial domain subband images. By this method, both spatial and spectral properties of the image are used. This compact form which has both spatial and spectral characteristics is more powerful than the methods that use only spatial statistics of the gray level images.

## 5.1 Decomposition Algorithm

Decomposition algorithm which is explained briefly in the following steps is important for the participation of the well known Fourier Transform technique. Another important property of this technique is the utilization of the basic high-pass and low-pass filtering which is not complicated in computational means. Decomposition scheme which is used to extract features from image involves 5 main steps:

1-transform the signal into Fourier Domain using DFT

2-decompose the signal into high-pass and low-pass parts

3-back transform the image into its original domain, the spatial domain

4-subsample the image by decimating the image by 2 if required

5-Finally acquire features specific to the sub image to best characterize the image, and build feature vector

The general subband decomposition approach and the concept of filter banks are explained and detailed in sections 3 and 4**.** In this chapter the same methods will be

explained utilizing DFT, ideal HP and LP filtering and sub sampling. These three operations together are stated in the following equations. An application of the same scheme for the image fusion accomplishes the task and proves the ability to be an alternative to other decomposition techniques such as wavelet [6, 7].

### 5.1.1 Discrete Fourier Transform

The DFT pair is given as

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N} \qquad k = 0,1,\ldots,N-1 \tag{5.1}$$

$$x[n] = \sum_{k=0}^{N-1} X[k] e^{j2\pi nk/N} \qquad n = 0,1,\ldots,N-1 \tag{5.2}$$

### 5.1.2 Ideal High-Pass and Ideal Low-Pass Filtering and Decimation

Let $X_{LP}[k]$ be the low pass filtered part of $X[k]$, the DFT of the $x[n]$. If the filter used is a low-pass symmetric, zero-phase one then the expression in terms of $X[k]$ is [6, 7].

$$X_{LP}[k] = \begin{cases} X[k] & \text{for } 0 \le k \le K, N-K \le k < N \\ 0 & otherwise \end{cases} \tag{5.3}$$

and the low-pass signal is

$$x_{LP}[n] = \frac{1}{N} \sum_{k=0}^{N-1} X_{LP}[k] e^{j2\pi nk/N} \qquad 5.4)$$

This is the output of the filter that has frequency response given by

$$H[k] = \begin{cases} 0 & \text{for } 0 \le k \le K, N-K \le k < N \\ 1 & otherwise \end{cases} \tag{5.5}$$

and impulse response have symmetry about the mid-point

$$h[n] = h[N-n] \tag{5.6}$$

For the ideal low-pass expression if $K = N/2$ chosen, taking every other sample, the decimated expression reduces to.

$$x'_{LP}[n] = x_{LP}[2n] = \frac{1}{N}\sum_{k=0}^{N-1} X_{LP}[k]e^{j2\pi nk/(N/2)} \tag{5.7}$$

And for the high-pass signal the processes are taken in the same way as above. Applying a high-pass filter $G$ to the frequency response of the signal $x[n]$, $X_{HP}[k]$ is obtained. The frequency response of the filter is defined as

$$G[k] = \begin{cases} 1 & \text{for} \quad 0 \le k < K, N-K < k < N \\ 0 & \quad\quad k = K, K+1, \ldots, N-K \end{cases} \tag{5.8}$$

Again $g[n]$ has symmetry about the origin, and is a zero-phase filter. That is

$$g[n] = g[-n] \tag{5.9}$$

$$X_{HP}[k] = G[k]X[k] \tag{5.10}$$

or equivalently

$$X_{HP}[k] = \begin{cases} X[k] & \text{for} \quad K \le k < N-K \\ 0 & \quad\quad otherwise \end{cases} \tag{5.11}$$

and high pass signal is obtained by inverse transforming $X_{HP}[k]$

$$x_{HP}[n] = \frac{1}{N}\sum_{k=0}^{N-1} X_{HP}[k]e^{j2\pi nk/N} \tag{5.12}$$

for a real signal $x[n]$, we have

$$X_{HP}[k] = X^*_{HP}[N-k] \tag{5.13}$$

Taking the relation in the Equation 5.7 into account the decimated high-pass signal can be reduced to the expression given in Equation 5.14 below as:

$$x'_{HP}[n] = \frac{e^{j\pi n}}{N} \sum_{k'=0}^{N/2} X_{HP}[k'+K] e^{j2\pi nk'/N} \tag{5.14}$$

### 5.1.3 Reconstruction in the Synthesis Bank

The reconstruction is the formation of the original image from its subbands. It is required in image fusion, image compression and coding processes. Although we don't have to reconstruct the original image after extracting the information needed from the subbands and building feature vectors, in the following the steps of reconstruction process in the low-pass and high-pass synthesis banks are given.

For reconstructing the low-pass signal these steps are followed,

$$X'_{LP}[k] = \begin{cases} \frac{1}{2} X_{LP}[k] & , \quad 0 \le k < N/4 \\ \mathrm{Re}[X_{LP}[k]] & , \quad k = N/4 \\ \frac{1}{2} X^*_{LP}[N/2-k] & , \quad N/4 < k < N/2 \end{cases} \tag{5.15}$$

In order to reconstruct $X_{LP}[k]$ from $X'_{LP}[k]$ we need additional information on the imaginary part of $X_{LP}[N/4]$. This can be considered to be the key data for perfect reconstruction. In summary, the reconstruction of the low-pass signal is achieved as follows:

1- High frequency components $(N/4 < k < 3N/4)$ are filled with zeros

2- Low frequency components are obtained by using equation (5.15) and $\mathrm{Im}\{X_{LP}[N/4]\}$

3- the low frequency components for $(3N/4 \le k < N)$ are the complex conjugates of the low frequency components for $(0 < k \le N/4)$ in the case of the DFT

4- $x_{LP}[n]$ is recovered by the inverse transform of $X_{LP}[k]$

The equation for recovering high pass signal can be written as

$$X'_{HP}[k] = \begin{cases} 0 & , \quad 0 \le k \le N/4 \\ \frac{1}{2} X_{HP}[\frac{N}{4} + k] & , \quad otherwise \end{cases} \tag{5.16}$$

The reconstruction of the high-pass signal is achieved as follows:

1- High frequency spectral components $(N/4 < k < 3N/4)$ are obtained by eq.(5.16)

2- other low frequency spectral components are set to zero

3- $x_{HP}[n]$ is recovered by the inverse transform of $X_{HP}[k]$

## 5.2 Ideal LP and HP Filters Used to Decompose and Reconstruct a Signal

The ideal LP and HP filters used in the decomposition and reconstruction processes are shown in the figure 5.1 and 5.2. These filters are symmetric, zero-phase and used to construct non-overlapping digital filter banks [7]. Impulse response of the wavelet filter Daubechies-4 is given in the Figure 5.3.

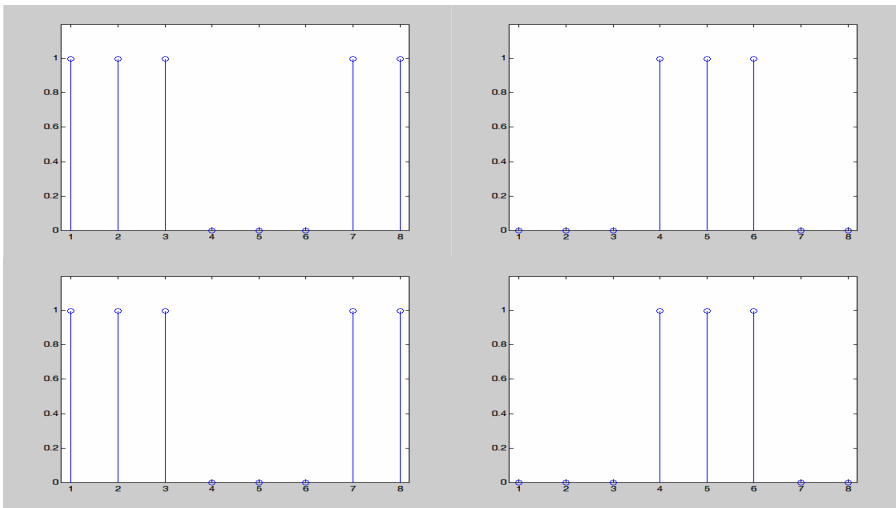

**Figure 5.1:** Frequency Responses of Ideal HP and LP Filters. Low-Pass Decomposition Filter: $H_0$, High-Pass Decomposition Filter: $G_0$, Low-Pass Reconstruction Filter : $H_1$, High-Pass Reconstruction Filter: $G_1$ ( from left to right and top to bottom).

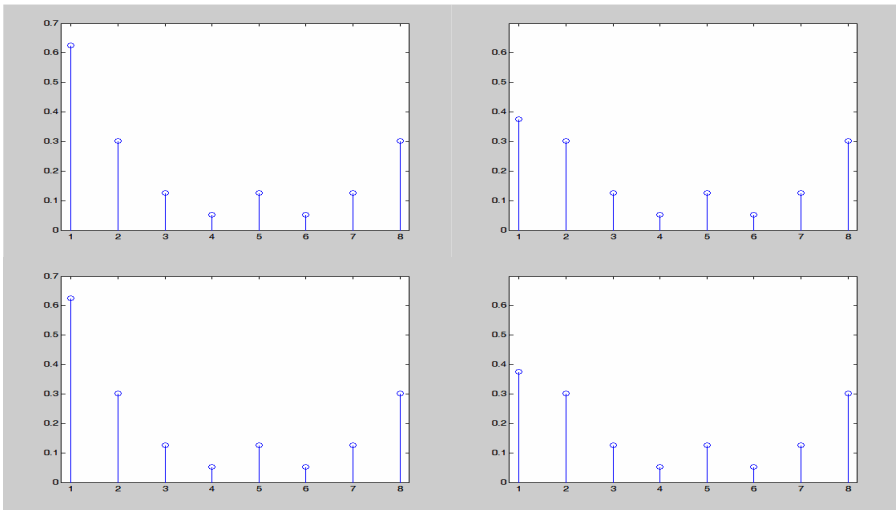

**Figure 5.2:** Impulse Responses of Ideal HP and LP Filters. Low-Pass Decomposition Filter: $H_0$, High-Pass Decomposition Filter: $G_0$, Low-Pass Reconstruction Filter : $H_1$, High-Pass Reconstruction Filter: $G_1$ (from left to right and top to bottom).

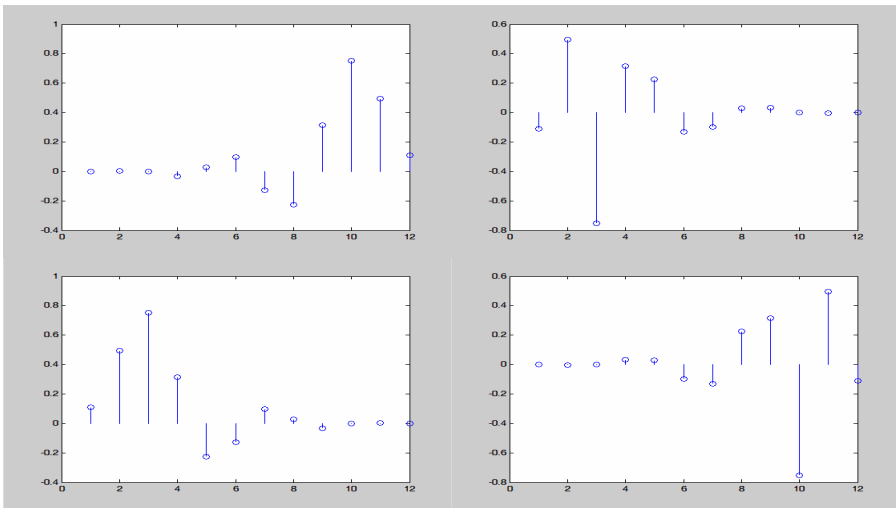

**Figure 5.3:** Daubechies-4 Filter Coefficients. Low-Pass Decomposition Filter, High-Pass Decomposition Filter, Low-Pass Reconstruction Filter, High-Pass Reconstruction Filter (from left to right and top to bottom).

### 5.2.1 An Example of a 1-Dimensional 3-Level Sub band Decomposition Using DFT

As an example a 1-Dimensional signal is decomposed at 3 levels into approximation and detail components,

$$x(n) = 1 + \cos(2\pi f_1 n) + \cos(2\pi f_2 n) + \cos(2\pi f_3 n) + \cos(2\pi f_4 n) + noise \qquad (5.17)$$

Where $n = 1,2,...,512$ and $f_1 = 3/512$, $f_2 = 5/512$, $f_3 = 10/512$, $f_4 = 150/512$ are normalized frequencies

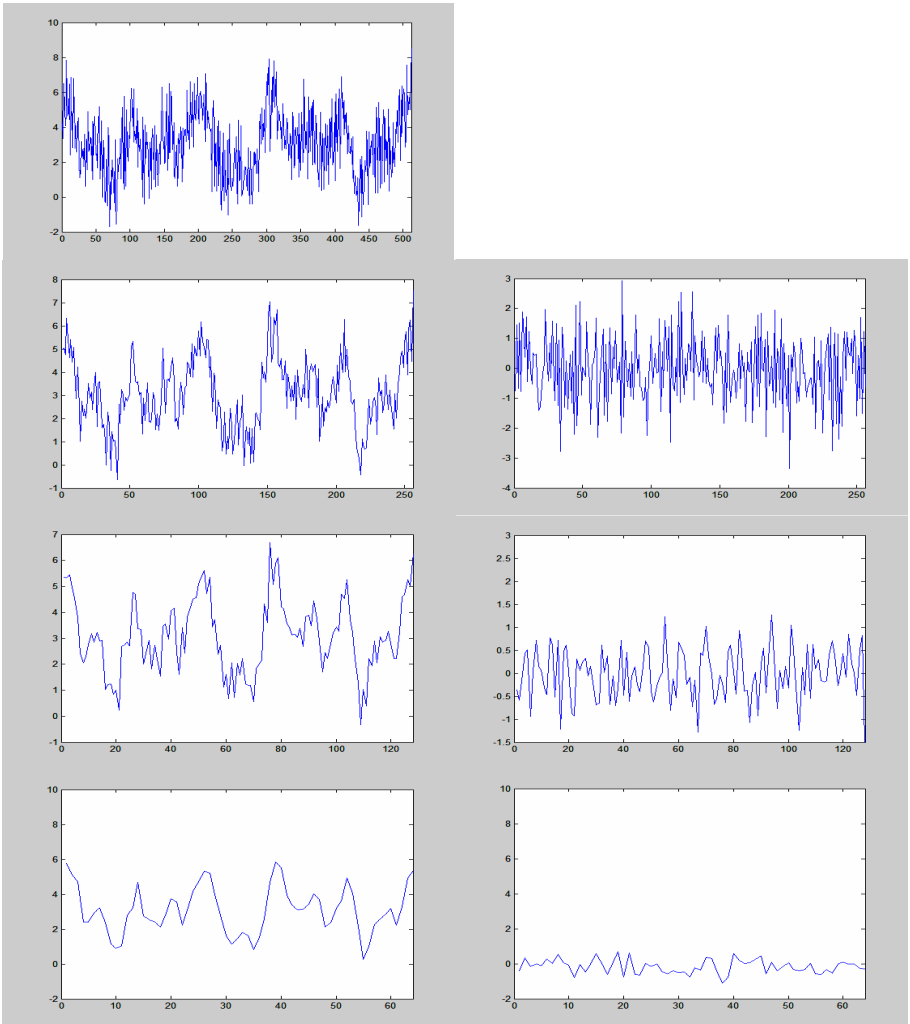

**Figure 5.4:** Example: 1-D Decomposition Using DFT Subband Decomposition. $x(n)$, $x_{LP}(n)$, $x_{HP}(n)$, $x_{LP,LP}(n)$, $x_{LP,HP}(n)$, $x_{LP,LP,LP}(n)$, $x_{LP,LP,HP}(n)$ (from left to right and top to bottom).

.

## 5.3 The Proposed DFT Transform Based Feature Extraction and Image Segmentation Algorithm

The feature extraction, segmentation and classification algorithm system is outlined in the following steps. Any such system is mainly composed of three parts:

- Pre-processing
- Processing
  - training phase
  - segmentation phase
- Post-processing

In most cases additional image processing operations are utilized. A typical pre-processing step includes the acquisition and preparation of the image and filtering operations such as mean, median, histogram equalization, contrast enhancement and denoising. A post-processing step may also include fundamental filtering operations such as mean, median, morphological filtering for better segmentation, image enhancement and representation processes [30, 31]. In order to get appropriate results we will apply median filtering operation to remove the small points, which still remains after segmentation. And a mathematical morphological operation will also be used to erode the points which are most probably not the interested objects. Test images used are 256 grey level images. First, we select images from Brodatz Texture Database. They are used for testing the performance of the proposed algorithm, to assure its accuracy. New images are synthesized using images from that database [2]. Images are 128x128 pixels, monochrome images. If they are not, they are resized by mathematical operations. The file formats of the artificial images are jpg, and gif. For the real images, to reduce the computation time, an interested area of the image is chosen, and the file type is generally converted from tiff to jpg, gif or bmp.

### 5.3.1 Flow Diagram of DFT Subband Decomposition

Flow Diagram of DFT Based Subband Decomposition Algorithm Proposed in this thesis is given in Figure 5.5. This algorithm is called as a subroutine at step 1.2 of the

training phase and at step 2.2 of the segmentation phase of the DFT based segmentation scheme.

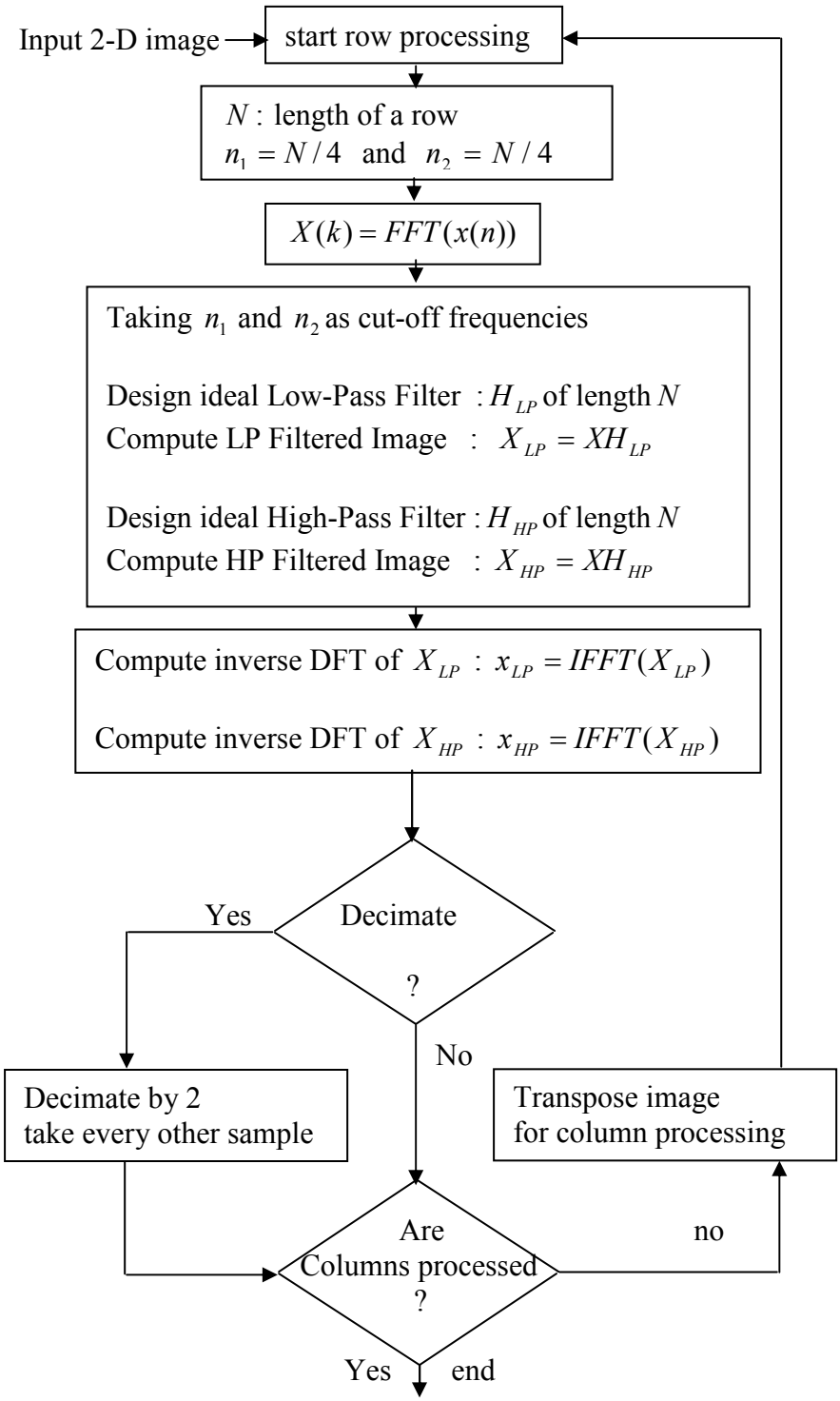


**Figure 5.5:** Flow Diagram of DFT Based Subband Decomposition Algorithm

### 5.3.2 Training Steps

In the test stage it is decided to use an 8x8 window to scan the image, after few experiments it is proven to be the best window size in both computation time, and accuracy rate for the 128x128 test image. At training phase the size of the windows used to scan the training image must be same as the size of the window used to test the image. Although this is not a must, it is experimentally observed that in that way it performed better. 16 images of size 8x8 pixel are cropped from each of the classes to train the algorithm and the features are kept as reference features. The feature vectors built at this stage include l1-norms and are normalized.

**1. Training Phase**

1.1 Select regions from image class i. compute feature vectors for each of 16 samples

1.2 Transmit each sub image window to DFT based decomposition algorithm

1.3 Build feature vector for class i sample j

1.4 Normalize the feature vector of class i sample j

1.5 Increment j by 1 and repeat for all samples

1.6 Increment i by 1 and repeat for all classes

1.7 Compute the mean value for all of the normalized feature vectors of class i

1.8 Increment i by 1 repeat for all classes

1.9 Pass normalized mean feature vectors of all classes to the testing step

### 5.3.3 Segmentation Steps

In a typical testing step we have an input image to be tested, reference feature vectors for each class. With these inputs segmentation algorithm outputs an image, either grey-level or binary, which is segmented. The input image is either an artificial image of size 128x128 pixel or a real SAR image. In both cases an 8x8 pixel window is run on the image starting from the top row and leftmost column of the image. At borders the symmetry of the image is taken for window operations. From two different modes of window sliding ways pixel by pixel and pixel group by pixel group we choose the first one, and decided for the pixel at the center of the window.

Since the window size is even, -not odd- we accepted the pixel point $(i+5, j+5)$ of the window $(i : i+8, j : j+8)$ as the center pixel and decision is made for that pixel.

**2. Segmentation Phase**

2.1 Run 8x8 pixel window on the 128x128 pixel image from left to right and top to bottom, in pixel by pixel manner.

2.2 At each pixel transmit the 8x8 sub-image window to DTF based decomposition algorithm as input

2.3 Build feature vector for the sub-image with center (i,j)

2.4 Normalize the feature vector of sub-image with center (i,j)

2.5 Compute the Euclidian distance $d$ between the normlized feature vector of class $d$ and the normalized feature vector of sub-image with center (i,j)

2.6 Increment $d$ by 1 and find distances for all classes

2.7 Compare the distances, find minimum distance to the test feature vector of sub-image with center (i,j)

2.8 Decide which class it belongs to and label the pixel (i,j) of test image window with the label of that class

2.9 Repeat steps from 2.1 to 2.8 until whole image is scanned by 8x8 windows

Subroutine called in 1.2 and 2.2 is depicted in the flow diagram shown in Figure 5.5. The steps of this routine are:

**3. DFT Low-Pass Decomposition**

3.1 Compute $N$ the length of one row of the input signal $x$

3.2 Compute length $N$ DFT of $X = FFT(x)$ of the signal $x$

3.3 Compute cutting frequencies of the LP filter $H_{LP}(k)$, $n_1 = N/4$, $n_2 = N/4$

3.4 Set coefficients of LP filter $0 \le k < n_1$ $n_2 < k < N$ to 1, set the others to 0

3.5 Compute $X_{LP}(k) = X(k)H_{LP}(k)$

3.6 Compute length $N$ inverse DFT of $X_{LP}$, $x'_{LP}(n) = IFFT(X_{LP}(k))$

3.7 Decimate $x_{LP}(n) = x'_{LP}(2n)$

3.8 Repeat steps form 3.1 to 3.7 for column processing

**4. DFT High-Pass Decomposition**

4.1 Compute $N$ the length of one row of the input signal $x$

4.2 Compute DFT of $X = FFT(x)$ of the signal $x$

4.3 Compute cutting frequencies of the HP filter $H_{HP}(k)$, $n_1 = N/4$, $n_2 = N/4$

4.4 Set coefficients of HP filter $n_1 < k < n_2$ to 1 and set the others to 0

4.5 Compute $X_{HP}(k) = X(k)H_{HP}(k)$

4.6 Compute inverse DFT of $X_{HP}$, $x'_{HP}(n) = IFFT(X_{HP}(k))$

4.7 Decimate $x_{HP}(n) = x'_{HP}(2n)$

4.8 Repeat steps form 4.1 to 4.7 for column processing

For Undecimated DFT, steps 3.7 and 4.7 are skipped.

# 6. EXPERIMENTAL RESULTS

In this section we apply the DFT based segmentation method both to the artificially synthesized images and to the remotely sensed images. The exact locations of the segments in the artificial images are known and therefore a performance measure can be driven and accuracy assessment is possible. Quantitative results that are obtained at the end of this stage are benefited when applying the segmentation algorithms to the SAR images. In the real remotely sensed image case, the assessments are made qualitatively and the qualities of the outputs are compared to see whether they coincide with the previous numerical results. In simulations, artificial images are built by composing different textures from Brodatz Album. A real SAR image from North Sea is studied for detecting oil spills on the sea surface. All of the images used are grayscale images.

The experiments are carried out in the following order:

Firstly the decimated version of the DFT based method is compared with DWT and then the undecimated version of DFT based method is compared with SWT. For a better visual presentation, median filtering and mathematical morphology are applied to the images as post processing operations. The resulting images are shown in the figures and accuracy rates are given in the tables for each experiment.

For either decimated or undecimated DFT based decomposition we have only one filter pair, but for both DWT and SWT we have to work with different wavelets from different families. The wavelets used are from Daubechies, Symlet, Coiflet and Biorthogonal wavelet families.

In the following four segmentation cases are studied

## 6.1 Simulation 1: 2-Classes Segmentation Problem

The first experiment is a simulation of an artificial image consisting two classes. The goal is to separate and label these regions. Firstly, we apply Decimated DFT based subband decomposition method and compare it to DWT based methods and then Undecimated DFT based subband decomposition method to SWT based methods.

### 6.1.1 Comparison of Decimated DFT and DWT Based Subband Decomposition Methods

From the Table 6.1 and Figure 6.1 we conclude that DWT based subband decomposition method regardless of the wavelet used, outperforms the Decimated DFT based subband decomposition method. The rates obtained by DWT vary from 0.8540 to 0.9100. The accuracy rate of Decimated DFT based subband decomposition method is poor, 0.7739. The decimation, sub sampling of the image, results in data loss and finally causes poor segmentation quality.

| Decomposition Method | | Accuracy Rates |
|---|---|---|
| Decimated DFT | | 0.7739 |
| DWT | db1 | 0.9100 |
| | db2 | 0.8540 |
| | sym2 | 0.8540 |
| | sym4 | 0.9204 |
| | sym5 | 0.9202 |
| | sym6 | 0.9206 |
| | coif1 | 0.9091 |
| | bior1.1 | 0.9100 |

**Table 6.1:** Accuracy Rates of Decimated DFT Based Subband Decomposition Method and DWT Based Subband Decomposition Methods

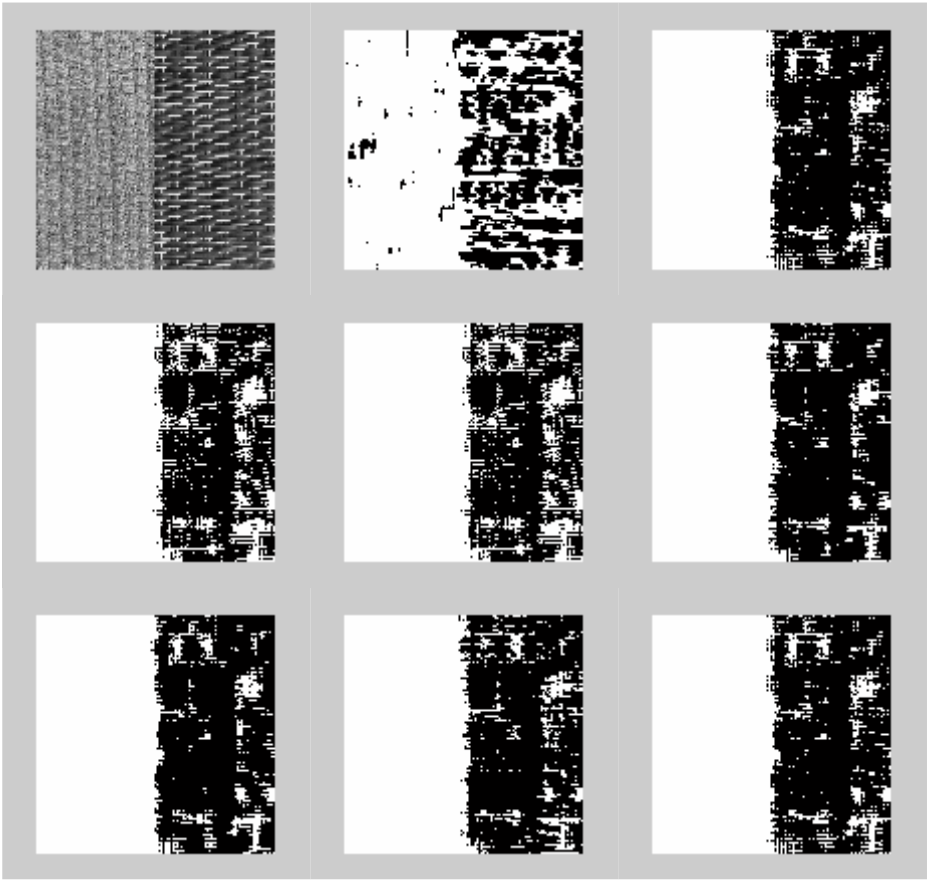

**Figure 6.1:** Synthesized Image, Decimated DFT and DWT Using: db1, db2, sym2, sym4, sym5, coif1, bior1.1 (from left to right and from top to bottom).

### 6.1.2 Comparison of Undecimated DFT and SWT Based Subband Decomposition Methods

The median filter used in this thesis, uses 8x8 image blocks, and replaces the value of the center pixel with the median of the 8x8 block. In morphological filtering an opening operation using a disk of diameter 3 as structural element is applied to the segmented image in order to get better segmentation. The accuracy rates are listed in Table 6.2. As it can be observed from the output images in Figure 6.2 and from the values given in the table, the Undecimated DFT based subband decomposition method gives better results than most of the wavelets. From the wavelets applied in the SWT based subband decomposition methods only db1 and bior1.1 are better than Undecimated DFT, the other wavelets except for these two have poorer performance. Moreover utilizing median filtering or morphological filtering increases the performance of all segmentation results. In both post-processing operations the

Undecimated DFT produces highest accuracy rates with the values 0.9712 and 0.9805. The outputs of the post-processed images are shown in Figures 6.3 and 6.4.

| Decomposition Method | | Accuracy Rates | | |
|---|---|---|---|---|
| | | | Median Filtering Applied | Mathematical Morphology Applied |
| Undecimated DFT | | 0.9247 | 0.9712 | 0.9805 |
| SWT | db1 | 0.9315 | 0.9681 | 0.9794 |
| | db2 | 0.9173 | 0.9572 | 0.9699 |
| | db3 | 0.9031 | 0.9458 | 0.9637 |
| | db4 | 0.8981 | 0.9466 | 0.9672 |
| | db5 | 0.8930 | 0.9438 | 0.9643 |
| | db6 | 0.8962 | 0.9477 | 0.9695 |
| | db7 | 0.8887 | 0.9413 | 0.9685 |
| | db8 | 0.8910 | 0.9471 | 0.9685 |
| | db9 | 0.8855 | 0.9421 | 0.9671 |
| | db10 | 0.8921 | 0.9466 | 0.9701 |
| | sym2 | 0.9173 | 0.9572 | 0.9699 |
| | sym3 | 0.9031 | 0.9458 | 0.9637 |
| | sym4 | 0.8878 | 0.9364 | 0.9606 |
| | sym5 | 0.8872 | 0.9366 | 0.9651 |
| | sym6 | 0.8777 | 0.9309 | 0.9609 |
| | coif1 | 0.9160 | 0.9551 | 0.9713 |
| | bior1.1 | 0.9315 | 0.9681 | 0.9794 |

**Table 6.2:** Accuracy Rates of Undecimated DFT and SWT Based Subband Decomposition Methods.

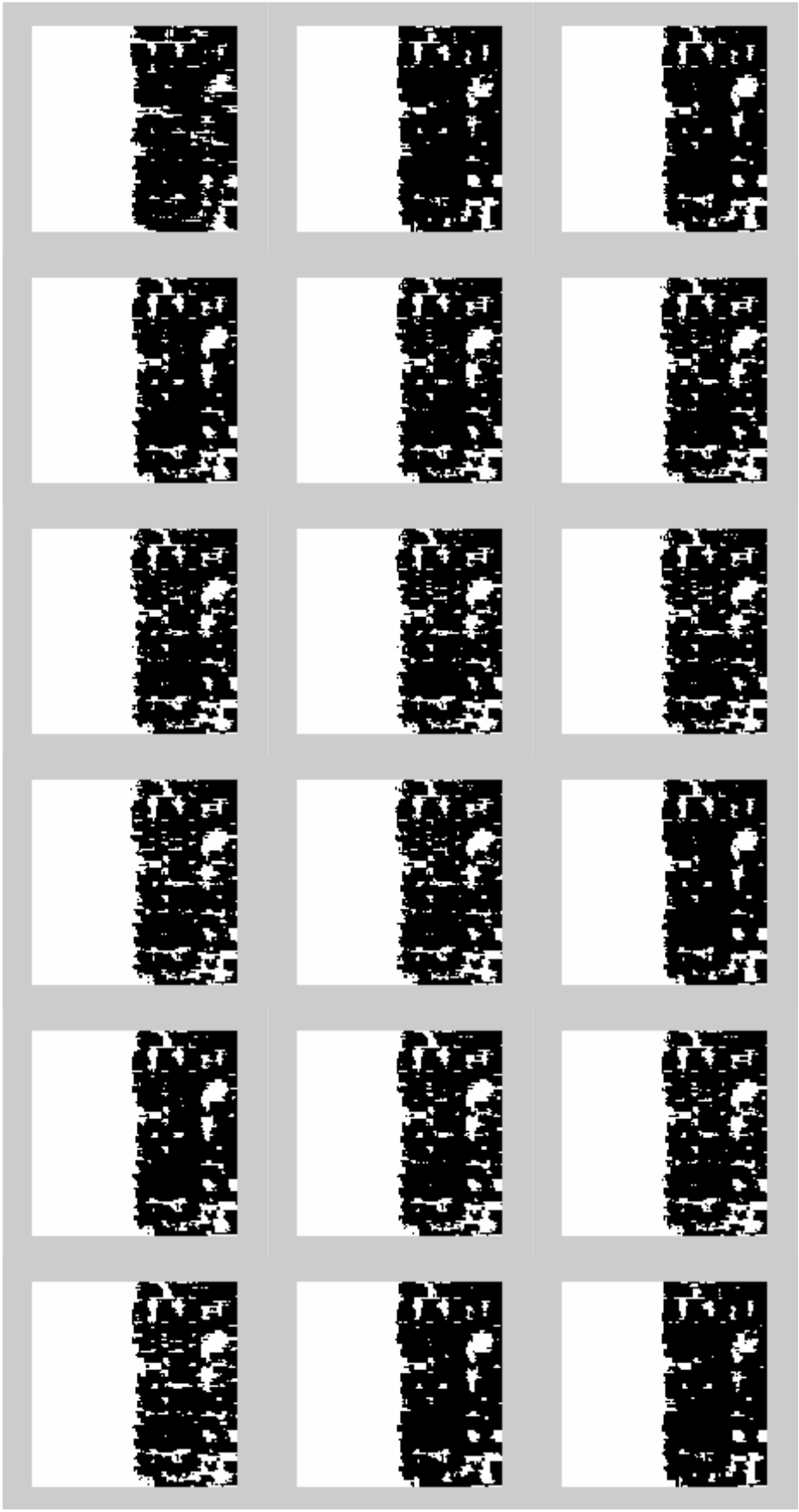

**Figure 6.2:** Undecimated DFT; SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, sym5, sym6, coif1, bior1.1 (from left to right and from top to bottom).

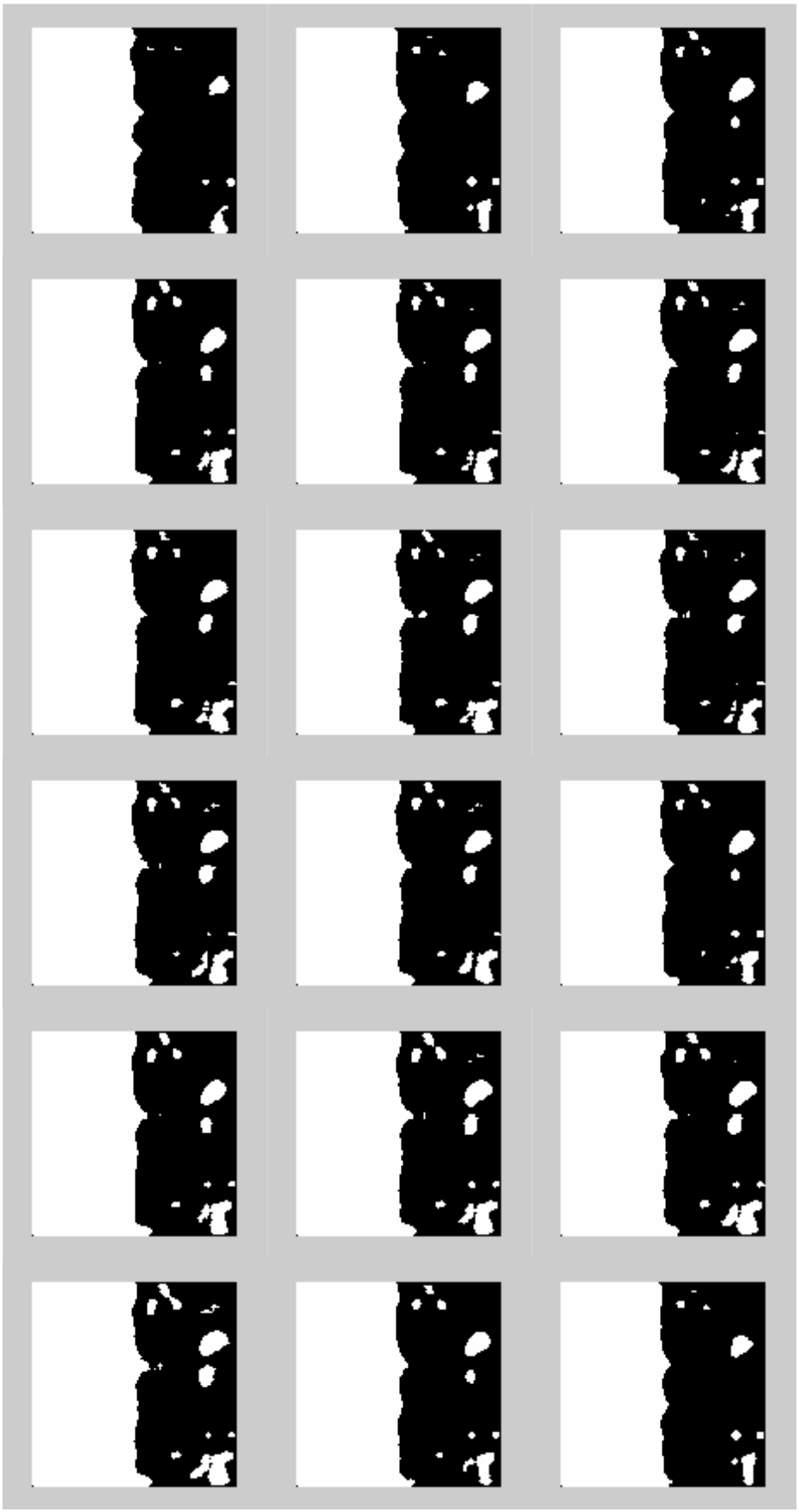

**Figure 6.3:** Median Filtering, Undecimated DFT; SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, sym5, sym6, coif1, bior1.1 (from left to right and from top to bottom).

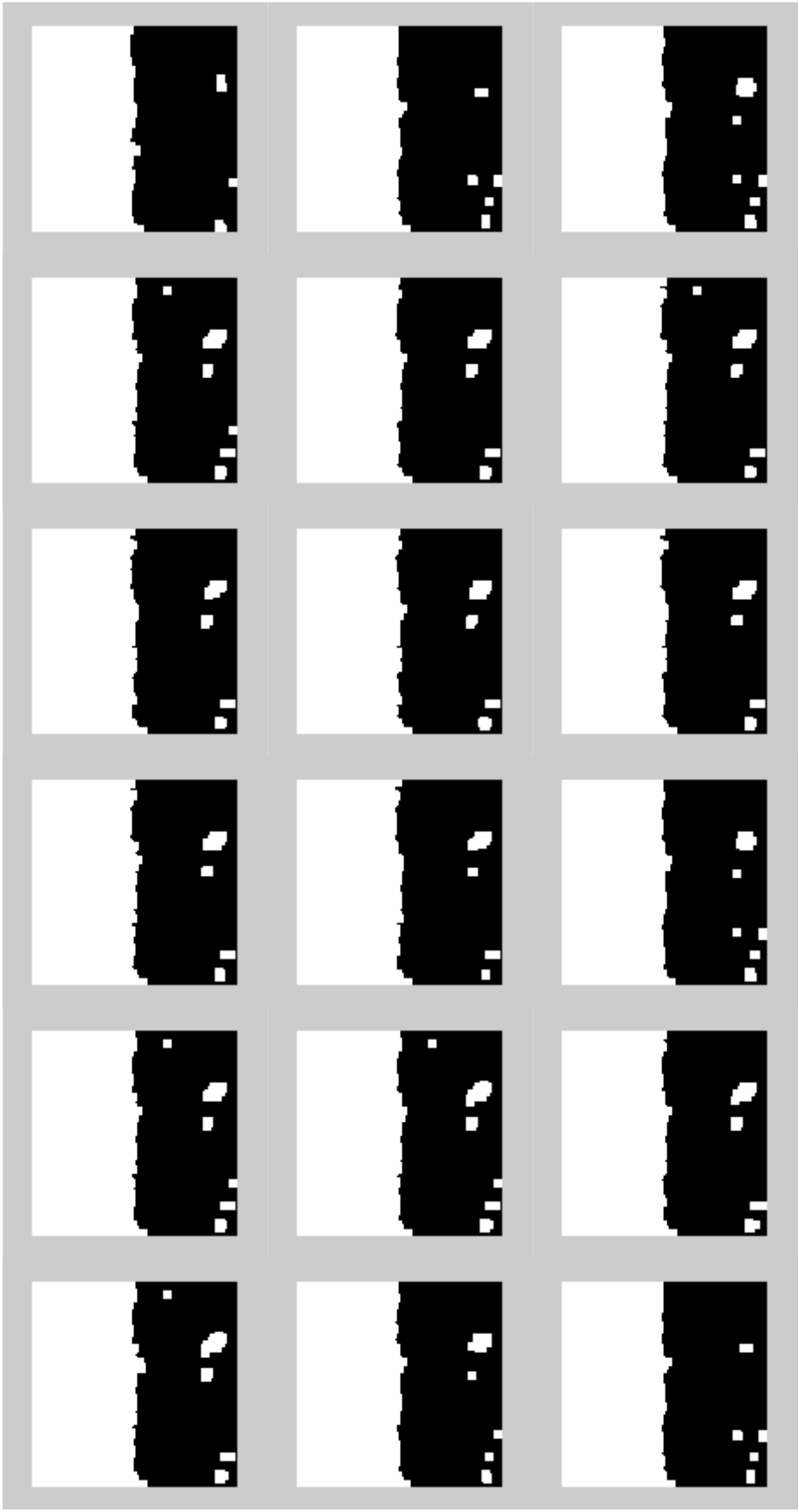

**Figure 6.4:** Morphological Filtering, Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym4, sym5, sym6, coif1, bior1.1 (from left to right and top to bottom).

## 6.2 Simulation 2: 4-Classes Segmentation Problem

This simulation is designed to segment four different textures. The input image which is 128x128 pixel is formed by four 64x64 pixel images. We compare only Undecimated DFT with SWT based decomposition methods.

### 6.2.1 Comparison of Undecimated DFT and SWT Based Subband Decomposition Methods

As it is seen from Figures 6.5 and 6.6, the performances of both Undecimated DFT and SWT based subband decomposition methods are satisfactory. The accuracy rate of DFT based method is 0.7645 while the results obtained from SWT are in the rage 0.7601-0.8219. Median filtering operation considerably increases the performances of all methods. But the increase of the performance after median filtering is better for all wavelets used in SWT then Undecimated DFT based subband decomposition method.

| Decomposition Method | | Accuracy Rates | |
|---|---|---|---|
| | | | After Median Filtering |
| Undecimated DFT | | 0.7645 | 0.8286 |
| SWT | db1 | 0.8219 | 0.9081 |
| | db2 | 0.8146 | 0.9070 |
| | db3 | 0.7978 | 0.8998 |
| | db4 | 0.7844 | 0.8934 |
| | db5 | 0.7762 | 0.8896 |
| | db6 | 0.7566 | 0.8841 |
| | db7 | 0.7601 | 0.8829 |
| | db8 | 0.7611 | 0.8868 |
| | db9 | 0.7620 | 0.8817 |
| | db10 | 0.7622 | 0.8762 |
| | sym2 | 0.8146 | 0.9070 |
| | sym3 | 0.7978 | 0.8998 |
| | sym5 | 0.7682 | 0.8895 |

**Table 6.3:** Accuracy Rates of Undecimated DFT and SWT Based Subband Decomposition Methods

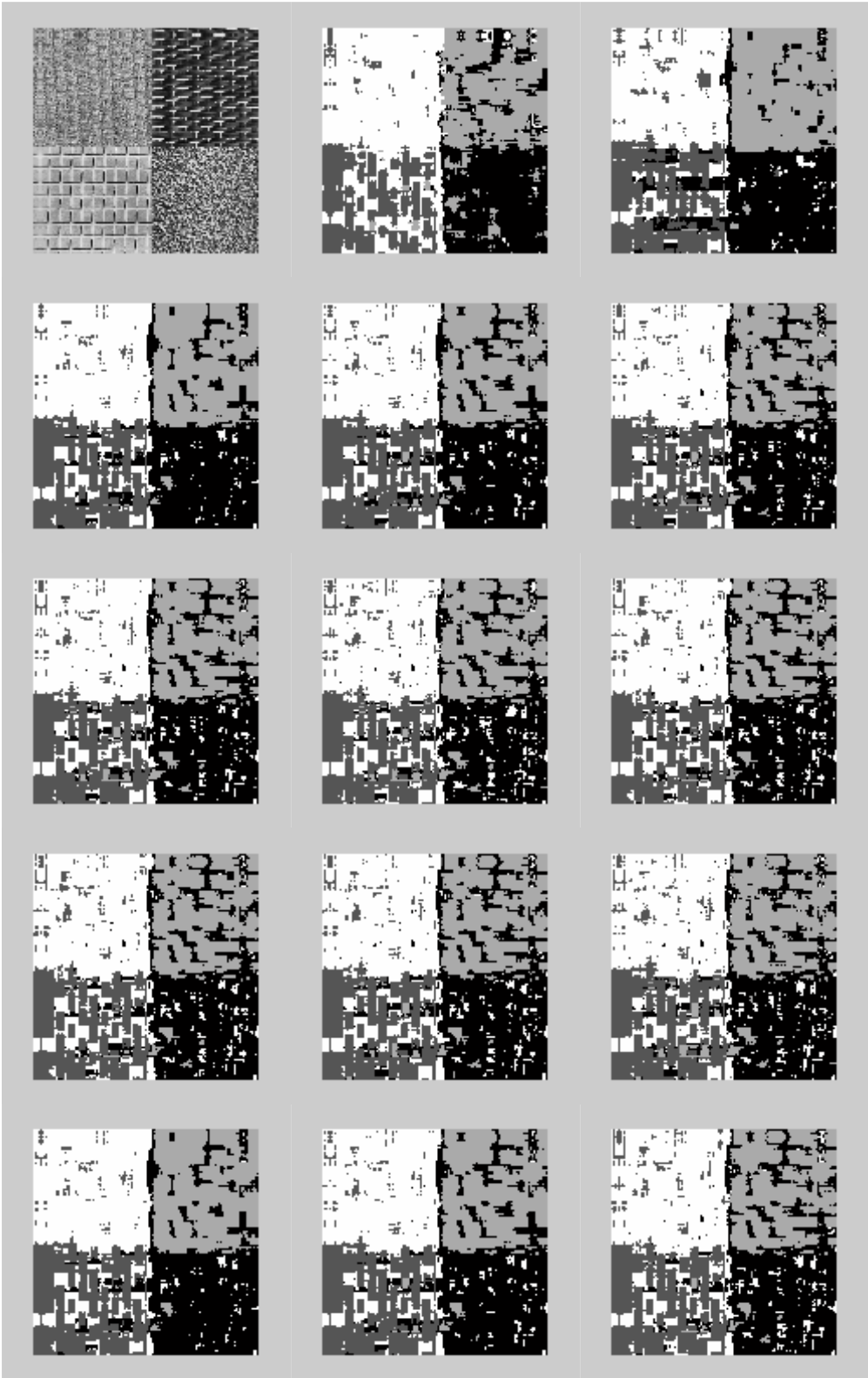

**Figure 6.5:** Synthesized Image. Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym5 (from left to right and top to bottom).

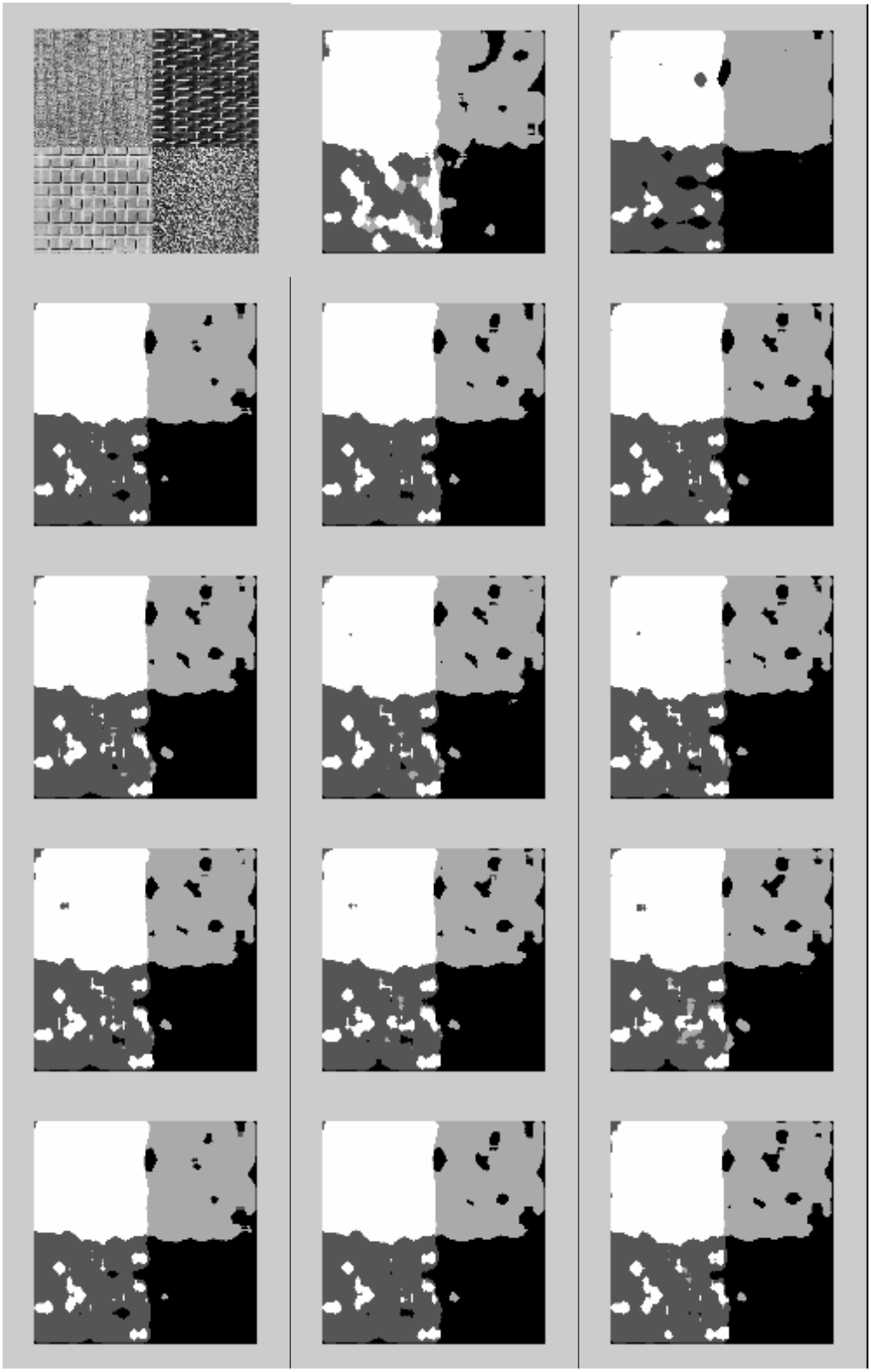

**Figure 6.6:** Median Filtered Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, db7, db8, db9, db10, sym2, sym3, sym5 (from left to right and top to bottom).

## 6.3 Simulation 3: Detection of 4 Objects on a Background

In this simulation it is intended to extract four different regions belonging to the same object from a uniform background. Textures of different sizes are placed on the background.

### 6.3.1 Comparison of Decimated DFT and DWT

For all wavelets DWT based method yields better results than Decimated DFT based subband decomposition method which is 0.9464 in Table 6.4. This result coincides with the results of Simulation1, where we have concluded that decimation decreases the performance of segmentation. The outputs are shown in Figure 6.7.

| Decomposition Method | | Accuracy Rates |
|---|---|---|
| Decimated DFT | | 0.9464 |
| DWT | db1 | 0.9723 |
| | sym2 | 0.9657 |
| | sym3 | 0.9659 |
| | sym4 | 0.9730 |
| | sym5 | 0.9732 |
| | sym6 | 0.9730 |
| | bior1.1 | 0.9723 |

**Table 6.4:** Accuracy Rates of Decimated DFT and DWT Based Subband Decomposition Methods

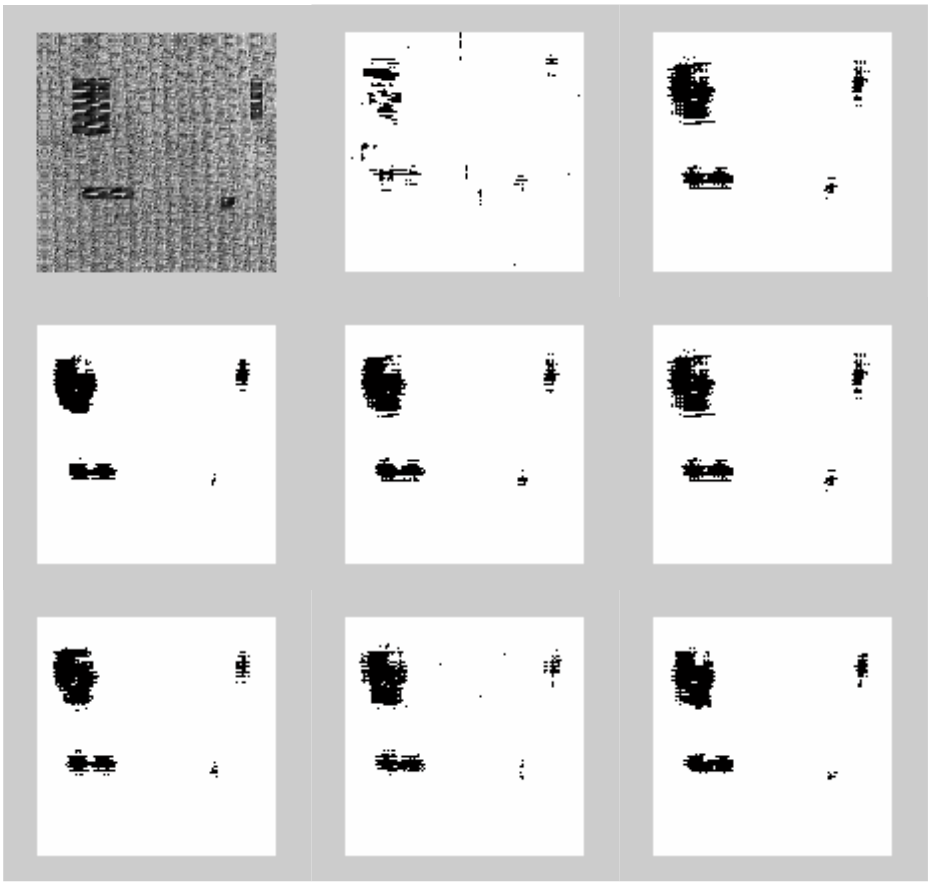

**Figure 6.7:** Decimated DFT and DWT Using: db1, sym2, sym3, sym4, sym5, sym6, bior1.1 (from left to right and top to bottom).

### 6.3.2 Comparison of Undecimated DFT and SWT Based Subband Decomposition Methods

As it is expected Undecimated DFT increases the performance of the segmentation. The rate of Undecimated DFT is 0.9738 while it was 0.9494 in Decimated DFT case. But the performance of Undecimated DFT is still less then the performance of all wavelets used in SWT. But this time the difference is in order of 0.001 which is a neglectable difference. Furthermore, both median filtering and mathematical morphology as post processing operators, increase performances of all methods. From the values listed in the columns of Table 6.5 and images shown in Figures 6.8, 6.9 and 6.10, it is seen that Undecimated DFT together with some post-processing operations can produce better results then SWT based decomposition methods.

| Decomposition Method | | Accuracy Rates | | |
|---|---|---|---|---|
| | | | Median Filtered | Mathematical Morphology |
| Undecimated DFT | | 0.9738 | 0.9753 | 0.9758 |
| SWT | db1 | 0.9755 | 0.9759 | 0.9760 |
| | db2 | 0.9781 | 0.9789 | 0.9777 |
| | db3 | 0.9773 | 0.9783 | 0.9740 |
| | db4 | 0.9757 | 0.9773 | 0.9745 |
| | db5 | 0.9748 | 0.9769 | 0.9736 |
| | db6 | 0.9739 | 0.9752 | 0.9729 |
| | sym2 | 0.9781 | 0.9789 | 0.9777 |
| | sym3 | 0.9773 | 0.9783 | 0.9740 |
| | sym4 | 0.9760 | 0.9774 | 0.9728 |
| | sym5 | 0.9747 | 0.9768 | 0.9745 |
| | sym6 | 0.9739 | 0.9752 | 0.9729 |

**Table 6.5:** Accuracy Rates of Undecimated DFT and SWT Based Subband Decomposition Methods.

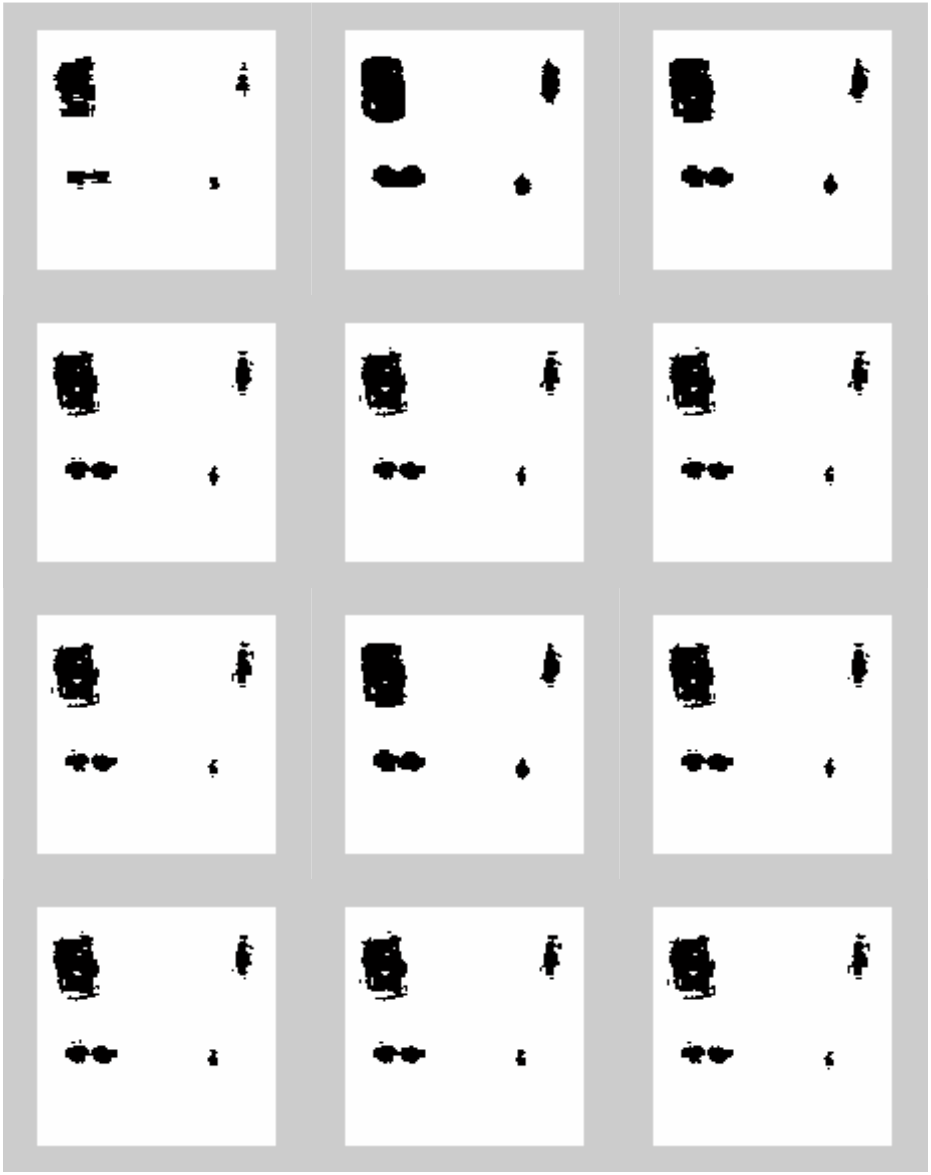

**Figure 6.8:** Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom).

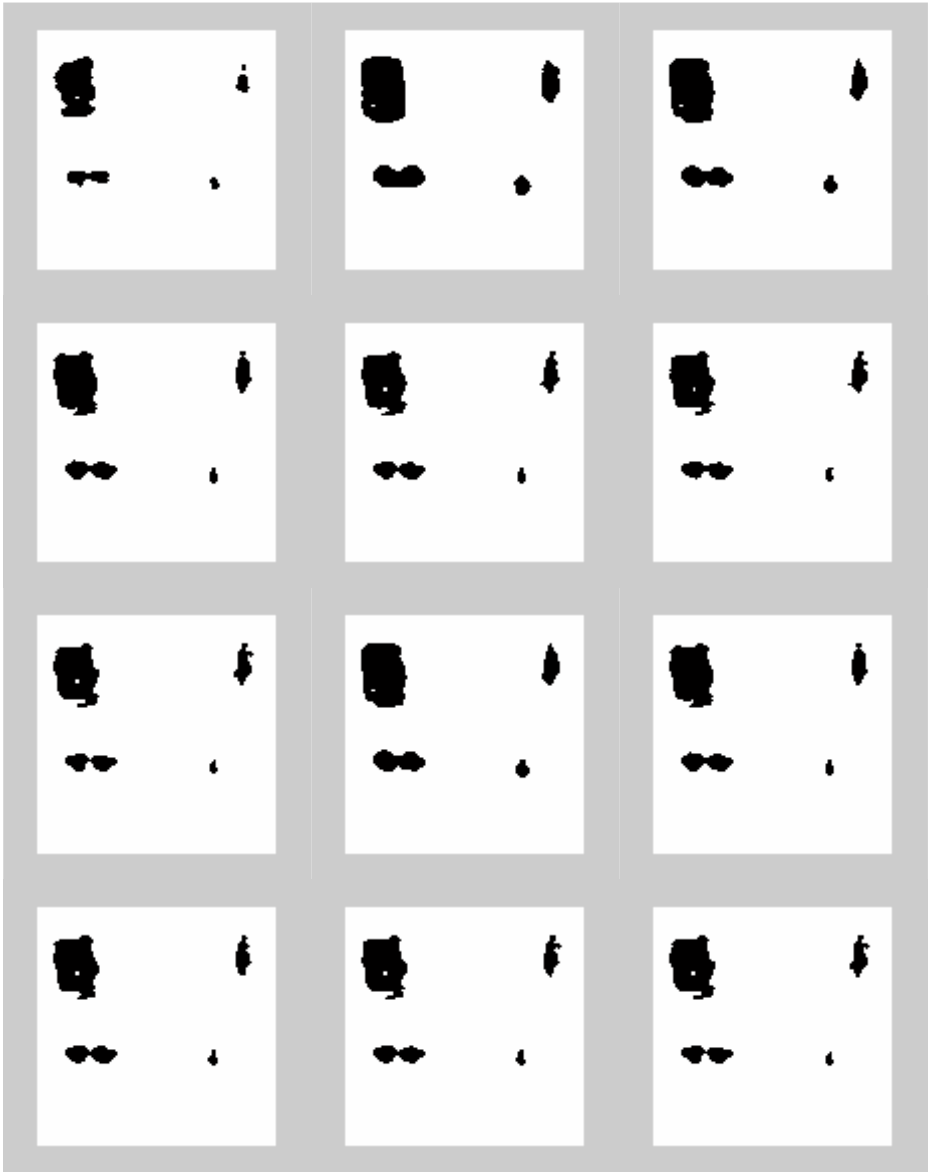

**Figure 6.9:** Median Filtered Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom).

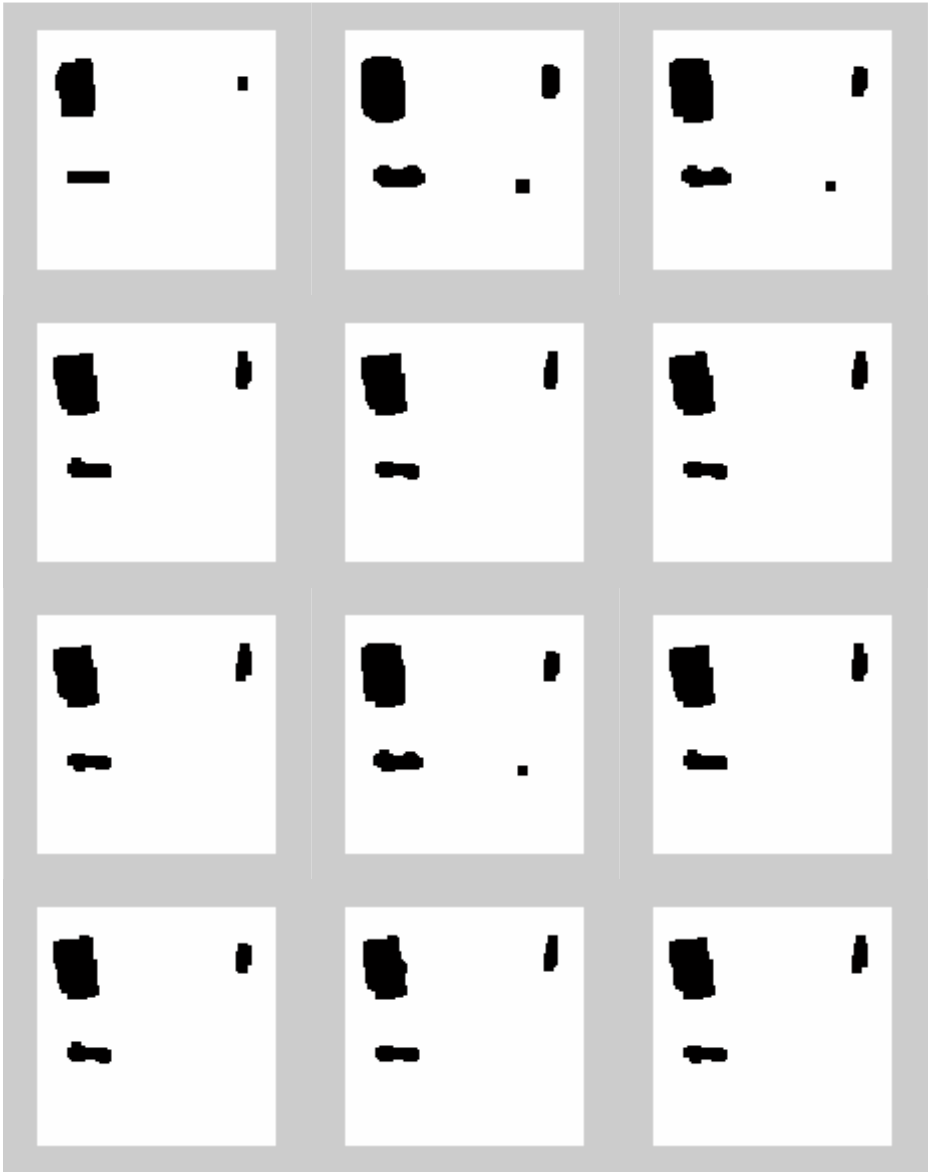

**Figure 6.10:** Mathematical Morphology of Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom).

## 6.4 Simulation 4: Detecting Oil Spills on the North Sea

A real SAR image of an oil spill accident is studied in this experiment. The image used is provided in the web site [http://earth.esa.int/ew/oil_slicks/noth_sea_96/]. The scene which is shown in the Figure 6.11 below belongs to North Sea and acquired from ERS-2 satellite on July 18, 1996. It is resized in MATLAB; dimensions of the original image are reduced to 128x128 pixels for segmentation process.

Median filtering is used to clear false detected points on the sea surface. The points which are scattered and do not have more points at the surrounding are successfully removed, since it is proven that median filtering is good at removing speckle noise. For all segmentation methods, the real SAR image is decomposed into subbands at 2-levels, by four transformation techniques, Decimated DFT, DWT, Undecimated DFT and SWT based subband decomposition methods respectively.

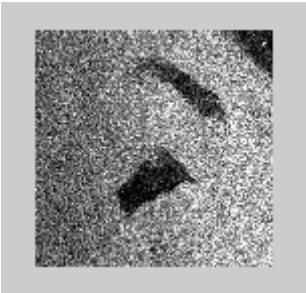

**Figure 6.11:** Real SAR Image Oil Spills at North Sea

### 6.4.1 Comparison of Decimated DFT, DWT, Undecimated DFT and SWT Based Subband Decomposition Methods

The top left image of the Figure 6.12 is the output of Decimated DFT based subband decomposition method and remaining images are products of DWT based subband decomposition methods. The image produced by the DFT based approach detected some small regions especially on the left side of the image as oil spills where these are not actual oil spills. However tree actual massive oil spills are successfully detected, and their shapes and boundaries are not certain. The db1 wavelet produces the best result among other wavelets. The boundaries of the image produced by db1 are uncertain. In addition to this disadvantage, almost all wavelets detect many points on the whole surface. These points are not actual oil spills, however, since the size of

these points are not large in size and scattered throughout the scene they can be removed by median filtering. Figure 6.13 shows the median filtered images. After median filtering, most of the small points are successfully cleared in all methods and improved quality can be visually observed. The masses detected as oil spills, where they are not actual oil spills, still remain in the resulting images of DWT using db3, db4, db5, sym3, sym4, sym5 as wavelets. But median filtering reduced the size of these regions in Decimated DFT method to an acceptable size and the shapes and locations of three oil spills are close to those of the real SAR image.

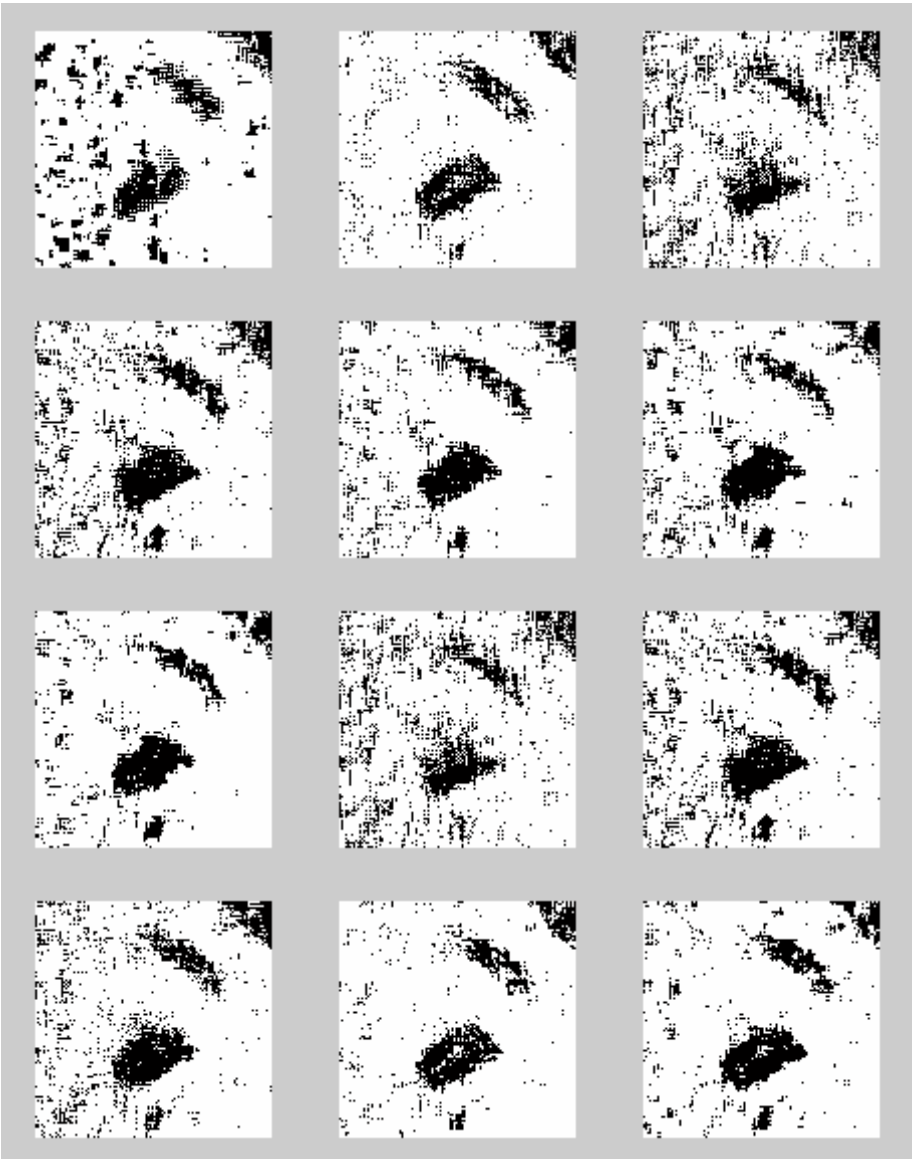

**Figure 6.12:** Decimated DFT and DWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom).

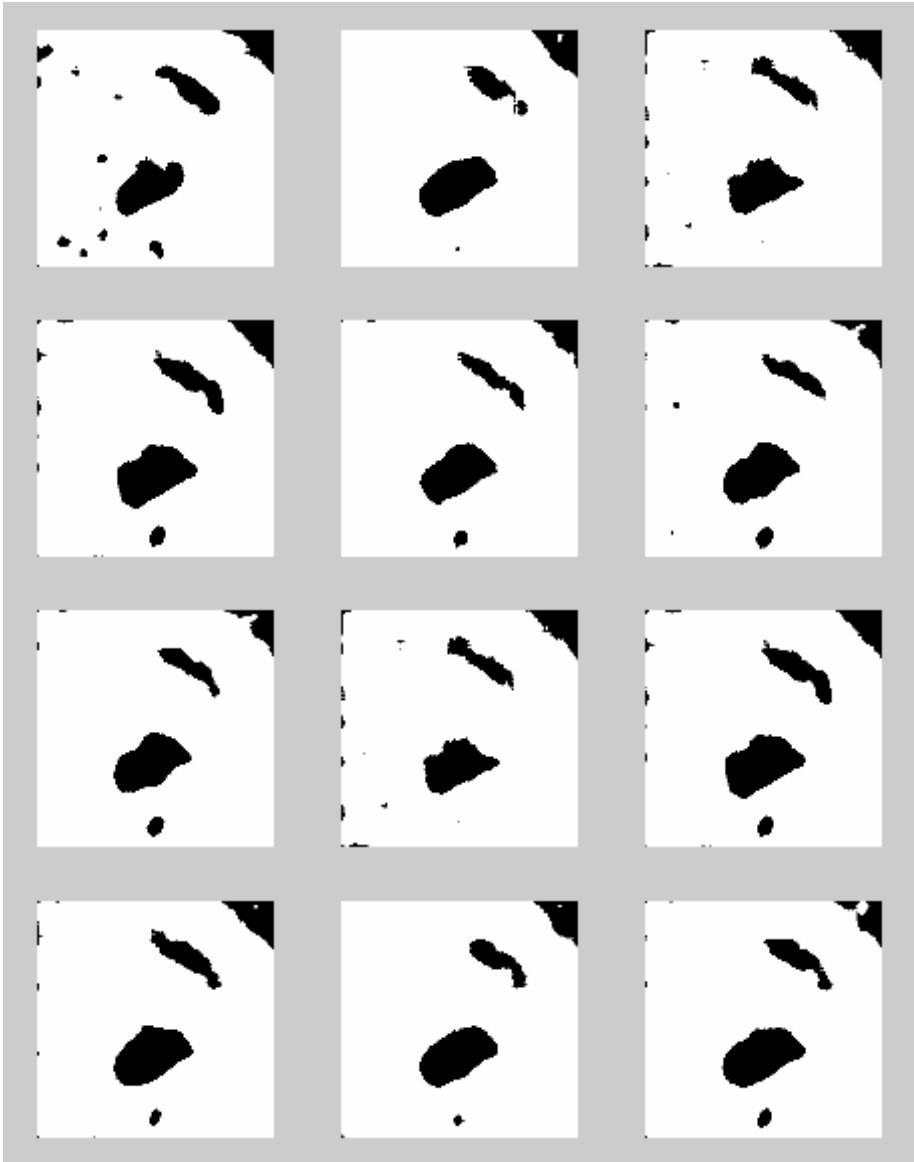

**Figure 6.13:** Median Filtered Decimated DFT and DWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom).

In Figure 6.14 and Figure 6.15 the images produces by Undecimated DFT and SWT using different wavelets are given. As it is expected the Undecimated DFT and SWT are better than Decimated DFT and DWT respectively since they retain much more information in the subbands. Moreover Decimated DFT in top left of the Figure 6.14

outperforms some SWT based decomposition methods. SWT using db1, db2, and sym2 seem to be better than Undecimated DFT, while in the images produced by SWT using db3, db4, db5, db6, sym3, sym4, sym5 and sym6 the masses detected are degraded and eroded. Incorrect small objects are removed with the help of median filtering as seen in the Figure 6.15. In Undecimated DFT and SWT using db3, db4, db5, sym3, sym4, sym5, sym6 wavelets a few small faulty regions exist. Besides this, SWT using db1, db2 and sym2 wavelets successfully clears all small objects and detect three actual oil spills with exact shapes and locations.

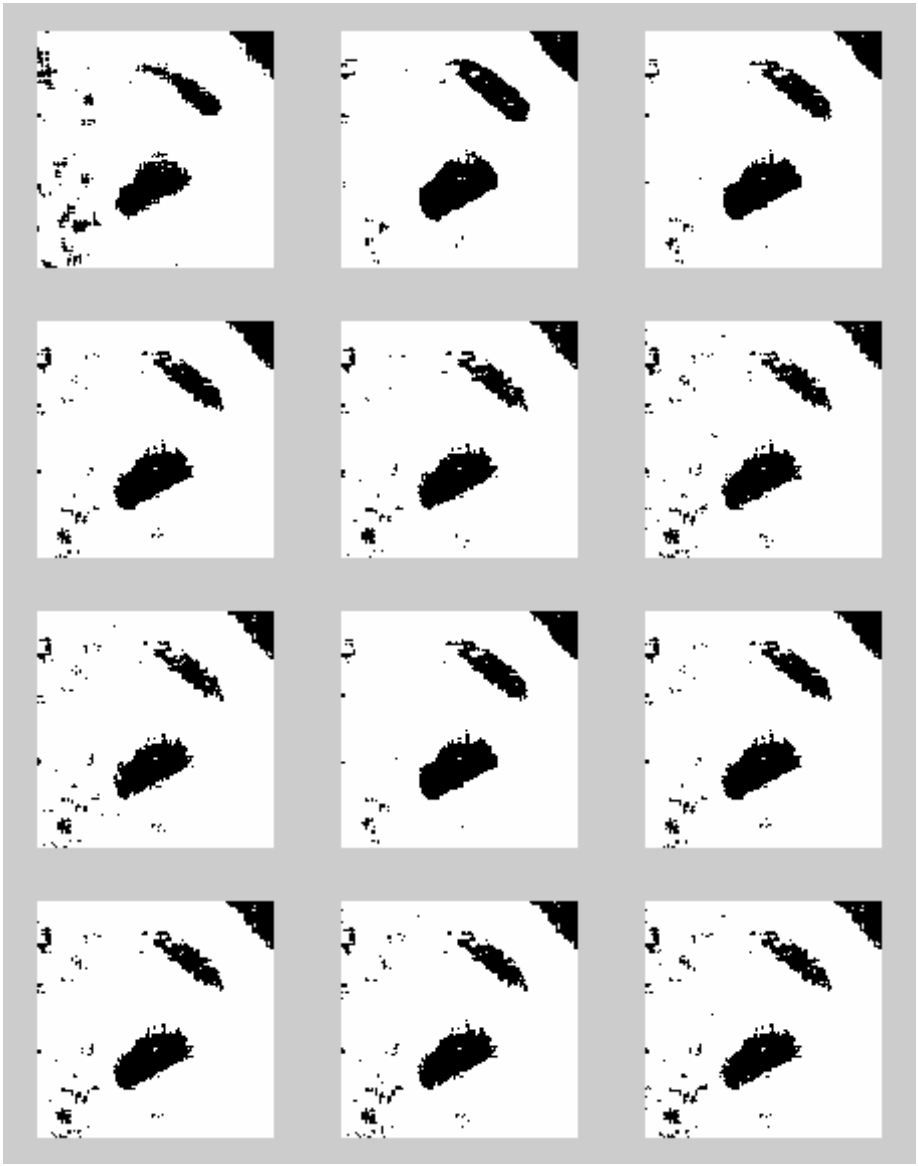

**Figure 6.14:** Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom).

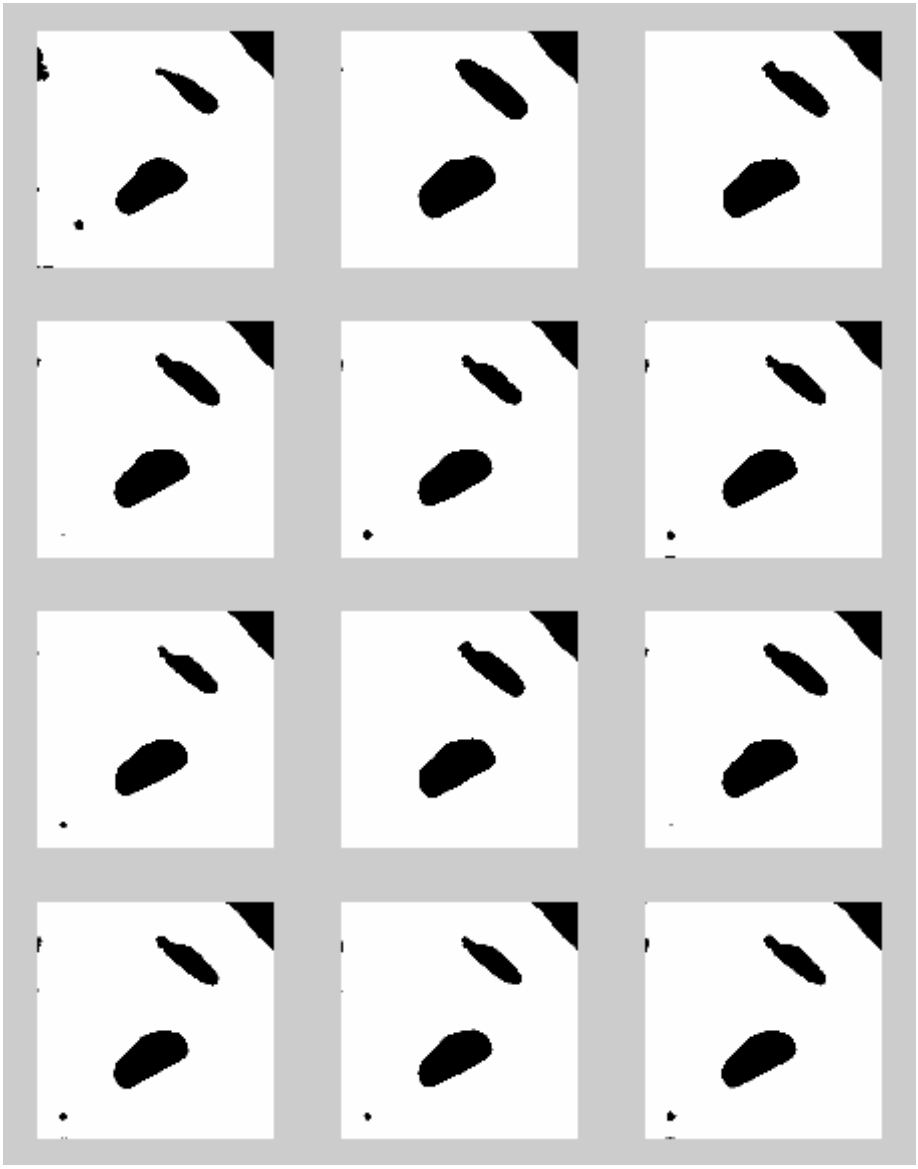

**Figure 6.15:** Median Filtered Undecimated DFT and SWT Using: db1, db2, db3, db4, db5, db6, sym2, sym3, sym4, sym5, sym6 (from left to right and top to bottom).

# 7. CONCLUSIONS AND DISCUSSION

A new DFT based subband decomposition scheme for segmentation of images is proposed in this thesis. Local energy distributions in spectral bands of signals are considered to be able to characterize the image. Feature sets are constructed using local energy distribution of the subbands of the image. Therefore feature extraction using subband decomposition is probably the most important phase of the overall segmentation phase. For efficient decomposition a 2-channel, zero-phase, non-overlapping digital filter bank is designed. The filters are ideal low-pass and high-pass filters which are designed in frequency domain. The size of the filters is always same as the size of the input signal and filters are applied to the frequency domain signal. After filtering in the frequency domain, images are transformed into spatial domain providing spatial-spectral features.

First we examined the effect of the down sampling on the performance of segmentation. For DFT based methods in case of down sampling we decimated the subband signal in the spatial domain at each level of decomposition. For the Wavelet Transform based method we have Wavelet Frames (SWT) which are over-complete counterparts of the critically sampled multirate analysis tool DWT. When exploring the effect of down sampling, the structure of the subband decomposition tree and general segmentation process did not change, only operations of down sampling are suppressed. Decimated DFT based method is compared to DWT and undecimated DFT based method is compared to SWT.

The experiments are carried out as follows; first we simulated a 2-class texture classification problem and tried to segment 2 regions in this synthetic texture. Second we applied same methods to another artificial image which is composed of 4 regions. In the third experiment, a problem which can be regarded as target detection problem, in which four objects belonging to the same texture are replaced on a uniform background, is simulated. Finally, the experiments are carried out on a real SAR image of an oil spillage accident which is a remotely sensed monochrome image with gray-levels. As post-processing operations we applied median filtering or morphological operations after classification.

The quantitative evaluation and performance assessment of the methods are carried out by computing confusion matrices. A weighted sum of the diagonal elements of this matrice is used to provide overall accuracy rate.

The results show that the proposed segmentation algorithm which uses DFT based subband decomposition may be considered as a good alternative to the conventional WT based approaches due to its less computation time, simple structure and sufficient segmentation performance.

We see that, as expected, decimating filter outputs degrades the objects in the resultant images. This is due to the lack of the transient information in the spatial domain of the sub sampled images. The major cause of the worse edge resolution and hence insufficient segmentation is due to sub sampling. Undecimated methods successfully detect edges and hence provide better segmentation then the decimated ones. However, it should be noted that in undecimated techniques at each decomposition level the size of the processed image is same as the size of the original image. This brings two conflicting results when compared to the multiresolution case. First one is a disadvantage for undecimated methods: in full rate case, at each step, images of the same size are filtered, this causes more computations than the multirate case in which at each level the size of the sub images is reduced. Second one can be considered as an advantage for undecimated methods: the length of filters used is always same at each level of decomposition, where for decimated case the length of the filter is reduced as the size of the image is reduced, that is filter coefficients are recalculated at each level and this causes computational complexity requiring efficient recursive mathematical manipulations as in DWT.

The experiments reveal that an average accuracy rate between the highest and lowest rates of different wavelet filters can be achieved using a single, simple yet effective symmetric, zore-phase, non-overlapping, 2-Channel QMF. WT based methods are tested using wavelets of different families. It is shown that they are only giving similar, and even in some cases significantly worse results than the DFT based method.

We have seen how various subband decomposition approaches yield different segmentation results for the same images. We demonstrate, in this thesis, that if wavelets either maximally sub sampled or over complete are replaced by a decimated

or an undecimated version of the DFT based subband decomposition using effective LP and HP filtering, then similar or even better segmentation results can be obtained.

From this study we can make two suggestions for the future research: First one may be the refinement of the ideal filters with a windowing, but that windowing should not violate the perfect reconstruction condition. The second one is on the design of the 2-D filters which are non-separable and can be directly applied to 2-D image. These filters are also directional and it is expected that these filters are capable of extracting directional properties such as edges those are neither horizontal nor vertical to be exactly located by the separable filters.

## REFERENCES


[1] **T. Lee and M. S. Lewicki**, "Unsupervised Image Classification, Segmentation, and Enhancement Using ICA Mixture Models," IEEE Trans. on Image Processing, Vol.11, No. 2, Feb. 2002, pp. 270-279

[2] **P. Brodatz**, Textures: A photographic album for artists and designers. New York: Dover, 1966.

[3] **Moffatt, D. L., J. D. Young, A. A.Ksienski, H. C. Lin, and C. M. Rhoods**, "Transient Response Characteristics in Identification and Imaging," IEEE Trans. Antennas and Propagation", Vol. 29, No. 2, 192–201, 1981.

[4] **Chen, K. M., D. P. Nyquist, E. J. Rothwell, L. L. Webb, and B. Drachman**, "Radar Target Discrimination by Convolution of Radar Returns with Extinction-Pulse and Single-Mode Extraction Signals," IEEE Trans. on Antennas and Propagation, Vol. 34, No. 7, 896– 904, 1986.

[5] **T.R. Reed and J.M.H. Du Buf**, "A review of Recent Texture Segmentation and Feature Segmantation Algorithms," CVGIP: Image Understanding, vol. 57, pp.359-372, May 1993. Also Available at http://www.ux.his.no/sigpro

[6] **M. Sezgin, I. Erer, O. K. Ersoy**, "A Novel DFT/RDFT Based Subband Representation for the Fusion of Remote Sensing Images", 2nd International Conference on Recent Advances in Space Technologies, 9-11 June 2005, Istanbul, Turkey.

[7] **M. Sezgin, I. Erer, O. K. Ersoy**, "DFT/RDFT Filter Bank Approach to Speckle Reduction in SAR Images.", IEEE Geoscience and Remote Sensing Symposium, IGAR SS 25-29 July, 2005, Seoul, Korea.

[8] **Ersoy, Okan K.**, Fourier-Related Transforms, Fast Algorithms, and Applications**,** Upper Saddle River, NJ : Prentice Hall PTR, c1997.

[9] **S. Sezen**, Application of Two-Dimensional Lattice Filter Structures to the Subband Decomposition Problem, M.Sc. Thesis, Bogazici University, 1995.

[10] **T. Randen, J. H. Husoy**, "Filtering for Texture Classification: A Comparative Study", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 21, No. 4, April 1999

[11] **S.G. Mallat**, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 11, pp.674-693, July 1989.

[12] **T. Randen and J.H. Husoy**, "Texture Segmentation Using Filters with Optimized Energy Separation," IEEE Trans. Image Processing, Vol. 8, No. 4, Apr. 1999.

[13] **Warin C., Punya T. And Yuttapong R.**, "Using Stationary Wavelet Transform in the Classification of Sar Images," available at: http://www.gisdevelopement.net/aars/acrs/1998/ps1/ ps1013pf.htm

[14] **M. Unser**, "Texture Classification and Segmentation Using Wavelet Frames", IEEE Trans. Image Processing, Vol. 4, pp. 1549-1560, Nov. 1995.

[15] **J.D. Johnson**, "A Filter Family Designed for Use in QMFB," Proc. Int'l Conf. Acoustic Speech, Signal Processing, Denver, Colo., pp.291-294, 1980.

[16] **H. Karakahya**, A Spectral-Special Classification Algorithm For Multispectral Remote Sensing Data, M.Sc. Thesis, Istanbul Technical University, June 2003.

[17] **Ryherd, S. and Woodcock, C**., 1996. "Combining Spectral and Texture Data in the Segmentation of Remotely Sensed Images," Photogrammetric Engineering and Remote Sensing, 62(2), 181-194

[18] **M. Tuceryan and A. K. Jain**. Texture Analysis. In C. H. Chen, L. F. Pau, and P. S. P. Wang, editors, Handbook of Pattern Recognition and Computer Vision, Chapter 2, pages 235_276. World Scientific, Singapore, 1993

[19] **R.Haralick**, "Statistical and Structural Approaches to Texture," Proc. of IEEE, vol. 67, no. 5, pp. 786-804, May 1979

[20] **A.Rosenfeld and L.Davis**, "Image Segmentation and Image Models," Proc. of IEEE, vol. 67, no. 5, pp. 764-772, May 1979.

[21] **Chrochiere R.E., S.A. Webber, J.L. Flanagan**, "Digital Coding of Speech in Subbands." Bell.Syst.Tech.J., Vol.55,pp1069-1085, October 1976.

[22] **Vetterli M**., "Multidimensional subband coding:some theory and algorithms," Signal Processing, Vol.6, no.3, pp 97-112,april 1984.

[23] **I. Celasun**, Design of Multidimensional Perfect Reconstruction Filter Banks with Compression Applications**,** Ph.D. Thesis, Bogazici University, 1996.

[24] **S. Zhou**, Y. V. Venkatesh, and C. C. Ko. "Texture Retrieval Using Tree-Structured Wavelet Transform," International Conference on Computer Vision, Pattern Recognition, and Image Processing (CVPRIP), March 2000

[25] **O. Rioul and P. Duhanel**, "Fast Algorithms for Discrete and Continuous Wavelet Transforms." IEEE Transactions on Information Theory, Special Issue on Wavelet Transforms and Multiresolution Signal Analysis, 38(2):569–586, 1992. I. Daubechies, S. Mallat, and A.S. Willsky, editors.

[26] **Akansu, Ali N., Richard A. Haddad**, Multiresolution signal decomposition: transforms, subbands, and wavelets San Diego : Academic Press, 2001.

[27] **T. Chang and C.C.J. Kuo**. "Texture Analysis and Classification with Tree Structured Wavelet Transform," IEEE Trans. on Image Processing, 2, 4:429-441, 1993

[28] **Duda, R.O., Hart, P.E., and Stork, D.G**., 2001. Pattern Classification John Wiley & Sons, New York.

[29] **Landgrebe, D.A.,** 2003. Signal Theory Methods in Multispectral Remote Sensing, John Wiley & Sons, New York

[30] **Russ, J.C.,** 1999. The Image Processing Handbook, CRC press, Boca Raton, FL

[31] **Soille, P. and Pesaresi, M**., 2002 "Advances in Mathematical Morphology Applied to Geoscience and Remote Sensing," IEEE Transactions on Geoscience and Remote Sensing, 40(9), pp 2042-2055

[32] **Song, J., and Delp, E.J**., 1992. "A Study of Generalized Morphological Filter," Circuits, Systems, and Signal Processing, 11(1),pp227-252.

[33] **Vaidyanathan, P.P**., "Quadrature Mirror Filter Banks, M-Band Extensions, and Perfect Reconstruction Techniques," IEEE Transactions on Acoustics, Speech, and Signal Processing Magazine, Vol. 4, No.3, pp. 4-20, July 1987

[34] **Vaidyanathan, P.P**., "Theory and Design of M-Channel Maximally Decimated Quadrature Mirror Filters with Arbitrary M, Having the Perfect Reconstruction Property," Transactions on Acoustics, Speech, and Signal Processing Magazine, Vol. ASSP-35, No.4, pp. 476-492, April 1987

[35] **Y.T. Chen**, Wavelet Basics, Kluver Academic Publishers, 1995

[36] **Papoulis**, A. Signal Analysis, Mc Graw Hill, New York, 1997

[37] **S.Mallat**, "Multifrequency Channel Decomposition of Images and Wavelet Models," IEEE Trans. Acoust. Speech, Signal Processing, vol 37, pp.14-38, Oct. 1991

[38] **Ali M. Reza**, From Fourier Transform to Wavelet Transform Basic concepts, white paper, October 27, 1999

[39] **Vaidyanathan, P.P. and Djokovic**, I., "Wavelet Transforms", in the Circuits and Filters Handbook. Edited by W.K. Chen, Chapter 5, pp 134-219, CRC Pres and IEEE Pres, Boca Rotan, Florida, 1995

[40] **Burrus, C.S., Gopinath RA and Guo H**., Introduction to Wavelets and Wavelet Transforms: a Primer. Prentice Hall Inc., Upper saddle River, New Jersey, 1998

[41] **Strang, G. And Nguyen, T.,** Wavelets and Filter Banks. Wellesley-Cambridge Pres, Wellesley, Massachusetts, 1996.

[42] **Coifman, R.R., D.L. Donoho**, "Translation invariant de-noising," Lecture Notes in Statistics, 103, pp. 125–150, 1995.

## BIOGRAPHY

Mehmet Enver Ergüven was born on May 10, 1980 in Diyarbakır. He graduated from Diyarbakır Cumhuriyet Fen Lisesi in 1999. He received his Bachelor of Science degree from Istanbul Technical University Electronics and Communication Engineering Department in 2003. He attended to Istanbul Technical University, Informatics Institute, Satellite Communication and Remote Sensing Graduate Program in 2003.