Chapter 1

Introduction

1.1 Motivation and background

Motivation behind the texture study comes from the viewpoint of computer vision and graphics as well as its huge application in the area of automatic visual inspection.

Generally objects are sensed through spatial filtering rather than higher level symbolic grouping processes [1]. Human visual system decomposes any picture into filtered images of various orientation and frequency similar to cell in visual cortex of macaque monkey [2, 3]. Julsez [4] showed that of two textures having the same 2^{nd} order statistics i.e.,., probability of finding a pair of pixel separated by a fixed distance and an orientation can not be discriminated preattentively. So the perception of texture is very much related with human vision. Two questions serve motivation for psychophysicists. What are the visual processes related to texture perception of human being? What image properties are essential for human texture perception?

Textural methods have been utilized in a variety of application domains. In some of the mature or established domain (such as remote sensing) texture already has played a major role, while in other disciplines (such as surface inspection for quality assurance) new applications of texture are coming up. Some of the fields where texture plays a major role are:

Remote Sensing.: Texture is extensively used in land usage classification where homogeneous regions with different types of terrains (such as wheat, bodies of water, urban regions, etc.) need to be identified.

Medical Image Analysis.: The applications involve the automatic extraction of features from the image which are then used for a variety of classification tasks, such as distinguishing normal tissue from abnormal tissue.

Surface Inspection. The applications include defect detection in image of textiles and automated inspection of carpet wear and automobile paints, metal surface etc.

Document Processing. Texture is used as a tool to segment document images to identify regions of interest.

Image and Video Retrieval. Texture is one of the main features used for measuring similarity between multimedia data.

1.1.1 Definition of texture

Though texture is one of the intuitive concepts, there is no universally accepted definition of texture till date [5]. But the researchers are agreed on two points [6]

1. Significant variation in gray levels between neighboring pixels indicates the existence of non-homogeneity at the limit of resolution.
2. Texture exhibits homogenous property at a particular scale, greater than the resolution of the image.

From the technical aspect texture is related to stochastic and regularity defined over a given area. More definitions are available in a review by Tuceryan and Jain [7]. Human being uses the concept of texture describing properties like directionality, smoothness, coarseness and regularity of a region.

Texture can also be described by extracting and defining its primitive or texels and placement rule of texels or structure. In strong texture, texel is well defined and structure is regular. But in weak or random texture, texel is difficult to define and correlation between primitives is also low. The size of the texel is large for coarse texture but it is small for fine texture. So texture measures are dependent on the size of the observation and also on the resolution of the input image.

Again dearth of suitable definition of texture points at complexity of modeling spatial distribution of gray levels. Two dimensional histograms and co-occurrence matrix have been widely used to capture two dimensional distribution of intensities.

Texture study can be classified into different branches like texture classification, texture segmentation, texture synthesis, recognizing 3-D shape from texture, etc. Identifying textured image on the basis of texture properties is known as texture classification. In an image consisting of multiple textures boundary between different textures or between different homogenous regions can be found out by the technique called as texture segmentation. The goal of texture synthesis is to obtain a surface closer to real look. The 'shape from texture' or 'shape from X' deals with extraction of 3-D shape information from various cues like shading, stereo and texture [8].

1.2 Literature review

Image segmentation may be defined as a process which partitions a digital image into disjoint regions [9]. The disjoint regions are characterized by textural properties in case of texture based image segmentation or texture segmentation. A region is a connected set of pixels. The segmentation can be performed using either of two ways: supervised or unsupervised. In a supervised segmentation, the identity and location of some of the object types or regions in the image are known *a priori* through a combination of field work, analysis of maps, personal experience etc. Now the analyst attempts to map regions of the remainder of the image depending upon these training sites composed of regions known *a priori*. In an unsupervised segmentation, the class identities of regions within a scene are not generally known *a priori*. The computer is required to group the pixels into different classes according to some criteria. Then the onus lies on the analyst to label these clusters. Texture segmentation consists of two steps namely feature extraction and classification. Feature extraction is usually done in the following ways,

1. Statistical method
2. Signal processing method
3. Geometrical method
4. Model based method

1.2.1 Statistical method

Any image consists of gray level at each pixel. After calculating feature from a local region about the pixel considered, statistics from this distribution of features is often used to characterize a texture. Depending upon the number of pixel considered within the local region, the statistical methods are classified into first order (one pixel), second order (two pixels) and higher order (three or more pixels) statistics.

First order statistics

Let us conceive image as a function $f(i, j)$ of two dimensional spatial coordinates i, j where $i = 0, 1......., M-1$ and $j = 0, 1......., N-1$. The function $f(i, j)$ can take discrete values $k = 0, 1........., G-1$ where G is the total number of gray levels. Histogram of an image often describes first order statistical information about an image since occurrence of gray value of a single pixel is considered. Now if the number of occurrences in histogram is divided by the image size (i.e., $M \times N$), it may represent approximately probability density (i.e., $p(k)$) of occurrence of the intensity levels. Shape as well as a number of measures including central moments calculated from this histogram characterizing a texture [10] are given below,

Mean $$\mu = \sum_{k=0}^{G-1} kp(k)$$

Variance $$\sigma^2 = \sum_{k=0}^{G-1} (k-\mu)^2 \, p(k)$$

Skewness $$\mu_3 = \sigma^{-3} \sum_{k=0}^{G-1} (k-\mu)^3 \, p(k)$$

Kurtosis $$\mu_4 = \sigma^{-4} \sum_{k=0}^{G-1} (k-\mu)^4 \, p(k) - 3$$

Energy $$E = \sum_{k=0}^{G-1} \left[p(k)\right]^2$$

Entropy $$H = -\sum_{k=0}^{G-1} p(k) \log_2 \left[p(k)\right]$$

Skewness, kurotosis, energy and entropy represent symmetricity, flatness, uniformity and variability of the histogram respectively. Lam and Li [11] observed that mean and variance of images normalized against both mean and variance [i.e.,., $\mu = 0$, $\sigma = 1$], give better texture characterizing property than actual mean and variance of original image. Apart from this Lowitz [12] proposed module and state of histogram as texture features. The textural features derived from first order statistics have the advantage of simplicity and easiness. Though Julsez [4] found that if two textures have differences in first order statistics, they can be discriminated preattentively. But two textures having same second order statistics can not be discriminated preattentively. Second order statistics relates statistics of pixel pair or joint probability distribution of two pixels. Gray level co-occurrence matrix (GLCM), gray level run length (GLRL) and gray level difference (GLD) are generally invoked to captivate second order statistics.

Gray level co-occurrence matrix method (GLCM)

Gray level co-occurrence matrix defines the frequency of occurrence of pixel-pair having a particular displacement and an orientation. The basic idea is to estimate second order joint conditional probability density function, $f(i, j \mid d, \theta)$. $f(i, j \mid d, \theta)$ indicates probability of movement from gray level i to gray level j, given the intersample distance d and the direction given by the angle θ. The all estimated $f(i, j \mid d, \theta)$ are put into a matrix called as co-occurrence matrix. Memory saving is 50% while using symmetrical co-occurrence matrix accumulating symmetrical elements of the co-occurrence matrix along first diagonal. Though Haralick [13] defined 14 features defined from co-occurrence matrix, four of these 14 features namely angular second moment (ASM), contrast (Con), correlation (Cor) and entropy (Ent) have been widely used [14, 15, 16]. Generally small displacement d values of co-occurrence matrix are concentrated near the diagonal whereas values will diverge for higher value of d. Using a suitable value of d for estimating co-occurrence matrix is cumbersome as small d suits finer texture and large d suits coarse texture. Optimal results have been obtained for smaller values of d [17, 18]. Also quantization of gray levels in image saves computational time for calculating GLCM but looses some textural information. For image size of 32×32 eight gray levels of the image were reported as sufficient for textural content by Ohanian and Dubes [15].

Gray level difference method (GLD)

Let us take $h\delta(i, j) = \left| h(i, j) - h(i + M, j + N) \right|$, when image gray level function is $h(i, j)$ for any displacement $\delta = (M, N)$ where M , N are integers. Texture feature can be extracted from the probability distribution function $g'(k,\delta) = P(h\delta(i,j) = k)$. With four possible versions of δ with intersample spacing distance as $(0,d)$, $(-d,d)$, $(d,0)$ and $(-d,-d)$, $g'(k,\delta)$ is referred to as the gray level difference density functions. Five texture features namely contrast, angular second moment, entropy, mean and inverse difference moment have been extracted for each one of these function [19].

Gray level run length method (GLRL)

The basis of the gray level run length method (GLRLM) lies in computing the number of gray level runs of various lengths. A gray level run is a set of linear adjacent image points having the same gray level value. The length of the run is the number of image points within the run. Each element $r(i,k\,|\,\theta)$ of the gray level run length matrix specifies the number of times an image contains a run of length i for gray level k in the angle θ direction. Each of these matrixes usually has been calculated for several values of θ. Features like short run emphasis, long run emphasis, gray level distribution, run length distribution and run percentages from these matrixes have been considered for texture discrimination [19]. GLCM has been found to be superior to GLRL and GLD in terms of performance [19].

Power spectrum

Features derived from power spectrum are given as following

1. Annular ring sample geometry which measures coarseness of the texture.
2. Wedge sampling geometry, which contains directional information.
3. Parallel-slit sampling geometry.

Eklundh [20] reported that phase information is not a useful texture measure. In this method sample power spectrum is computed as below,

$\phi(u,v) = \hat{F}(u,v)\hat{F}^*(u,v) = \left|\hat{F}(u,v)\right|^2$, where $\wedge$ denotes complex function, $*$ denotes complex conjugate, ϕ denotes sample power spectrum, F denotes Fourier transform of the image. Some details about the features extracted from power spectrum are included in the Section 2.2.

Autocorrelation

From structural point of view, texels are placed in accordance to some predefined rules. Now for coarse texture the size of the texels is large whereas for fine texture it is small. Autocorrelation function of an image $I(x, y)$ is given by

$$\rho(x, y) = \frac{\sum_{u=0}^{M} \sum_{v=0}^{N} I(u,v) I(u + x, v + y)}{\sum_{u=0}^{M} \sum_{v=0}^{N} I^2(u,v)}$$

where image size is $M \times N$. When primitives are relatively large autocorrelation function will drop of slowly but for fine texture the process is very first. For regular texture there are periodical ups and downs like peaks and valleys in autocorrelation function. People have defined this feature as a measure of coarseness. Kaizer [21] defined $1/e$ drop off distance to be a measure of texture coarseness.

1.2.2 Signal processing method

Human recognition of textured images has similarity with signal processing based features. Signal processing methods can be of two types namely spatial domain filters and Fourier domain filters.

Fourier domain filters

In the Fourier domain of filters the discrete Fourier transform of an image $f(i, j)$ is denoted by $F(u,v) = \frac{1}{N^2} \sum_{i,j=0}^{N-1} f(i, j) e^{-2\pi\sqrt{-1}(ui + jv)}$ for $0 \le u, v \le N - 1$

where $N \times N$ is the size of the image. Fourier power spectrum is $|F|^2 = FF^*$ [where $*$ denotes the complex conjugate]. For finer texture, high $|F|^2$ concentrated near origin whereas for coarser texture it is distributed. Later on wedge and ring filter have been introduce to capture property in the transformed domain. Ring filter is described as following, $\phi_{r_1 r_2} = \sum_{\substack{r_1^2 \le u^2+v^2 \le r_2^2 \\ 0 \le u,v \le N-1}} |F(u,v)|^2$

for various values of inner and outer ring radii i.e.,., r_1 and r_2. Features based on wedge shaped samples are of the form

$$\phi_{\theta_1 \theta_2} = \sum_{\substack{\theta_1 \le \tan^{-1}(v/u) \le \theta_2 \\ 0 \le u,v \le N-1}} |F(u,v)|^2$$

for various values of span of the wedge i.e.,., θ_1 and θ_2.

$\phi_{r_1 r_2}$ is a measure of coarseness whereas $\phi_{\theta_1 \theta_2}$ gives directionality present in an image.

Power spectrum based methods have following disadvantages:

1. Discrete Fourier transform assumes an image to be repetitive though in fact it is only semi-repetitive. Spurious values in the transform domain due to discontinuity at the extremities of the image result poor performance.
2. Because of the fact that image is modeled as sum of sinusoids in Fourier transform, in many cases, the performance of features based on transforms is worse than statistical features in extraction of textural property in terms of classification accuracy.

Apart from Fourier transform two dimensional pseudo-Wigner distribution (PWD) was used for texture segmentation. The drawbacks of this method are high dimensionality of feature space as its space to space/spatial frequency representation and erroneous results for complex textures. Local frequency analysis of human visual system could be replicated by continuous spectrum of frequency analysers using two dimensional Gabor functions [22]. Human vision system also interprets an image with a set of filters varying in center frequency and orientation like Gabor filters. Though Turner [23] showed meaningful texture features by using Gabor power spectrum the main disadvantage lies in large number of Gabor filters required to capture small change of center frequency and orientation. Therefore efforts have been made to reduce the number of filters required through some combination or coding. Jain and Farrokhinia [24] proposed texture features as average absolute deviation from mean in Gaussian weighted window moving upon nonlinearity added even symmetric Gabor filtered images. Number of filter selection is done when summed power of the selected filters becomes a given fraction (95%) of the total power of the all filters. Addition of pixel coordinate location with features extracted from even symmetric Gabor filter improves the result to an extent. Bovik et al. [25] implemented a peak searching scheme in global power spectrum to select the number of filters. Thereafter feature selection was done taking the maximum filter response at any position of the image. Du Buff [26] reduced the number of higher order Gabor features by application of least-squares approximation by considering local Gabor power spectrum as two dimensional array in log-polar frequency coordinates. Real and complex moments are also compared for describing the shape of the local power spectrum [27, 28]. The disadvantage of these higher order Gabor features is that elongated regions are often detected on the texture boundaries due to implementation of region based segmentation algorithm [29] for sensitivity of some parameters to the mixture of two power spectrum at the boundary. Local spectral dissimilarity estimate was invoked for a simple polar-complex boundary detection [30]. Apart from all these types on features of Gabor power spectrum, the phase spectrum carries important texture information. But this is paralysed by classical two dimensional phase unwrapping problem. Although phase gradient [25] is one of the tools to circumvent these problems. Quality of local phase features is reduced by noise and jitters in structural elements of the texture. The disadvantages of the Gabor filters are as following,

1. Redundancy of features derived from Gabor filter at different scales or channels occurs due to non-orthogonality of Gabor filters [31]
2. Resolutions of analysis i.e. width of analysis window is the same at all locations in the time frequency plane.

Above discrepancies of Gabor filter pave the way for wavelet which is basically bandpass filter. Advantages of using wavelet filters are given as below,

1. Due to orthogonality no interference is produced between different wavelet filters.

2. Unlike Fourier transform, wavelet as well as Gabor transform is space to space and spatial frequency transformation.

3. Width of analysis window is varied depending upon frequency content of the signal that is why it is used in multiresolution analysis.

Convolution of a signal $f(x)$ with a family of basis functions $\psi_{2^s,\tau}(x)$; wavelet decomposition of a signal is done through the following formula,

$$\left\langle f(x), \psi_{2^s,\tau}(x) \right\rangle = \int_{-\infty}^{\infty} f(x)\psi_{2^s,\tau}(x)dx \qquad (1.1)$$

where s, τ are translation and dilation parameters.

Pair of a low pass filter and a high pass filter is used to calculate wavelet coefficients through pyramidal algorithm of wavelet decomposition as given in Eq. (1.1). In spite of using wavelet functions quadrature mirror filters (QMF) can be used to implement wavelet transform as shown in Fig. 1.1.

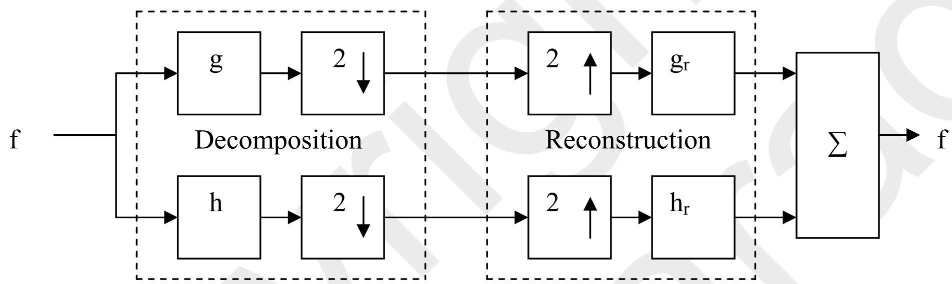

Fig. 1.1. Illustration of wavelet based signal decomposition and reconstruction; g, g_r – lowpass filter; h, h_r – highpass filter; (2↓) –downsampling (decimation by 2); (2↑) –upsampling

For two dimensional images, the wavelet decomposition is achieved with separable filtering along the rows and along the columns of an image like Fig. 1.2 [32].

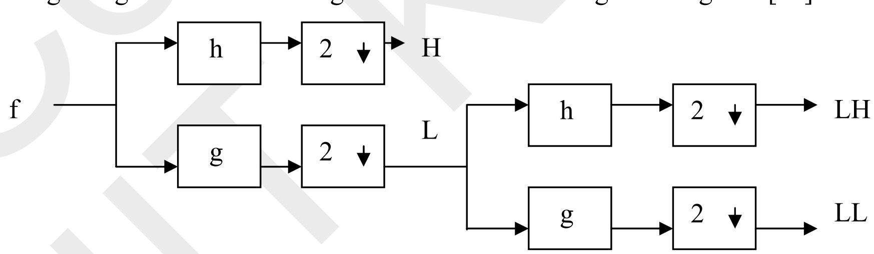

Fig. 1.2. Signal analysis of two levels of dyadic wavelet decomposition; L- Low pass or Approximation components; H- High pass or Detail components.

Wavelet analysis therefore means decomposition of image in a set of independent, spatially oriented frequency channels. HH, HL, LH and LL sub images represent high frequencies in horizontal and vertical directions, horizontal high frequencies, vertical high frequencies and lowest frequencies in both direction respectively. In subsequent scale of analysis the image LL undergoes the decomposition using the same g and h filters. Always the lowest frequency component is placed in the upper left corner of the image. Every stage of analysis produces four sub images whose size is reduced twice as that of in the previous scale. The number of levels of decomposition is guided by the size of the image. Among the different types of wavelet functions symmetric wavelet has got edge over non symmetric wavelet filter due to linear phase properties of symmetric filters. Porter and Canagarajah [33] divided an image into textured and smooth regions looking at whether the ratio of energy levels of different combination of wavelet coefficients of the image is greater than, equal to or lesser than a certain threshold. Thereafter suitable features are selected for the regions to perform segmentation. Wavelet transform based on octave band decomposition has been successful as texture feature [32]. However Chang and Kuo [34] observed that texture feature may be prevalent in intermediate frequency band of tree-structured wavelet transform rather than pyramid-structured wavelet transform. This aspect motivates researchers toward wavelet packet transform [35, 36] where sub band decomposition of wavelet transform is not restricted to be dyadic. Inaccurate texture edge localization [37] results from critically sampled filter banks. The overcomplete wavelet representation i.e., wavelet frame is one of the remedy for solving this problem [38].

Spatial domain filtering

In earlier days for obtaining edge density map spatial filtering was done. Fine texture has more edge density than that of coarser version. For this purpose Robert's and Laplacian operator [8, 39] are used. Malik and Parona [40] proposed following models of spatial filtering,

1. Input image is convoluted with a bank of even symmetrical filter followed by half wave rectification.

2. Removal of spurious responses in a localized area.

3. Detection of boundary between different textures.

The $(p+q)^{\text{th}}$ moments over an image region R for an image $I(x,y)$ are given below,

$$M_{pq} = \sum_{(x,y)\in R} x^p y^q I(x,y)$$

If moments are calculated for every point in the image and R denotes local rectangular area, then the above equation implies filtering of input image by a set of spatial masks. The resulting filtered images can therefore be used for extraction of textural features. Moment based features are successfully used for texture segmentation [41]. Apart from these there are wide varieties of spatial filters like Laws, DCT etc. Laws proposed microstatistical features after convolution of input image with twenty five two dimensional masks of size 5×5 obtained by outer product of five one dimensional masks of size 1×5. The microstatic features were extracted in the form of sum of the squared or absolute values of the micro feature plane. Later on eight number of two dimensional DCT basis masks of size 3×3 are constructed from three one dimensional filtering masks h_1=[1 1 1], h_2=[1 0 -1], h_3=[1 -2 1] leaving out the low frequency component as suggested by Ng et al. [42] for texture feature extraction. As these types of spatial masks are separable, minimal time and memory requiring implementation scheme is proposed [43].

1.2.3 Geometrical method

Texture is composed of smallest repetitive elements called as texel. Once the texture primitive is found, there are two ways to represent a texture. Statistical features extracted from the texture primitive are often used to identify the texture. Determination of the placement rules is another way to describe a texture. The latter approach is known as geometrical or syntactic method of analyzing texture. There are a number of methods for extracting texel and defining its placement rules. Among these Voronoi tessellation method is widely used.

Voronoi tesselation method

Extraction of texture primitive or texture element as proposed by Tuceryan and Jain [44] by using the properties of the voronoi tessellation has the following features

1. Suitable property for defining local spatial neighborhood.
2. Local distribution of texels are reflected in the shapes of voronoi polygons.

After extracting texels tessellation is constructed. Texture primitive can be defined as the highest gradient points or structures like line segments or closed boundaries. Steps for texel extraction are given as follows,

1. It starts with application of laplacian of Gaussian (LOG) filtering on input image. To an extent laplacian of Gaussian filter can be approximated by difference of Gaussian filter which in turn depends upon the size of the two Gaussian filter.
2. Pixels on a local maximum of filtered image is selected if the magnitude at the pixel location is greater than that of its six to eight nearest neighbors. This results a binary image.
3. Suitable algorithm for connected components analysis is applied on the binary image using eight nearest neighbors. Therefore texture primitive (token) as each connected components.

Thereafter voronoi tessellation of obtained token is constructed. Uniform texture regions with texels with similar is found extracting features from voronoi cell.

In this regard moments of area of the voronoi polygons constitute a set of features considering both i.e., spatial distribution and shapes of the texels in the textured image [45]. Unlike voronoi tessellation LOG filter at multiple scales is used for extracting texture elements [46]. Tomita and Suji [47] described texture token extraction by doing medial axis transform on the connected components of a segmented image. Placement rules defined by a graph for ideal textures are transformed to get the graphs for observable or real textures which is conceived as distorted version ideal textures [48]. According to Fu [49] texture is perceived as string defined by tree grammar whose terminal symbols are the texture element or primitive. This method is useful for analyzing and generating textures. Jayamurthy [50] describes texture $t(x, y)$ given a texture primitive $h(x,y)$ and placement rules $p(x,y)$ by $t(x,y)=h(x,y)\otimes p(x,y)$ where $p(x,y)=\sum\delta(x-x_n, y-y_n)$ and $\otimes$ indicates convolution.

Matsuyama et al. [51] extract texture primitives from regular texture by region growing. Toriwaki et al. [52] proposed adjacency of texture primitives by extending the concepts of voronoi neighbor, Gabriel neighbor relating neighbhor to connected components in digitized image.

In natural textures, texel as well as its placement rule is very difficult to find out. So the application of Geometrical methods is limited to some artificial and regular textures.

1.2.4 Model based method

Model based methods often try to interpret a texture by stochastic and generative model [53, 54, 55]. Before analyzing any image these model parameters need to be estimated. Computational complexity incorporated for estimation of parameters pose a problem. Autoregressive (AR) model expresses intensity of image pixels as a weighted sum of neighboring pixel intensities. Conceiving image (I) as zero mean random field, AR causal model can be given as

$$I_s=\sum_{t\in N_s}\phi_t I_t+e_s$$

where I_t is intensity at site t, N_s is neighborhood of s, e_s depicting as independent and identically distributed noise and ϕ is the vector containing model parameters. More efficient and simpler causal AR model as compared to other non causal spatial interaction method was proposed by Hu and Dennis [56]. Texture segmentation was carried out by estimating five parameters relating to four neighborhood pixels for causal half plane AR model neighborhood. Mean square estimation is employed for estimating the parameters. The extensive study of moving average AR model and autoregressive moving average model representation are found in [57]. Markov random field [58] model try to determine the probability of pixel belonging to particular state given the states of its neighboring pixels. Relation between probability and energy of an image is given by

$p \propto \exp(-e/c)$, c is constant. So an image pixel generated by MRF with less energy is more probable. Hidden Markov Model has a advantage (HMM) of extracting underlying fundamental structure of an image which may not be directly observable. It is shown to outperform autocorrelation method [59]. Segmentation can be done by maximizing a posterior probability for Markov Random Field (MRF) and Gaussian random field models [60]. MAP uses Gibbs random field as conditional density function is not accurately estimated by MRF. Computational complexity may rise due to large number of parameters defining interaction within and between color bands. In multiresolution technique [61] GMRF parameter estimation from fine resolution image is followed by segmentation from coarser to finer resolution.

Fractal dimension shows similarity with human recognition system for estimating surface roughness. Human vision system successfully characterizes natural textures having linear log Gabor power spectrum. In this regard fractal dimension shows some similarity with human vision in spite of having scale invariance [62]. Textured surfaces are often modeled by 2-D fractional Brownian motion (fBm) model [63]. The disadvantage of fBm is that it is isotropic. Research is ongoing to invent the suitable extended self similarity (ESS) for non-isotropic texture. fBm describes fractal dimension as [63] $D = E + 1 - H$, where E is Euclidean dimension and H is the Hurst parameter. Though various definitions of FD exit [64], it fails to segment all types of texture.

Actually 2-D models [65, 66, 67, 68] are better to characterize texture 1-D models though question lies in the choice of 2-D neighborhood (orientation and size) with respect to pixel under consideration. Tou et al. [65] represents 2-D model to two 1-D linear processes.

Model based methods are seldom used nowadays due to huge computational burden required for estimation of the model parameters.

1.3 Motivation

Advantages and disadvantages of each feature extraction methods have been discussed in a number of review papers [7]. Out of these, signal processing method is used by researchers extensively due to its similarity with Human Visual System (HVS) as pointed out by Julsez [4]. Classification can be carried out in two ways, namely, supervised and unsupervised. Before taking the problem of unsupervised classification, a small section on the application of supervised classification is also presented in the thesis. Multilayer perceptron (MLP), Bayesian network and probabilistic neural network (PNN) have been implemented here for classifying edge based textural features of different types of defective fluted ingots in supervised framework [69]. This work which is briefed as below, has been done as an extension of a previous study [70].

Fluted ingots are manufactured for different types of locomotive wheels. In order to prepare good quality wheels, there should not be any defect on the surface and at the interior part of the fluted ingots. Surface defects are monitored by camera-based system whereas internal holes and cracks are monitored by ultrasound testing. Now inspection for defective ingots is carried out generally at the final stage of production of locomotive wheels. But if at the last step a fluted ingot has been found as defective, it should be eliminated incurring huge loss. To avoid such an occurrence automated classification of defective ingot has been implemented at early stages of production cycle. Here two kinds of defect found in frontal surface of fluted ingot, namely, blade defect and metallurgical defect are considered. The ingot is usually cut into three pieces called cheese by band saw blade. Poor quality of band saw blade is mainly responsible for blade defect. On the other hand improper material composition at early stage of production cycle leads to metallurgical defect. After capturing analog image of frontal surface view of fluted ingot under adequate illumination through CCD camera, it is digitized through image grabber card. Further the digital image is preprocessed through operations like median filtering and contrast stretching. Textural features have been extracted to quantify the properties of a surface. Textural features may be of different types varying from features based on localized histogram map, edge density map, Fourier texture spectrum, co-occurrence matrix. Among the above mentioned techniques edge density map has been invoked for feature extraction. From the edge density map of a given image number of edge pixels has to be counted when edge density map of a pixel crosses some given threshold. Number of edge pixels found within a block serves as a feature for processing in non-overlapping blocks. The feature set thus evolved has been applied as input to Multilayer

Perceptron (MLP), Bayesian classifier and Probabilistic Neural Network (PNN). MLP is widely used ANN based technique where the objective is to optimize weights of the internal layers given a set of input and output. Levenberg Marquardt method has been used for training. On the other hand Bayesian classifier uses decision function as logarithm of the product of prior and likelihood probability of a pattern. Once covariance and mean of a class prototype is calculated from training dataset, unknown data can be classified into a class. Probabilistic Neural Network consists of three layers of input, pattern and class. During training unity spread of the Gaussian window provides maximum smoothness in the estimate of the probability density function without sacrificing performance. The goal of this study is to identify good ingots, ingots with blade defect and metallurgical defects. Each image of the frontal surface of fluted ingot has been divided into non-overlapping blocks of size 32×32 before applying it as input to different classifier. The input block of size 32×32 has been further sub-divided into non overlapping 8×8 block to calculate the edge count prevailing in each of sixteen ((32/8)X(32/8)) number of sub-blocks. Training has been carried out with exhaustive data sets. As far as the performance is concerned, the percentage of misclassified blocks for PNN is lesser than that of Bayesian network and MLP. Misclassification occurs when distinguishing between ingots having blade defect and metallurgical defect whereas good ingot and defective ingots are perfectly classified. But this type of supervised classification is paralyzed by the requirement large amount of training dataset.

There are various unsupervised classification or clustering techniques namely k-means, fuzzy c-means, hierchical clustering, self organizing map etc. Among the different classification techniques k-means and fuzzy c-means are widely used due to simplicity in spite of being sub-optimal [71].

Majority of the texture based techniques consider that images of a particular object are captured from the same viewpoint. But it is not a practical scenario. So always scale, translation, rotation and illumination are the key parameters related with a textured image. Invariant texture studies [72] deal with invariance due to these parameters. Scopes are there to extend these invariant methods to 3-D texture and perspective transformation. More attention should be paid to affine models. Rotation invariant textural method can be of four types - statistical, structural, signal processing and model based methods.

In signal processing techniques a set of Gabor filters are proposed to obtain invariant texture features [72]. Porter and Canagarajah [73] developed the wavelet transform for invariant texture analysis based on the Daubechies 4 tap wavelet filter for 3 level wavelet decomposition. The horizontal and vertical components at each frequency are combined to get the rotation invariants leaving out HH channel at each decomposition level since it mostly contains noise. The drawback of this method is that the directional information is lost when the channels are combined.

Texture is a neighborhood property. So contextual information always helps in grouping neighboring pixels or local region to the same class [56]. But many a time it assigns dubious label to pixel or local region for three or more texture discrimination. Contextual information often enhances the classification accuracy with the addition of little computational complexity. Contextual structure of Bayesian method is modeled by Markov Random Field (MRF) for contextual labeling [56]. This results in smooth boundary and high computational complexity. Feature extraction in smaller blocks for better resolution introduces noise causing less stability. Spatial contextual information often is used to attenuate this variation or noise in the feature space [56].

Texture segmentation is basically subdivided into two methods: region based segmentation and boundary based segmentation. In region based segmentation generally similar pixels or local areas belonging to same region are classified as member of same class. Drawback of this method is that number of distinct texture category present along with threshold for similarity measure should be specified beforehand. Also boundary of different textural regions must be closed. In case of boundary based segmentation main idea is to detect difference of texture in the adjacent regions without the prior knowledge about the number of texture categories present. However, if boundaries have gap, then two regions with different textures may not be identified as a separate closed region. The boundary based and region based techniques are merged by Jain [7] for cleaner and robust segmentation. The integration of boundary based segmentation and region based segmentation improves the resulting segmentation in most cases. Similar to this, watershed technique merges the two techniques for better results. Watershed treats every image as topographic representation in three dimensions namely two spatial coordinates and one gray level. But watershed produces over segmentation for occurrence of multiple minima due to presence of noise in any image. Though marker controlled watershed is often used to inhibit local minima, watershed in graph theory framework of the shortest-path forest transform (IFT) is guaranteeing for optimal solution [74].

Feature level fusion belongs to a broad class of multisensor fusion when sensor data are non-discrimatory [75]. There are various reasons for fusion of features like improvement in high frequency resolution of Gabor filter derived features [76]. Apart from classification accuracy, quantitative tools used for evaluation of accuracy of fused features, are methods like Fisher's criterion, Dempster–shafer's evidential reasoning etc.

Nowadays a large number of OSF (Oral Submucous Fibrosis) cases are transformed into oral cancer which is one of the deadliest disease across the world. Image of OSF captured by any means contains some natural pattern or texture. The rotated versions of these original pattern may be located anywhere within the image. So in that context plain or rotational invariant textural features help segmentation of this image into constituent layers to an extent. Current domain of confirmative diagnosis sometimes extends from expertise of medical experts to qualitative histopathological evaluation of the biopsies. There is a dearth of established techniques for quantitative evaluation of histopathological features like thickness of different constituent layers. This could be only possible if Light Microscopic (LM) images of OSF is segmented into constituent layers. Still precancerous stage detection is an open problem as it contains mixed features of normalcy as well as pro or pre malignancy. Transmission Electron Microscopic (TEM) images of incisional biopsies of OSF is heterogenous mixture of axial, transverse and collagen free sections. The simultaneous presence of three sections really pervades the process of stage detection for malignancy. In that respect segmentation of TEM images of OSF into constituent sections may ease the further processing.

1.4 Objectives of the thesis

Though good amount of research has been carried out in texture segmentation still it has got lacuna in many theoretical and application aspects like universal definition of texture, rotational invariance in block based processing, optimum fusion strategies, feature normalization, feature minimization, affine transform for invariance in 3-D, heavy computational burden of texture algorithms, block size vs. resolution for feature extraction etc. Among these the following aspects have been taken as objective of this thesis.

- Simple and computationally attractive rotational invariant features of rotated textured images (2-D) as well as actual rotation (3-D) of the textured surfaces.
- Suitable measures for determining the number of texture categories present in an image.
- Efficient and suitable contextual features for texture segmentation in non overlapping blocks.
- Hybrid method of image segmentation with reasonable classification accuracy and computational time.
- Multifarious feature level fusion and their evaluation.

1.5 Contribution of the thesis

The contribution of the thesis is summarized as below.

- Rotational invariant features based on sixty four two dimensional DCT basis filtering masks of size 8×8 has been proposed. The efficacy of proposed features is shown in case of texture classification and unsupervised texture segmentation.
- Effect of adding contextual features has been studied for automatic segmentation.
- The above two techniques have been successfully applied to Transmission Electron Microscopic images of different stages of OSF.
- Results of Hybrid segmentation algorithm using watershed in the shortest path forest transform as compared to Region growing algorithm (RGA) have been presented for segmentation of precancerous stage of OSF.
- Fusion of proposed rotational invariant features based on DCT, contextual features and features extracted from region grown images has been proposed while segmenting Light Microscopic images of precancerous lesions of OSF.

1.6 Organization of the thesis

The thesis has been divided into six chapters including the introduction as the first chapter.

In chapter 2, initially rotational invariant features using DFT encoded even symmetric Gabor filtering in non overlapping blocks followed by fuzzy c-means clustering has been studied. Number of texture categories presenting an image has been determined automatically. Thereafter in this chapter statistical features of mean and variance extracted from the input image after convolution with the horizontal and vertical version of the original sub-mask of Discrete Cosine Transform (DCT) basis filtering masks of size 8×8 have been proposed as rotational invariant textural features. Minimum distance classifier has been applied to classify the query images based on the proposed features. Motivated by success of proposed features in texture classification an unsupervised texture segmentation approach using rotational invariant energy features based on DCT is also proposed in chapter 2. The DCT based rotational invariant features used here show interesting results compared to plain wavelet transform and conventional DCT features with and without rotation. The proposed method is applied to texture image database ranging from natural texture to natural scene. The algorithm performs well with the addition of white Gaussian noise also. Another objective of the chapter 2 is to describe the efficacy of proposed rotational invariant energy features based on DCT in comparison to feature extraction methodologies based on rotational invariant wavelet transform and wavelet packet transform for segmenting Transmission Electron Microscopic (TEM) images of Oral Submucous Fibrosis (OSF).

Chapter 3 presents a study of different texture feature extraction techniques in wake of contextual information for blockwise feature extraction. An additional feature of positional information is incorporated in conventional features based on Laws, DCT, Traditional Gabor filter and co-occurrence method. The classification accuracy is validated in presence of noise. The study has been conducted on mostly natural monochrome images and some biomedical images.

Chapter 4 concentrates on segmentation of histological images of oral sub-mucous fibrosis (OSF) into its constituent layers. In this regard hybrid segmentation algorithm shows very interesting results. The segmentation results depict superiority of hybrid segmentation algorithm (HSA) in comparison to region growing algorithm (RGA).

The aim of chapter 5 is to segment light microscopic images of Oral Submucous Fibrosis (OSF) into its constituent layers. In this regard fusion of features based on Region Growing Algorithm (RGA) and context enhanced rotational invariant Discrete Cosine Transform (DCT) has been studied.

Finally chapter 6 concludes the thesis and is pointing at some of the future directions.