ROTATION AND SCALE INVARIANT TEXTURE CLASSIFICATION

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology
In
Electronics & Instrumentation Engineering

By

ANUBHAV AGARWAL (10407008)

&

MADHUKAR MANOHAR (10407033)

Department of Electronics & Communication Engineering

National Institute of Technology
Rourkela

2008

ROTATION AND SCALE INVARIANT TEXTURE CLASSIFICATION

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Bachelor of Technology
In
Electronics & Instrumentation Engineering

By

ANUBHAV AGARWAL (10407008)

&

MADHUKAR MANOHAR (10407033)

Under the Guidance of
Prof. G.S. Rath

Department of Electronics & Communication Engineering

National Institute of Technology
Rourkela

2008

NATIONAL INSTITUTE OF TECHNOLOGY
ROURKELA

CETRIFICATE

This is to certify that the Thesis Report entitled "***Rotation and Scale Invariant Texture Classification***" submitted by **Anubhav Agarwal (10407008) & Madhukar Manohar (10407033)** in partial fulfillment of the requirements for the award of Bachelor of Technology degree in Electronics and Instrumentation Engineering during session 2007-2008 at National Institute Of Technology, Rourkela (Deemed University) is an authentic work by them under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other university/institute for the award of any Degree of Diploma.

Prof. G.S. RATH
Dept. of Electronics & Communication Engineering
National Institute of Technology
Rourkela-769008
Email: gsrath@nitrkl.ac.in

Date:

CONTENTS

ACKNOWLEDGEMENT

We would like to articulate our deep gratitude to our project guide **Prof. G.S. Rath** who has always been our motivation for carrying out the project. It is our pleasure to refer Microsoft Word exclusive of which the compilation of this report would have been impossible. An assemblage of this nature could never have been attempted with out reference to and inspiration from the works of others whose details are mentioned in reference section. We acknowledge out indebtedness to all of them. Last but not the least, our sincere thanks to all of our friends who have patiently extended all sorts of help for accomplishing this undertaking**.**

ANUBHAV AGARWAL
MADHUKAR MANOHAR

ABSTRACT

Texture classification is very important in image analysis. Content based image retrieval, inspection of surfaces, object recognition by texture, document segmentation are few examples where texture classification plays a major role. Classification of texture images, especially those with different orientation and scale changes, is a challenging and important problem in image analysis and classification. This thesis proposes an effective scheme for rotation and scale invariant texture classification. The rotation and scale invariant feature extraction for a given image involves applying a log-polar transform to eliminate the rotation and scale effects, but at same time produce a row shifted log-polar image, which is then passed to an adaptive row shift invariant wavelet packet transform to eliminate the row shift effects. So, the output wavelet coefficients are rotation and scale invariant. The adaptive row shift invariant wavelet packet transform is quite efficient with only O (n*log n) complexity. The experimental results, based on different testing data sets for images from Brodatz album with different orientations and scales, show that the implemented classification scheme outperforms other texture classification methods, its overall accuracy rate for joint rotation and scale invariance being 87.09 percent.

Chapter 1

INTRODUCTION

1.1 Back ground

Image texture is an important surface characteristic used to identify and recognize objects. Texture is difficult to be defined. It may be informally defined as a structure composed of a large number of more or less ordered similar patterns or structures. Textures provide the idea about the perceived smoothness, coarseness or regularity of the surface. Texture has played an increasingly important role in diverse applications of image processing such as in computer vision, pattern recognition, remote sensing, industrial inspection and medical diagnosis.

Texture is the visual cue due to the repetition of image patterns which may be perceived as being directional or non-directional, smooth or rough, coarse or fine, regular or irregular, etc. The objective of the problem of texture representation is to reduce the amount of raw data presented by the image, while preserving the information needed for the task. In image processing texture analysis is aimed at three main issues: texture synthesis, segmentation and classification.

Texture synthesis is an alternative way to create textures. Synthetic textures can be made of any size without visual repetition as in original texture images. Potential applications of texture synthesis are image de-noising, compression, etc.

Texture segmentation is an important topic in image processing. It aims at segmenting a textured image into several regions without a priori knowing the textures. An effective and efficient texture segmentation method will be very useful in applications like the analysis of aerial images, biomedical images and seismic images as well as the automation of industrial inspections.

Texture classification involves deciding what texture category an observed image belongs to. In order to accomplish this, one needs to have an a prior knowledge of the classes to be recognized. Once this knowledge is available and the texture features are extracted, one then uses classification techniques in order to do the classification.

1.2 Objective

Many algorithms for texture classification are not rotation and scale invariant. The efficiency of a texture classification/segmentation algorithm can be increased by using a module for feature extraction followed by classification. This will be particularly useful for very large images such as those used for medical image processing, remote-sensing applications and large content based image retrieval systems. The objective of this thesis is to develop such a module for the rotation and scale invariant texture classification.

1.3 Thesis Contribution

This thesis addressed the problem of rotation and scale invariance in image analysis and classification. First I briefly reviewed the standard 2D wavelet packet decomposition techniques. Then, I define an algorithm to extract the rotation and scale invariant log-polar wavelet energy signatures for a given image. The feature extraction process involves applying a *log-polar transform* and an *adaptive row shift invariant wavelet packet transform* to obtain rotation and scale invariant wavelet coefficients. This feature extraction process is quite efficient with only O $(n * \log n)$ complexity (where n is the number pixels in the given image). Also, the construction of a feature vector using most dominant log-polar wavelet transform extracted from each sub band of wavelet coefficients, provides an effective and small number of features for rotation and scale invariant texture classification. The performance of implemented algorithm was tested by a number of experiments using a set of distinct natural textures selected from the Brodatz album.

The experimental results, based on different testing data sets for images with different orientations and scales, show that the implemented classification scheme using log-polar wavelet transform is quite robust to noise and very efficient. The overall accuracy of 87.09 percent for joint rotation and scale invariance was achieved with a vector of only 128 features.

1.4 Organization of Thesis

The remainder of the thesis is organized as follows. Chapter 2 gives a brief introduction about texture classification methods. It also presents the recent developments in rotation and scale invariant texture classification and discrete wavelet transform. Chapter 3 describes the method for rotation and scale invariant texture classification. It explains about log-polar transform and wavelet packet transform. Chapter 4 describes about the proposed algorithm for rotation and scale invariant texture classification. Chapter 5 gives the Simulation results for the implemented and proposed algorithm for rotation and scale invariant texture classification. Then we made a conclusion to our work.

Chapter 2

LITERATURE REVIEW

This chapter gives a review of existing literature about texture classification. A small overview about discrete wavelet transforms and wavelet families of image processing are also presented.

2.1 Texture Classification

A host of literature is available on texture analysis. Texture segmentation and classification methods can be broadly follow two approaches: Spatial-domain approach and Transform domain approach.

2.1.1 Spatial-domain Approach

The approach includes the following:

Structural texture analysis: This method considers texture as a composition of primitive elements arranged according to some placement rule. These primitives are called texels. Extracting the texels from the natural image is a difficult task. Therefore this method has limited applications.

Statistical methods: This is based on the various joint probabilities of gray values. Gray Level Co-occurrence Matrices (GLCM) estimate the second order statistics by counting the frequencies for all the pairs of gray values and all displacements in the input image. Haralick proposed several texture features that can be extracted from the co- occurrence matrices such as uniformity of energy, entropy, maximum probability, contrast, inverse difference moments, correlation and probability run lengths.

Model based methods include fitting of model like Markov random field, autoregressive, fractal and others. The estimated model parameters are used to segment and classify textures.

2.1.2 Transform-domain Approach

This approach analyses texture in various transform domains usually implemented through various filters/filter banks. T Randen and J. H Husoy gave an excellent review of the various filtering techniques for texture classification and compared the performance of the techniques. Filtering approach includes Laws mask, ring/wedge filters, dyadic Gabor filter banks, wavelet transforms which is explained in detail in the next section, quadrature mirror filters, DCT, eigen filters etc.

Laws was one of the pioneers of the filtering approach. He proposed nine 3X3 masks to accentuate the texture features. The response of each filter mask was used to extract the texture energies. Coggins and Jain have suggested seven dyadically spaced ring filters and four wedge-shaped orientation filters for feature extraction. More recent developments are based on Gabor filters and wavelets. Dyadic Gabor filter banks have been used to extract texture features. These filters give maximum resolution in both spatial and frequency domains and are highly desirable for texture analysis. There is also evidence that Gabor filters provide good models for the response profiles of many cortical cells in the human visual cortex. The Gabor filter is of the form of a 2-D Gaussian modulated complex sinusoidal in the spatial domain.

$$h(x,y) = g(x',y')e^{-2\pi j(Ux+Vy)} \tag{2.1}$$

Where (x′, y′) = (xcosφ + ysinφ, −xsinφ + ycosφ) are rotated co ordinates, and

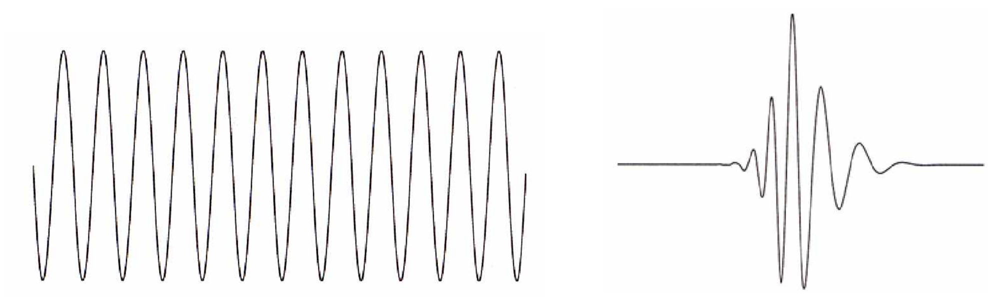
$$g(x,y) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{(\frac{x}{\lambda})^2 + y^2}{2\sigma^2}\right)} \tag{2.2}$$

Where λ defines the aspect ratio, σ the scale factor and (U, V) defines the position of the filter in the frequency domain. The scale factor is typically determined by the center frequency of the filter. A fixed set of filters is usually chosen to generate features for texture classification. These filters are centered at the required frequencies and orientations to obtain the optimum coverage of the frequency domain.

2.2 The Discrete Wavelet Transform

The transform of a signal is just another form of representing the signal. It does not change the information content present in the signal. The Wavelet Transform provides a time-frequency representation of the signal. It was developed to overcome the short coming of the Short Time Fourier Transform (STFT), which can also be used to analyze non-stationary signals. While STFT gives a constant resolution at all frequencies, the Wavelet Transform uses multi-resolution technique by which different frequencies are analyzed with different resolutions.

A wave is an oscillating function of time or space and is periodic. In contrast, wavelets are localized waves. They have their energy concentrated in time or space and are suited to analysis of transient signals. While Fourier Transform and STFT use waves to analyze signals, the Wavelet Transform uses wavelets of finite energy.



(a) (b)

Figure 2.1 Demonstrations of (a) a Wave and (b) a Wavelet.

The wavelet analysis is done similar to the STFT analysis. The signal to be analyzed is multiplied with a wavelet function just as it is multiplied with a window function in STFT, and then the transform is computed for each segment generated. However, unlike STFT, in Wavelet Transform, the width of the wavelet function changes with each spectral component. The Wavelet Transform, at high frequencies, gives good time resolution and poor frequency resolution, while at low frequencies; the Wavelet Transform gives good frequency resolution and poor time resolution.

2.2.1 The Continuous Wavelet Transform and the Wavelet Series

The Continuous Wavelet Transform (CWT) is provided by equation 2.3, where x (t) is the signal to be analyzed. Ψ (t) is the mother wavelet or the basis function. All the wavelet functions used in the transformation are derived from the mother wavelet through translation (shifting) and scaling (dilation or compression).

$$X_{WT}(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \cdot \psi^{*}\left(\frac{t-\tau}{s}\right) dt \qquad (2.3)$$

The mother wavelet used to generate all the basis functions is designed based on some desired characteristics associated with that function. The translation parameter **τ** relates to the location of the wavelet function as it is shifted through the signal. Thus, it corresponds to the time information in the Wavelet Transform. The scale parameter s is defined as |1/frequency| and corresponds to frequency information. Scaling either dilates (expands) or compresses a signal. Large scales (low frequencies) dilate the signal and provide detailed information hidden in the signal, while small scales (high frequencies) compress the signal and provide global information about the signal. Notice that the Wavelet Transform merely performs the convolution operation of the signal and the basis function. The above analysis becomes very useful as in most practical applications, high frequencies (low scales) do not last for a long duration, but instead, appear as short bursts, while low frequencies (high scales) usually last for entire duration of the signal.

The Wavelet Series is obtained by discretizing CWT. This aids in computation of CWT using computers and is obtained by sampling the time-scale plane. The sampling rate can be changed accordingly with scale change without violating the Nyquist criterion. Nyquist criterion states that, the minimum sampling rate that allows reconstruction of the original signal is 2ω radians, where ω is the highest frequency in the signal. Therefore, as the scale goes higher (lower frequencies), the sampling rate can be decreased thus reducing the number of computations.

2.2.2 The Discrete Wavelet Transform

The Wavelet Series is just a sampled version of CWT and its computation may consume significant amount of time and resources, depending on the resolution required. The Discrete Wavelet Transform (DWT), which is based on sub-band coding, is found to yield a fast computation of Wavelet Transform. It is easy to implement and reduces the computation time and resources required.

The foundations of DWT go back to 1976 when techniques to decompose discrete time signals were devised. Similar work was done in speech signal coding which was named as sub-band coding. In 1983, a technique similar to sub-band coding was developed which was named pyramidal coding. Later many improvements were made to these coding schemes which resulted in efficient multi-resolution analysis schemes.

In CWT, the signals are analyzed using a set of basis functions which relate to each other by simple scaling and translation. In the case of DWT, a time-scale representation of the digital signal is obtained using digital filtering techniques. The signal to be analyzed is passed through filters with different cutoff frequencies at different scales.

2.3 DWT and Filter Banks

2.3.1 Multi-Resolution Analysis using Filter Banks

Filters are one of the most widely used signal processing functions. Wavelets can be realized by iteration of filters with rescaling. The resolution of the signal, which is a measure of the amount of detailed information in the signal, is determined by the filtering operations, and the scale is determined by up sampling and down sampling (sub sampling) operations.

The DWT is computed by successive low pass and high pass filtering of the discrete time-domain signal as shown in figure 2.2. This is called the Mallat algorithm or Mallat-tree decomposition. Its significance is in the manner it connects the continuous- time multiresolution to discrete-time filters. In the figure, the signal is denoted by the sequence x[n], where n is an integer. The low pass filter is denoted by G_0 while the high pass filter is denoted by H_0. At each level, the high pass filter produces detailed information; d[n], while the low pass filter associated with scaling function produces coarse approximations, a[n].

At each decomposition level, the half band filters produce signals spanning only half the frequency band. This doubles the frequency resolution as the uncertainty in frequency is reduced by half. In accordance with Nyquist's rule if the original signal has a highest frequency of ω, which requires a sampling frequency of 2ω radians, then it now has

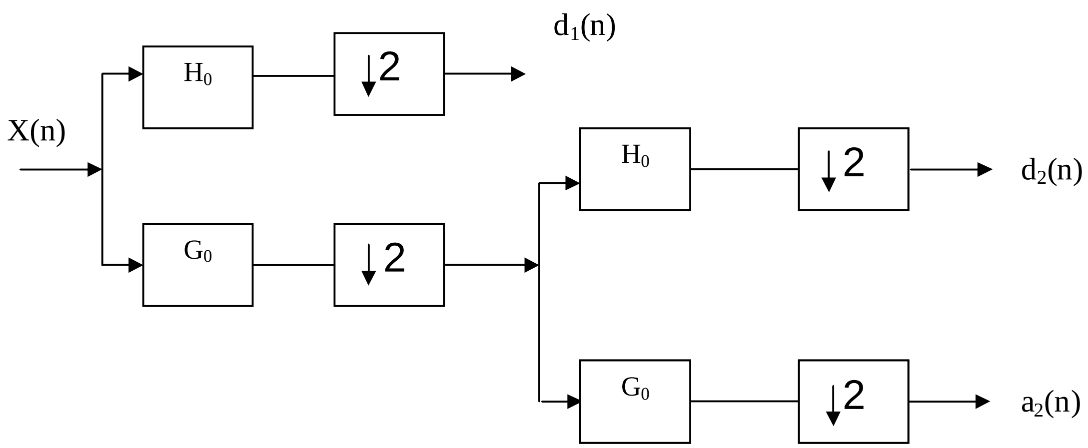


Figure 2.2 Two-level Wavelet Decomposition Tree.

a highest frequency of ω/2 radians. It can now be sampled at a frequency of ω radians thus discarding half the samples with no loss of information. This decimation by 2 halves the time resolution as the entire signal is now represented by only half the number of samples. Thus, while the half band low pass filtering removes half of the frequencies and thus halves the resolution, the decimation by 2 doubles the scale.

With this approach, the time resolution becomes arbitrarily good at high frequencies, while the frequency resolution becomes arbitrarily good at low frequencies. The filtering and decimation process is continued until the desired level is reached. The maximum number of levels depends on the length of the signal. The DWT of the original signal is then obtained by concatenating all the coefficients, a[n] and d[n], starting from the last level of decomposition.

Figure 2.3 shows the reconstruction of the original signal from the wavelet coefficients. Basically, the reconstruction is the reverse process of decomposition. The approximation and detail coefficients at every level are up sampled by two, passed through the low pass and high pass synthesis filters and then added. This process is

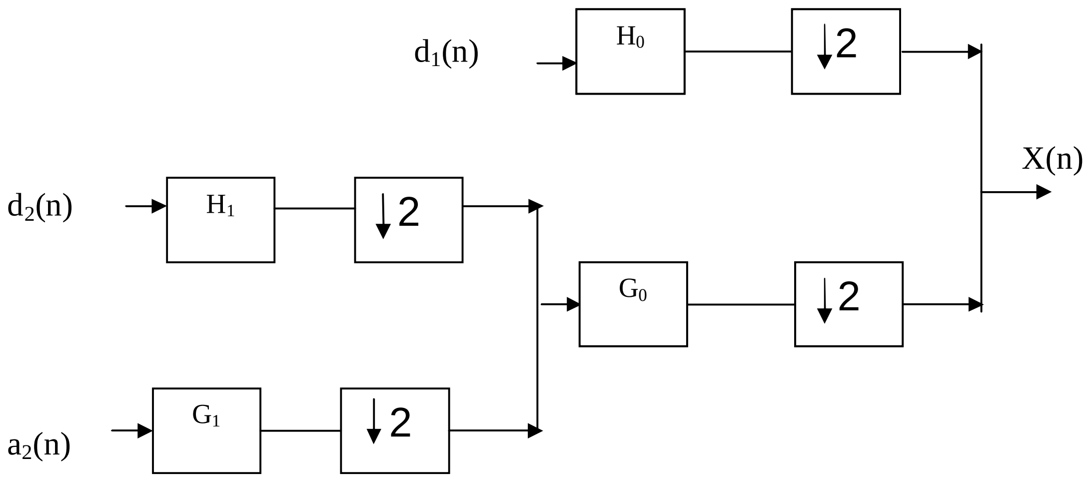


Figure 2.3 Two-level Wavelet Reconstruction Tree.

continued through the same number of levels as in the decomposition process to obtain the original signal. The Mallat algorithm works equally well if the analysis filters, G_0 and H_0, are exchanged with the synthesis filters, G_1 and H_1.

2.3.2 Conditions for Perfect Reconstruction

In most Wavelet Transform applications, it is required that the original signal be synthesized from the wavelet coefficients. To achieve perfect reconstruction the analysis and synthesis filters have to satisfy certain conditions. Let $G_0(z)$ and $G_1(z)$ be the low pass analysis and synthesis filters, respectively and $H_0(z)$ and $H_1(z)$ the high pass analysis and synthesis filters respectively. Then the filters have to satisfy the following two conditions as given:

$$G_0(-z)\, G_1(z) + H_0(-z).\, H_1(z) = 0 \qquad (2.4)$$

$$G_0(z)\, G_1(z) + H_0(z).\, H_1(z) = 2z^{-d} \qquad (2.5)$$

The first condition implies that the reconstruction is aliasing-free and the second condition implies that the amplitude distortion has amplitude of one. It can be observed that the perfect reconstruction condition does not change if we switch the analysis and synthesis filters.

There are a number of filters which satisfy these conditions. But not all of them give accurate Wavelet Transforms, especially when the filter coefficients are quantized. The accuracy of the Wavelet Transform can be determined after reconstruction by calculating the Signal to Noise Ratio (SNR) of the signal. Some applications like pattern recognition do not need reconstruction, and in such applications, the above conditions need not apply.

2.3.3 Classification of wavelets

We can classify wavelets into two classes: (a) orthogonal and (b) biorthogonal. Based on the application, either of them can be used.

(a) Features of Orthogonal Wavelet Filter Banks

The coefficients of orthogonal filters are real numbers. The filters are of the same length and are not symmetric. The low pass filter, G_0 and the high pass filter, H_0 are related to each other by

$$H_0(z) = z^{-N} G_0(-z^{-1}) \qquad (2.6)$$

The two filters are alternated flip of each other. The alternating flip automatically gives double-shift orthogonality between the low pass and high pass filters, i.e., the scalar product of the filters, for a shift by two is zero. i.e., $\sum G[k]\, H[k-2l] = 0$, where $k, l \in Z$. Filters that satisfy equation 2.6 are known as Conjugate Mirror Filters (CMF). Perfect reconstruction is possible with alternating flip.

Also, for perfect reconstruction, the synthesis filters are identical to the analysis filters except for a time reversal. Orthogonal filters offer a high number of vanishing moments. This property is useful in many signal and image processing applications. They have regular structure which leads to easy implementation and scalable architecture.

(b) Features of Biorthogonal Wavelet Filter Banks

In the case of the biorthogonal wavelet filters, the low pass and the high pass filters do not have the same length. The low pass filter is always symmetric, while the high pass filter could be either symmetric or anti-symmetric. The coefficients of the filters are either real numbers or integers.

For perfect reconstruction, biorthogonal filter bank has all odd length or all even length filters. The two analysis filters can be symmetric with odd length or one symmetric and the other anti-symmetric with even length. Also, the two sets of analysis and synthesis filters must be dual. The linear phase biorthogonal filters are the most popular filters for data compression applications.

2.4 Wavelet Families

There are a number of basis functions that can be used as the mother wavelet for Wavelet Transformation. Since the mother wavelet produces all wavelet functions used in the transformation through translation and scaling, it determines the characteristics of the resulting Wavelet Transform. Therefore, the details of the particular application should be taken into account and the appropriate mother wavelet should be chosen in order to use the Wavelet Transform effectively.

Figure 2.4 illustrates some of the commonly used wavelet functions. Haar wavelet is one of the oldest and simplest wavelet. Therefore, any discussion of wavelets starts with the Haar wavelet. Daubechies wavelets are the most popular wavelets. They represent the foundations of wavelet signal processing and are used in numerous applications. These are also called Maxflat wavelets as their frequency responses have maximum flatness at frequencies 0 and π. This is a very desirable property in some applications.

The Haar, Daubechies, Symlets and Coiflets are compactly supported orthogonal wavelets. These wavelets along with Meyer wavelets are capable of perfect reconstruction. The Meyer, Morlet and Mexican Hat wavelets are symmetric in shape. The wavelets are chosen based on their shape and their ability to analyze the signal in a particular application.

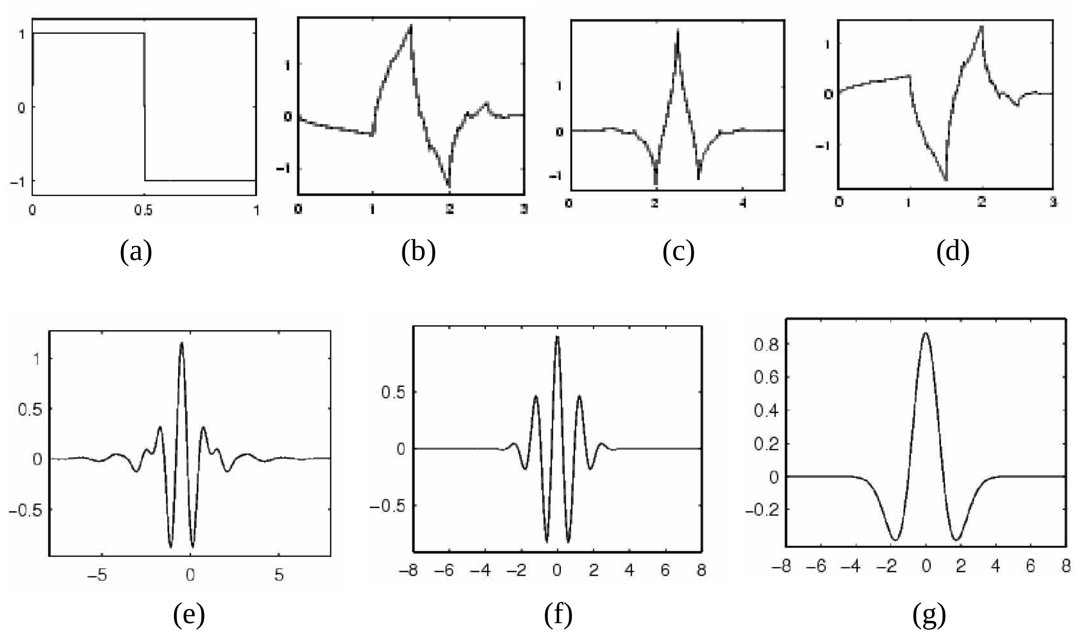


Figure 2.4 Wavelet families (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symlet2 (e) Meyer (f) Morlet (g) Mexican Hat.

Discrete Wavelet Transform (DWT) provides a tractable way of decomposing an image into different frequency subbands at different scales. The conventional wavelet transform decomposes a signal into a set of frequency channels that have narrower bandwidths in the lower frequency region. The transform is suitable for signals consisting primarily of smooth components so that their information is concentrated in the low frequency regions. So it cannot be applicable to all classes of texture images, more specifically images which have middle frequency components. Tianhorng Chang and C.C. Jay Kuo have proposed tree structured wavelet transform. The key difference between this algorithm and the traditional DWT algorithm is that the decomposition is no longer simply applied to the low frequency sub images recursively. Instead, it can be applied to the output of any frequency subband based on some energy criteria.

2.5 Rotation and Scale Invariant Texture Classification

All the above methods assumed that the texture images have same orientations and same scales. However, this assumption is not realistic. The images obtained from digital cameras are in different orientations and with different scales. Number of works can be found in the literature, which specifically address the problem of rotation and scale invariant texture recognition. F.S. Cohen, Z. Fan have proposed the 2D Gaussian Markov random field (GMRF) model with a likelihood function to incorporate and estimate rotation and scale parameters. Haley and Manjunath employ a complete space-frequency Gabor wavelet model for rotation-invariant texture classification with very promising results.

This method is as follows.

1) Map the image samples into Gabor space.

2) To facilitate discrimination between textures, transform the Gabor coefficients into micro features that contain local amplitude, frequency, phase, direction, and directionality characteristics. These micro features are spatially localized and do not characterize global attributes of textures.

3) Derive the macro features from the estimated selected parameters of the micro model. Then classification of texture samples is performed based on the rotation invariant components of the macro model. A potential disadvantage is that the outputs of Gabor filters are not mutually orthogonal. In addition, designing of Gabor filters are computationally intensive. R. Manthalkar, P.K. Biswas and B.N. Chatterji have proposed rotation and scale invariant texture features using Discrete Wavelet Packet Transform (DWPT). DWPT decomposes each decomposed level of the image into 4 sub bands LL, LH, HL and HH. For a decomposition of level d wavelet packet transform provide 4^d sub images. It is found that using the features from the HH channel of each level of decomposition can degrade the classification performance because these channels contain the majority of noise in the image. For these reasons HH channel information is not used for classification purpose.

4) Features are calculated from each sub image as follows.

$$f_n = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} |x(i,j)| \tag{2.7}$$

$$fnstd = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (|x(i,j)| - f_n)^2 \tag{2.8}$$

Where x (i, j) are the wavelet coefficients. M, N are the size of the subimage.
Rotation invariant features are obtained as follows.

$$F_n = 0.5 * [f_{nHL} + f_{nLH}] \tag{2.9}$$

$$F_{nstd} = 0.5 * [f_{nstdHL} + f_{nstdLH}] \tag{2.10}$$

For scale invariance, Discrete Fourier Transform (DFT) is applied on rotation invariant features. This operation removes the dependence of the feature values on scale.

Chi-Man Pun and Moon-Chuen Lee have proposed log-polar wavelet energy signatures for rotation and scale invariant texture classification. Rotation and scale invariant feature can be obtained by applying the wavelet packet transform on log-polar transform of the input image and its one row circular shift. Then energy signatures are computed from the sub image outputs of DWPT. Then the feature vector is computed from these energy signatures for rotation and scale invariant texture classification.

Chapter 3

METHOD FOR ROTATION AND SCALE INVARIANT TEXTURE CLASSIFICATION

3.1 Introduction

As mentioned in the earlier chapter most of the texture classification methods are not rotation and scale invariant. Rotation and scale invariant features are obtained by applying the wavelet packet transform on log-polar transform of the input image and its one row circular shift as shown in Figure. 3.1. The block diagram of the method is shown in Figure. 3.1. The various operations involved in the method are explained below.

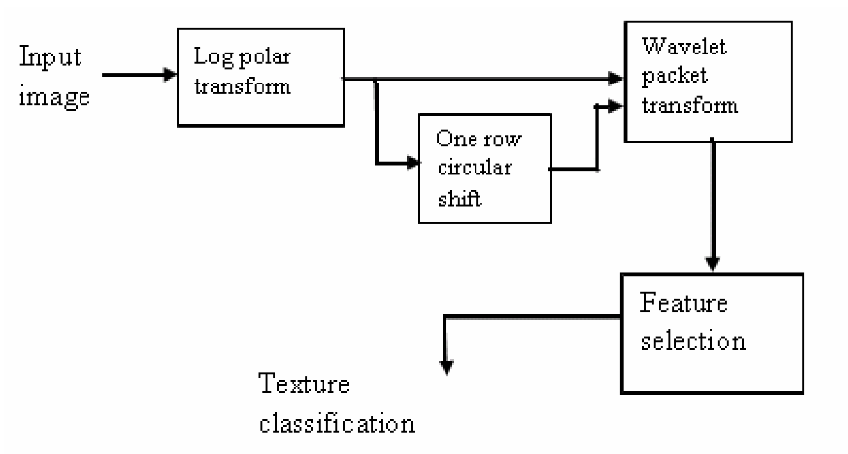


Figure 3.1: Block diagram explaining rotation and scale invariant feature extraction

3.2 Discrete Wavelet Packet Transform

The standard 2D discrete wavelet packet transform (DWPT) is a generalization of 2D discrete wavelet transform (DWT) that offers a richer range of possibilities for image analysis. In 2D-DWT analysis, an image is split into an approximation and three detail images. The approximation image is then itself split into a second-level approximation and detail images, and the process is recursively repeated. So, there are (n + 1) possible ways to decompose or encode the image for an n-level decomposition. In 2D-DWPT analysis, the three details images as well as the approximation image can also be split. So, there are 4^n different ways to encode the image, which provide a better tool for image analysis.

The standard 2D-DWPT can be described by a pair of quadrature mirror filters (QMF) H and G. The filter H is a low-pass filter with a finite impulse response denoted by h (n). And, the high pass G with a finite impulse response is defined by:

$$G(n) = (-1)^{n} \cdot h(1-n) \qquad \text{for all n.} \qquad (3.1)$$

The low-pass filter is assumed to satisfy the following conditions for orthonormal representation:

$$\sum_{n} h(n)h(n+2j) = 0 \qquad \text{For all } j \neq 0. \qquad (3.2)$$

$$\sum_{n} \left|h(n)\right|^{2} = 1 \qquad (3.3)$$

$$\sum_{n} h(n)g(n+2j) = 0, \qquad \text{For all j.} \qquad (3.4)$$

2D discrete wavelet transform (DWT) decomposes the image into four frequency bands *LL, LH, HL* and *HH*. The *LL* band is decompose into second level *LL, LH, HL* and *HH* frequency bands and the process is recursively repeated. The standard 2D discrete wavelet packet transform (DWPT) is a generalization of 2D discrete wavelet transform (DWT). In 2D-DWPT analysis all frequency bands (*LL, LH, HL and HH*) decompose to next de-composition levels. 2D-DWPT mathematical formulas are defined as follows.

$$W_{4k,(i,j)}^{p+1}(m,n) = \sum_{m} \sum_{n} h(m)h(n)W_{k}^{p}(m+2i, n+2j) \qquad (3.5)$$

$$W_{4k+1,(i,j)}^{p+1}(m,n) = \sum_{m} \sum_{n} h(m)g(n)W_{k}^{p}(m+2i, n+2j) \qquad (3.6)$$

$$W_{4k+2,(i,j)}^{p+1}(m,n) = \sum_{m} \sum_{n} g(m)h(n)W_{k}^{p}(m+2i, n+2j) \qquad (3.7)$$

$$W_{4k+3,(i,j)}^{p+1}(m,n) = \sum_{m} \sum_{n} g(m)g(n)W_{k}^{p}(m+2i, n+2j) \qquad (3.8)$$

Where $W_0^0(i, j)$ = x (i , j) is given by the intensity levels of the image x.

j = Level of 2D DWPT;

i = Sub image at j^{th} level;

K = Filter length;

g (m) = Impulse responses of the low-pass filter G (Z);

h (n) = Impulse responses of the high-pass filter H (z);

Wj = Image at the jth level of DWPT with W^0 as input image;

2D-DWPT can be implemented by a pair of quadrature mirror filters (QMF) lowpass filter G (z) and high pass filter H (z). The low pass filter impulse response g (n) and the high pass filter impulse response h (n) are related by using equation (3.1). This decomposition algorithm is illustrated by the block diagram shown in Figure. 3.2. Each decomposition comprises of two stages. Stage 1 performs horizontal filtering, and stage 2 performs vertical filtering.

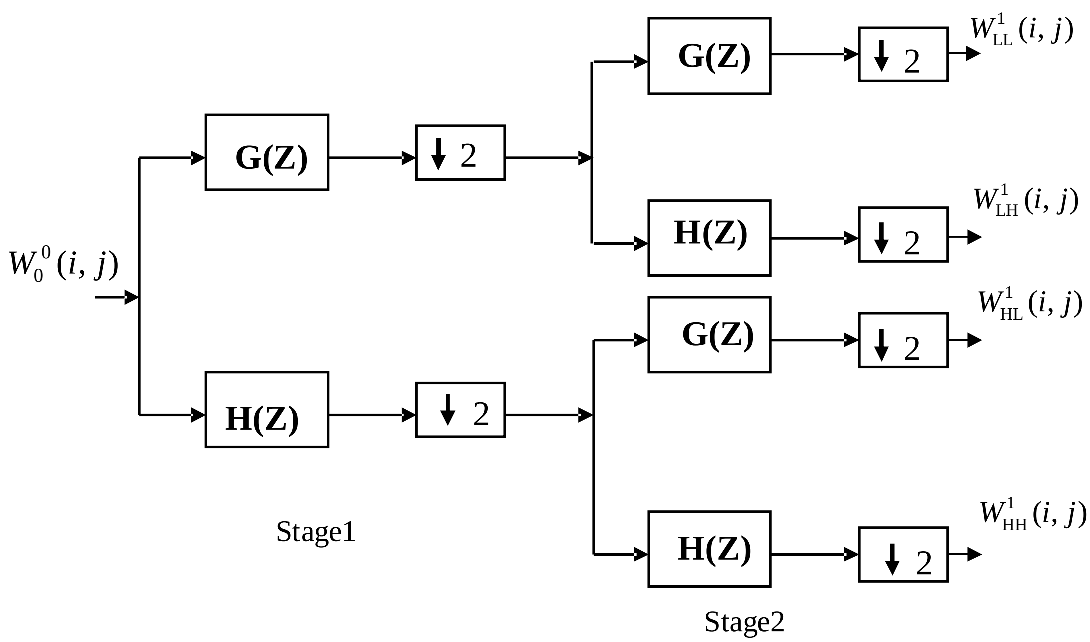
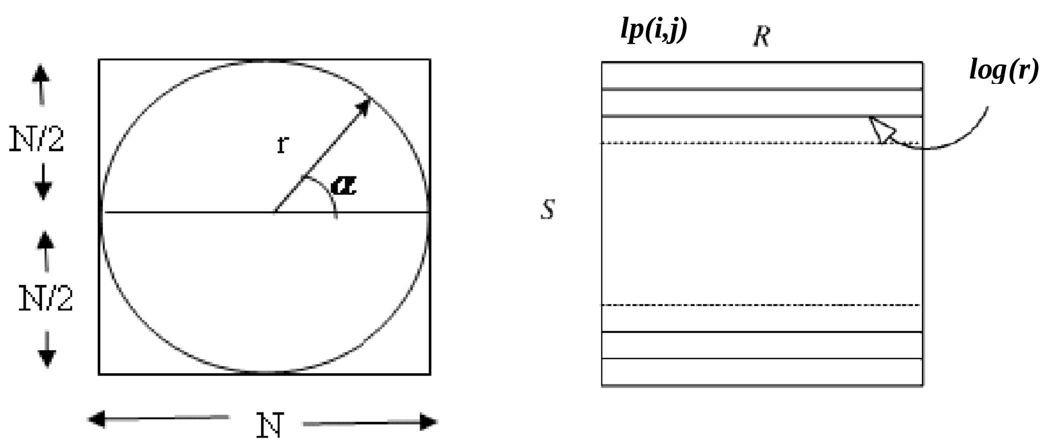


Figure 3.2: First level 2D Wavelet Packet Decomposition

3.3 Log-Polar Transform

Log-polar transform converts the rotation and scale variations in an image into row shifted images. Log-polar transform involves two steps. In the first step, the image is divided into S × N/2 polar grids where S is the number of points along a circle and N/2 is the maximum radius of the circle as shown in Figure 3.3. The polar image p *(α ,r)* is given by

$$p(\alpha, r) = f\left(\left\lfloor \frac{N}{2} \right\rfloor + \left\lfloor r\cos\left(\frac{2\pi\alpha}{S}\right)\right\rfloor, \left\lfloor \frac{N}{2} \right\rfloor - \left\lfloor r\sin\left(\frac{2\pi\alpha}{S}\right)\right\rfloor\right) \tag{3.9}$$



(a) (b)

Figure 3.3: Log-polar transform of *NXN* image (f (x, y)) into *SXR* log-polar image (lp (i, j)) by first (a) using radius as scanline for sampling N times the circle to produce a polar form P (α ,r), and (b) applying quantization on the logarithm of all radii to produce the log-polar image

Where

α = Angle,

r = Radius,

r = 1 to N/2

And α = 1 to S.

In the second step, logarithm functions are applied to all radii values in the polar form and their outputs are then quantized into R bins. Hence, an S x R log-polar image for the given N x N image is produced (as shown in Fig. 3.3b).

The procedure can be formally defined as follows:

$$lp(i, j) = p\left(i, \left\lfloor \frac{\log_2(j+2)}{\log_2(R+2)} \times \left\lfloor \frac{N}{2} \right\rfloor \right\rfloor \right) \tag{3.10}$$

For i=0……. S-1 and j=0……. (R-1).

As shown in Fig. 3.4, the log-polar images of a texture image with different rotation angles and scales seem having only row shifts when compared with the log-polar image of the original texture. The log-polar transform is also quite efficient with only O(n) computational complexity (where n is the number pixels in the given image).

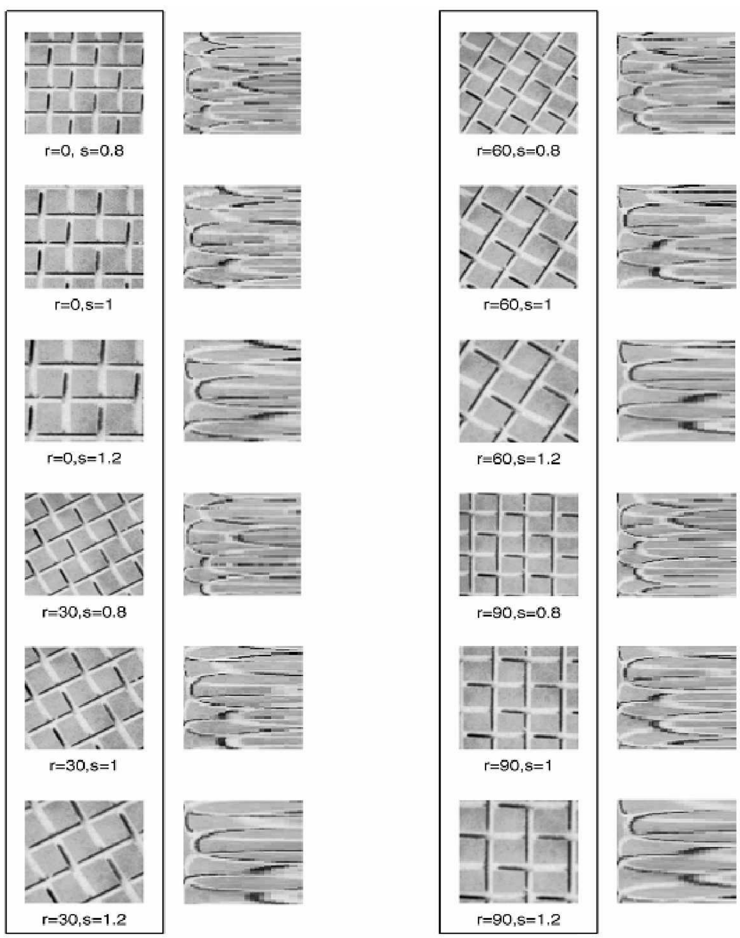


Figure 3.4: A sample texture (D1) from the Brodatz album in different rotation angles (r in degrees) and scales (s) and their corresponding log-polar images

In this section, we define an algorithm to extract the rotation and scale invariant feature for a given image, which can be obtained by applying a log-polar transform on the image, followed by adaptive row shift invariant wavelet packet transform (as shown in Fig. 3.5). The procedure can be formally defined as follows:

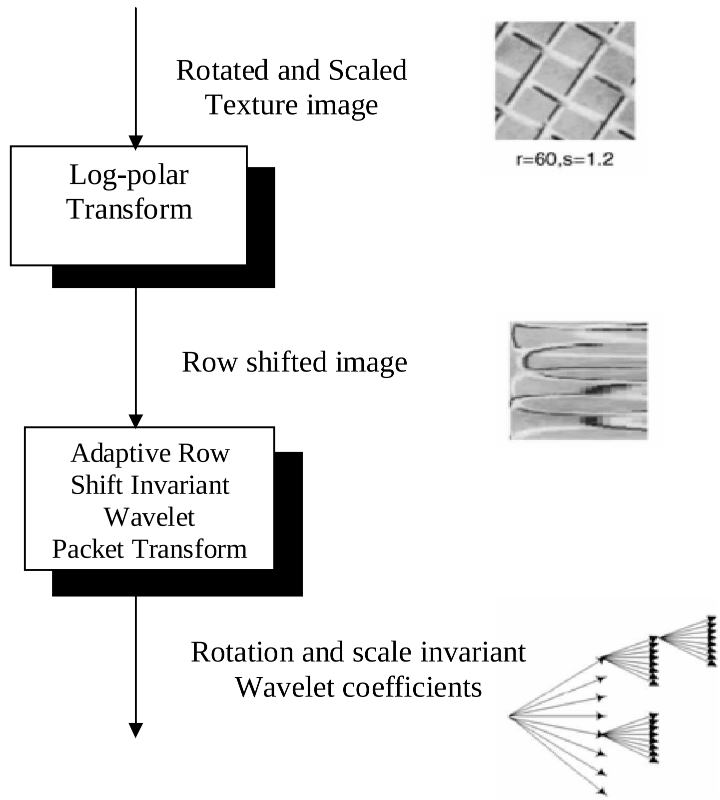


Figure 3.5: Decomposition of rotation and scale invariant wavelet coefficients from a rotated and scaled texture image.

3.3 Row Shift Invariant Wavelet Packet Transform

After applying the log-polar transform operation, a rotated and scaled image would be converted into a corresponding log-polar image which is rotation invariant and nearly scale invariant. However, any orientation changes would cause a row shifting in the log-polar image. Simple wavelet packet decomposition of row shifted log-polar image may not be much help for rotation and scale invariant texture classification.

Many shift-invariant wavelet decomposition algorithms have been proposed and are shift invariant in both rows and columns. These algorithms generate more redundant wavelet coefficients which are not suitable for the row-shifted output image produced by the log-polar transform. The row shift problem produced by the log-polar transform can be eliminated by redundant set of wavelet packet coefficients for one additional row circular shift of the log-polar image in each level.

$W_{i,LL,0}^{j+1}$, $W_{i,LH,0}^{j+1}$, $W_{i,HL,0}^{j+1}$, $W_{i,HH,0}^{j+1}$ are the output sub-images of the wavelet packet transform of the input image.

$W_{i,LL,1}^{j+1}$, $W_{i,LH,1}^{j+1}$, $W_{i,HL,1}^{j+1}$, $W_{i,HH,1}^{j+1}$ are the output sub-images of the wavelet packet transform of one row circular shift of the input image. This algorithm is illustrated in Figure.3.4. The mathematical formulas are defined as follows.

$$W_{\mathrm{i,LL,0}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n h(m)h(n)W_{i,0}^{j}(m+2i, n+2j) \tag{3.11}$$

$$W_{\mathrm{i,LH,0}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n h(m)g(n)W_{i,0}^{j}(m+2i, n+2j) \tag{3.12}$$

$$W_{\mathrm{i,HL,0}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n g(m)h(n)W_{i,0}^{j}(m+2i, n+2j) \tag{3.13}$$

$$W_{\mathrm{i,HH,0}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n g(m)g(n)W_{i,0}^{j}(m+2i, n+2j) \tag{3.14}$$

Where m = [N/2^{p+1}]-1, n= [M/2^{p+1}] and $W_0^0(i, j)$= x (i, j) is given by the gray levels of the image x.

Since we just keep one out of two rows, these coefficients appear the same $W_{\mathrm{i}}^{\mathrm{j}}(m,n)$ if is circularly shifted by 0, 2, 4…2^n rows. In order to have row shift invariance, we need to compute another four periodic images each with one row shift:

$$W_{\mathrm{i,LL,1}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n h(m)h(n)W_{i,1}^{j}(m+2i, n+2j) \tag{3.15}$$

$$W_{\mathrm{i,LH,1}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n h(m)g(n)W_{i,1k}^{j}(m+2i, n+2j) \tag{3.16}$$

$$W_{\mathrm{i,HL,1}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n g(m)h(n)W_{i,1}^{j}(m+2i, n+2j) \tag{3.17}$$

$$W_{\mathrm{i,HH,1}}^{\mathrm{j+1}}(m,n) = \sum_m \sum_n g(m)g(n)W_{i,1k}^{j}(m+2i, n+2j) \qquad (3.18)$$

In a similar manner, these coefficients appear the same if $W_{\mathrm{i}}^{\mathrm{j}}(m,n)$ is circularly shifted by 1,3, 5,…….. 2^n +1 rows, respectively.

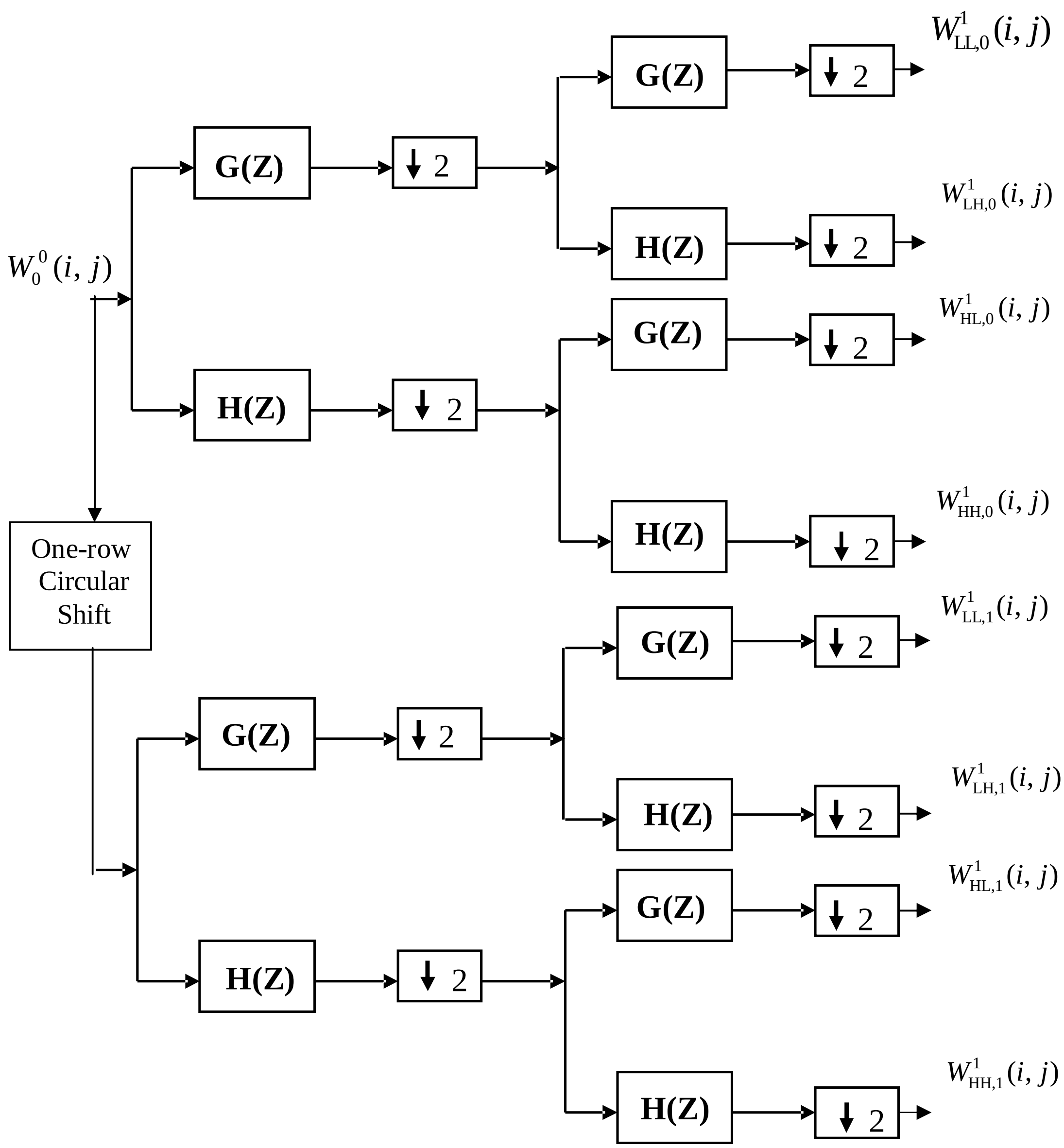


Figure 3.6: Oct-tree DWPT Decomposition

3.4 Extraction of Rotation and Scale Invariant Image

Log-polar transform eliminates the rotation and scale variations in an image. But at the same time it produces row shifted log-polar images. This problem is eliminated by applying the wavelet packet transform on the log-polar image and its one row shift in each level as shown in Figure.3.6.

Chapter 4

METHOD FOR TEXTURE CLASSIFICATION

4.1. Introduction:

As digital images become more widely used, digital image analysis must find more tools to work on them. Texture analysis is a huge challenge nowadays, since simple images may be considered as a mosaic of textures separated by some boundaries. That is why both texture retrieval and classification, combined with image segmentation, may be very powerful in image analysis. Texture retrieval (i.e. to find the N most similar textures to a query texture among a large set of data textures) can be used by internet applications in the general context of Content-Based Image Retrieval, whereas texture classification (i.e. among N classes of textures, to select the most probable one where the query texture could lie) has a more local use, since a "class of texture" has a loose sense, depending on the application.

In texture classification, we shall only consider the k-nearest neighbor classifier, Mahalanobis classifier, since it provides an efficient and robust scheme. Then in both texture classification and retrieval, we need a function measuring similarity between the query image and the images of the database. This function should measure a distance between some features of the textures. Therefore the method consists in two main steps: Feature extraction (FE) and Similarity Measurement (SM).

Comparing textures requires a definition of a texture. It can be defined as a homogeneous and coherent field of an image. But this is not a satisfactory definition, since it is vague and not general at the same time. Actually there does not exist a proper one. That is why in this report we shall consider gray-level images corresponding intuitively to homogeneous and coherent textures only. Two models of textures are to be considered: the periodic textures such as tiles and fabric, which can be studied by frequency analysis, and random textures such as grass and metal, which can be analyzed by statistical descriptions. However these two models correspond to extreme cases which do not match with real natural images. Thus we shall try to fuse the two approaches thanks to the wavelet transform.

The general method of FE and SM requires the features to be small-sized (much smaller than the image), requires the features to correctly characterize the texture, and requires the similarity measure to be precise and small when the features are close and big otherwise. It can also be asked to the features to be translation and rotation invariant, in order to regard two textures as equivalent, one deriving from the other by translation or rotation.

4.2. Gray-level statistics Methods

4.2.1 Feature Extraction

With the row shifted log-polar image obtained from the log-polar transform as the input to the adaptive row shift invariant wavelet packet transform, the row shift problem of the log-polar image is properly solved. So, the generated wavelet coefficients are rotation invariant and nearly scale invariant now. However, the large number of wavelet coefficients is not suitable for robust texture classification.

4.2.2 Similarity Measurement

After having extracted features, our next task is to find a similarity measure d such that ***d (x, vi)*** is small if and only if $\boldsymbol{x}$ and $\boldsymbol{v_i}$ are close. The simplest similarity measure is the Euclidean distance:

$$d^2(x(i), v(j)) = \sum_{k=1}^{k} \left|(x_k(i) - v_k(j))\right|^2 \qquad (4.1)$$

It is relevant to introduce a normalized Euclidean distance is also known as modified Mahalanobis classifier defined by:

$$d^2(x(i), v(j)) = \sum_{k=1}^{K} \frac{\left| x_k(i) - v_k(j) \right|^2}{\operatorname{var} x_k} \qquad (4.2)$$

Where $\boldsymbol{var x_k}$ is the empirical variance of $\boldsymbol{x_k}$ over the data base, i.e.

$$\operatorname{var} x_k = \sum_{m=1}^{M} \left| x_k(i) - m_k \right| \qquad (4.3)$$

Where $\boldsymbol{m_k}$ is the mean given by

$$m_k = \frac{1}{M} \sum_{m=1}^{M} x_k(i) \qquad (4.4)$$

The reason for introducing this distance is of statistical matter. The $\boldsymbol{(x_k(i))_i}$ are considered as M realizations of a random variable $\boldsymbol{x_k}$, such that the $\boldsymbol{x_k}$ are independent. Then $\boldsymbol{var x_k}$ is only the squared empirical standard deviation of $\boldsymbol{x_k}$. When the $\boldsymbol{x_k}$ are not independent, one prefers considering the Mahalanobis distance, which is more specific to the classification purpose. If $\boldsymbol{C}$ is a class of textures and $\boldsymbol{m_C}$ is the mean signature of class $\boldsymbol{C}$, the Mahalanobis distance is given by:

$$d^2(x_i, C) = (x_i - m_C)^t \, \Sigma^{-1} (x_i - m_C) \qquad (4.5)$$

Where $\boldsymbol{\Sigma}$ is the empirical covariance matrix of $\boldsymbol{x}$ on class $\boldsymbol{C}$. Note that if the features are independent, the covariance matrix is diagonal with diagonal elements $\boldsymbol{var x_k}$ we come back to the normalized Euclidean distance. Anyway we shall always consider that the $\boldsymbol{x_k}$ are independent, and then in this section we shall always use the normalized Euclidean distance.

The texture classification algorithm is explained below.

4.3 Texture Classification Algorithm

The steps are as follows:

Step 1. Mean, μ of the class is calculated.

Step 2. Covariance of the class is then calculated.

Step 3. After that Eigen values and corresponding Eigen vectors are calculated.

Step 4. All eigen vectors are arranged in descending order.

Step 5. Ten numbers of highest eigen vectors are selected which is represented as matrix $\mathbf{A}$.

Step 6. Calculate $\mu_f = A^t . \mu$ for each class.

Step 7. An unknown sample is selected and represented as $\mathbf{X}$**.**

Step 8. Calculate $X_{fi} = A_i^t . (X - \mu_i)$.

Step 9. Corresponding distance d_i between X_{fi} and μ_{fi} is then calculated.

Step 10. Unknown is classified to the original class if d_i is minimum.

Chapter 5

SIMULATION RESULTS

For the implementation of the above algorithms (chapter3 and chapter4) in MATLAB following database was created. Natural texture images were taken form Brodatz's album which is shown below.

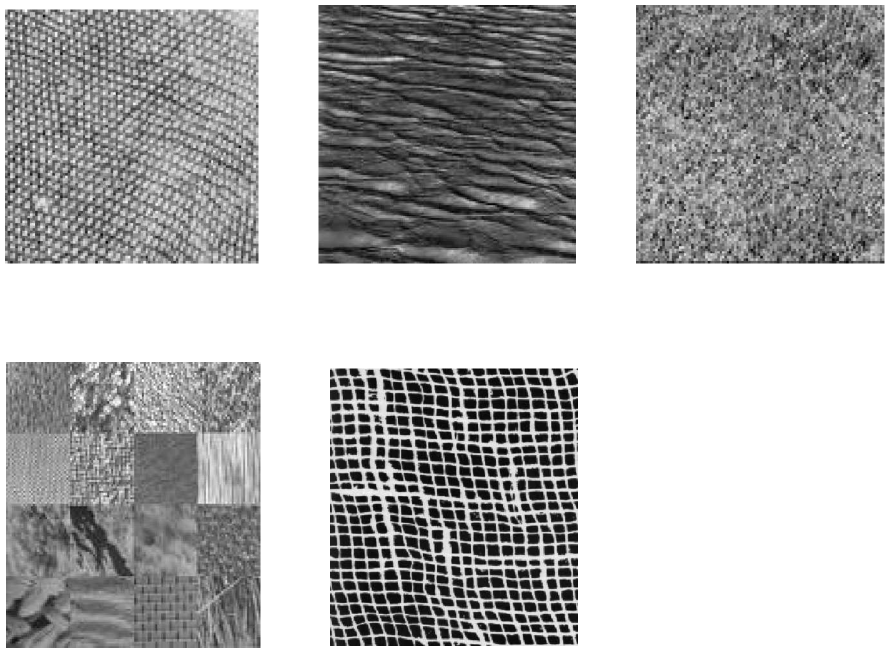


Figure 5.1 Four classes of textures from the Brodatz album. **ROW 1:** Image 1, Image 2, Image 3. **ROW 2:** Image 4, Image 5

The experiments were carried out with the objectives:

1) To investigate the texture classification performance based on the implemented log-polar wavelet feature.

2) To investigate the texture classification performance of the implemented method on texture images with rotation and scale changes.

3) To compare the proposed method with other texture classification method and demonstrate the noise robustness of the proposed method.

The effectiveness of the implemented method for rotation and scale invariant texture feature extraction was tested using natural texture images from Brodatzs texture album. The result for texture Image 5 (Figure. 5.2) is presented here. The Image 5 texture image with a rotation angle 30^0 (Figure. 5.3) was read row-wise in MATLAB and input image was applied to log-polar transform followed by row shift invariant wavelet packet transform. Figure. 5.4 show the MATLAB simulation results of the log-polar transform. Figure 5.5 show the three level wavelet packet decomposition of the log-polar image.

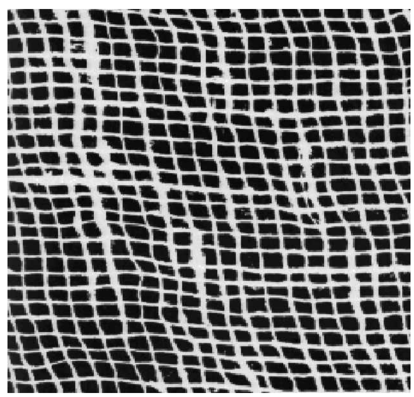


Figure 5.2. Image 5 texture from the Brodatz album.

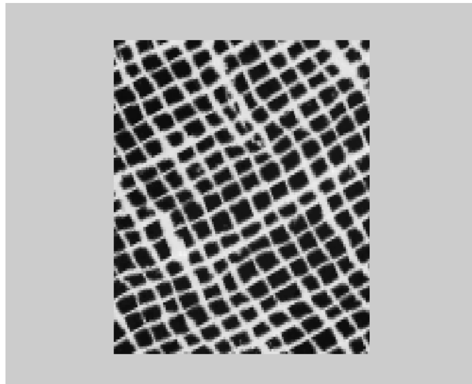


Figure 5.3 Image 5 texture image with 30^0 rotation angle.

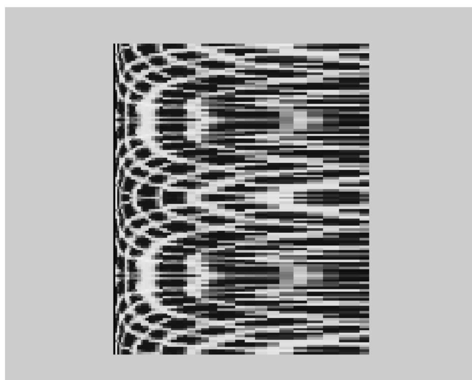


Figure.5.4 Rotation invariant log-polar image of Image 5

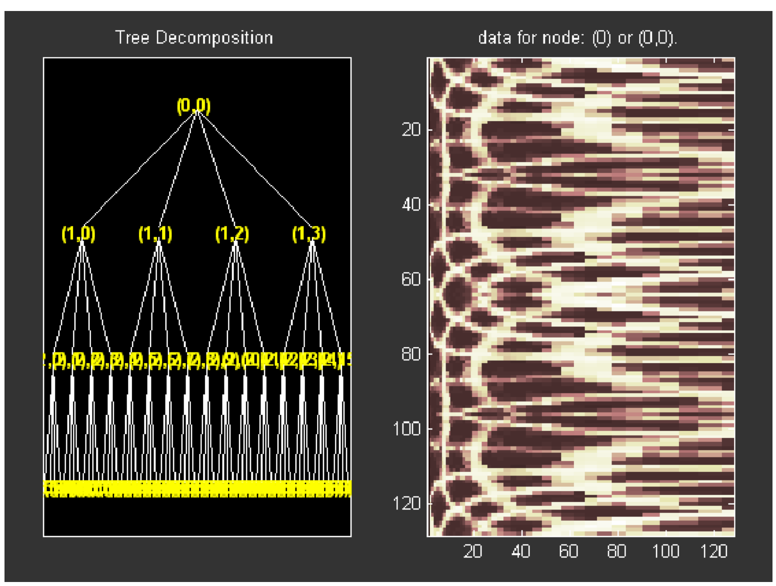


Figure 5.5 Input texture Image 5 (Log-polared) of DWPT.

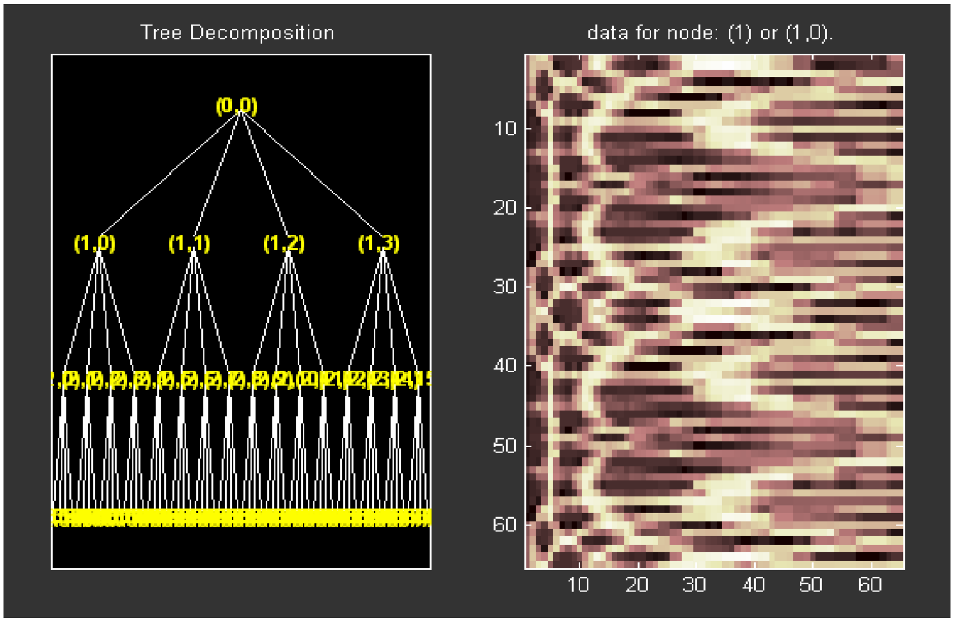


Figure 5.6 MATLAB Simulation for 1 Level DWPT LL Band texture Image 5.

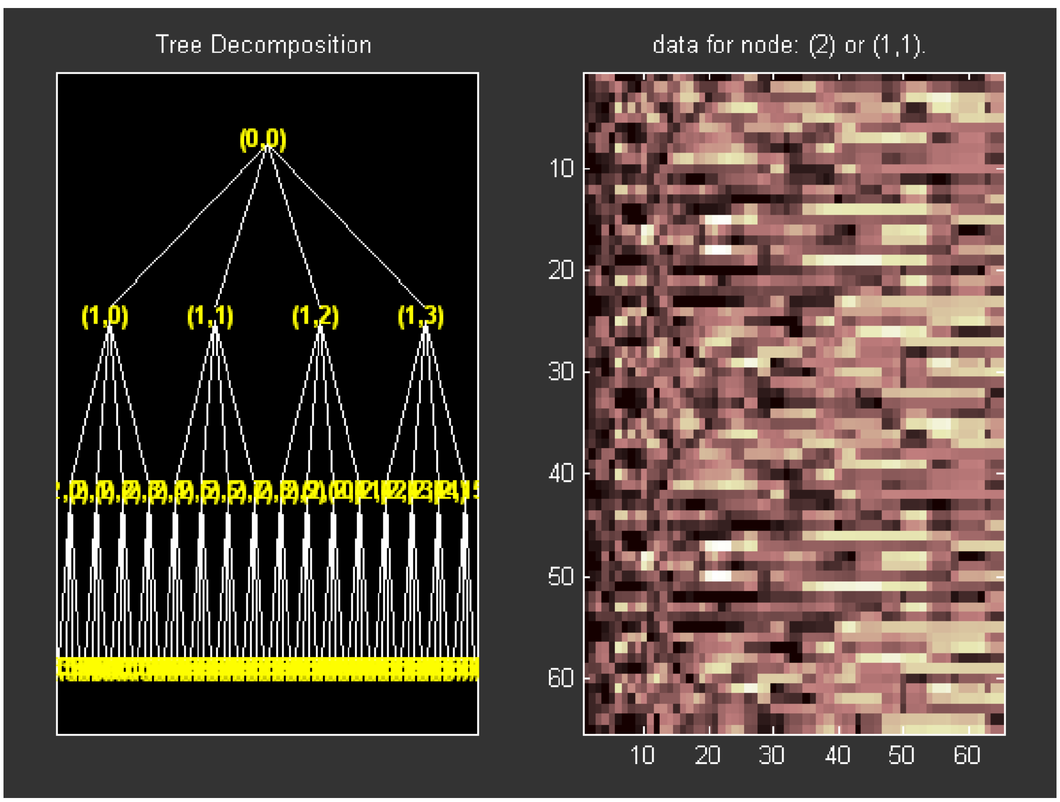


Figure 5.7 MATLAB Simulation for 1 Level DWPT LH Band texture Image 5.

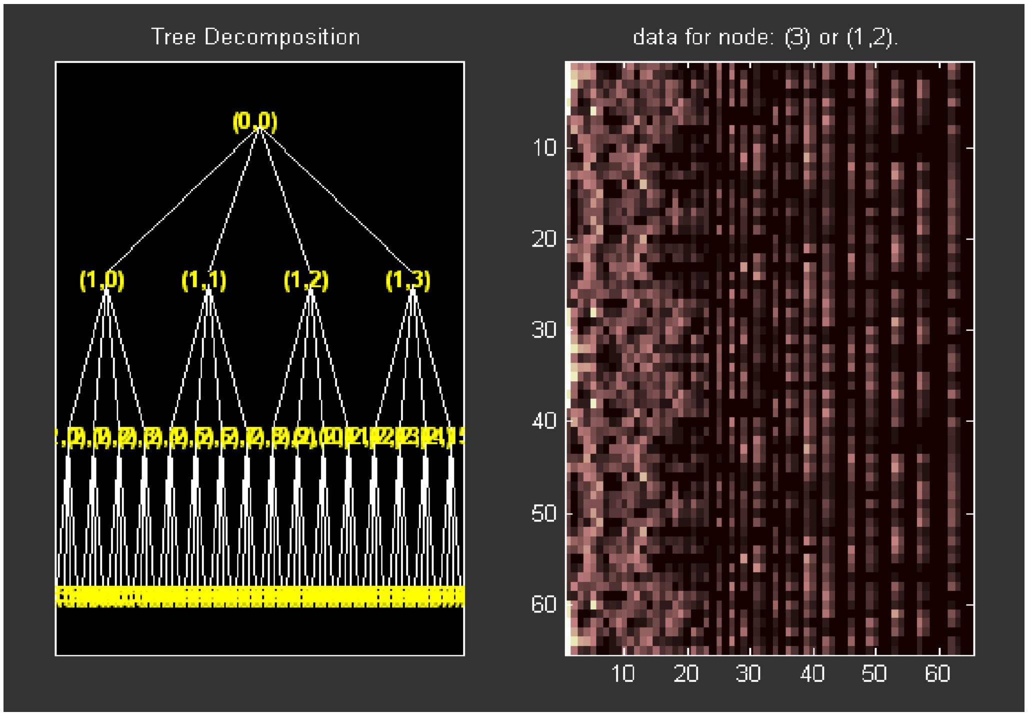


Figure 5.8 MATLAB Simulation for 1 Level DWPT HL Band texture Image 5

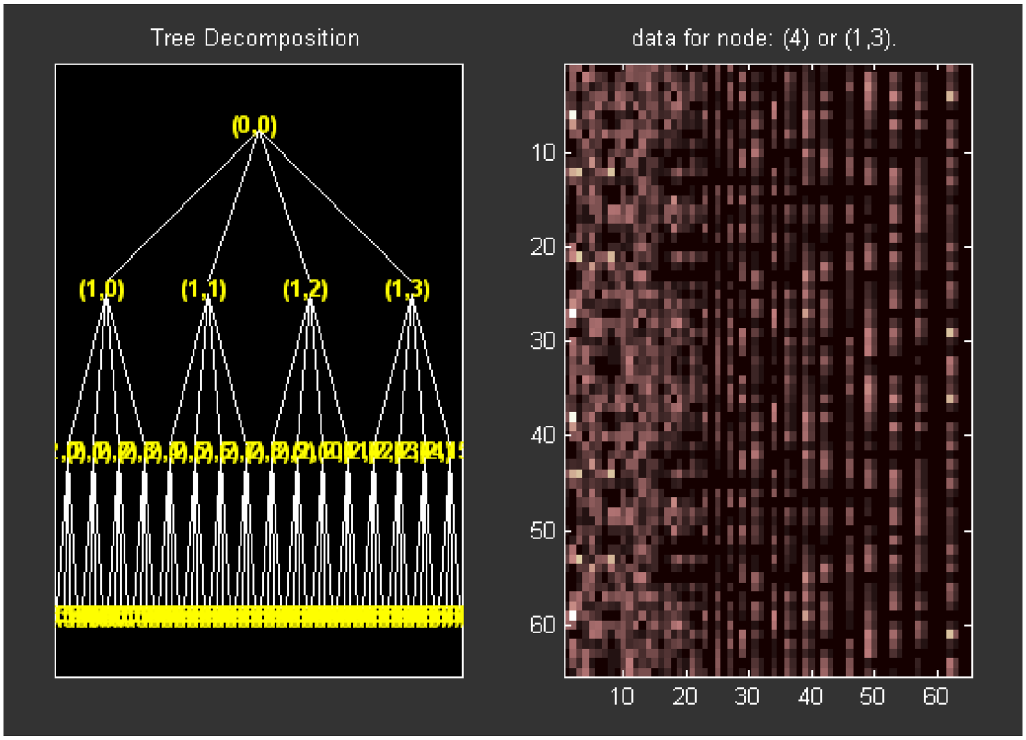


Figure 5.9 MATLAB Simulation for 1 Level DWPT HH Band texture Image 5.

Table: Classification Rates

Texture ID	Total No. of Samples	Error % age	Classification Rate %age
Image 1	10	15.10	84.90
Image 2	10	7.5	92.5
Image 3	10	9.2	90.8
Image 4	10	19.8	80.2
Image 5	10	12.95	87.05
Overall	50	12.91	87.09

From the table it is seen that with variations the proposed architecture gives 87.09 % classification with scale and rotation variations.

.

Chapter 6

CONCLUSION

This chapter gives a summary of the work presented in this thesis.

6.1 Summary

This thesis addressed the problem of rotation and scale invariance in image analysis and classification and implemented an effective algorithm for rotation and scale invariant texture classification. First we briefly reviewed the standard 2D wavelet packet decomposition techniques. Then, we defined an algorithm to extract the rotation and scale invariant log-polar features for a given image. The feature extraction process involves applying a *log-polar transform* and an *adaptive row shift invariant wavelet packet transform* to obtain rotation and scale invariant wavelet coefficients. This feature extraction process is quite efficient with only O $(n * \cdot log n)$ complexity (where n is the number pixels in the given image). The performance of the implemented algorithm was tested by a number of experiments using the 5 distinct natural textures selected from the Brodatz album.

The experimental results, based on different images with different orientations and scales, show that the implemented classification scheme is quite robust to noise. The overall accuracy of **87.09** percent for rotation and scale invariance was achieved with a vector of only 128 features.

BIBLIOGRAPHY

[1] R. C. Gonzalez, R. E.Woods, *Digital Image Processing*, Pearson Education 2002.

[2] M. Tuceryan, and A.K. Jain, "Texture analysis," *The Handbook of pattern Recognition and Computer vision*, pp. 235-276, World Scientific, 1993.

[3] Chi-Man Pun and Moon-Chuen Lee, "Log-Polar Wavelet Energy Signatures for Rotation and Scale Invariant Texture Classification," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 25, pp. 590-603, May 2003.

[4] R. Haralick, K. Shanmugam and I. Dinstein, "Texture Features for Image Classification," *IEEE Trans. on Systems, Man, Cybernetics*, vol. 3, pp. 610-621, Nov. 1973.

[5] R.L. Kashyap and A. Khotanzed, "A Model-Based Method for Rotation Invariant Texture Classification," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, pp. 472-481, July 1986.