Early Detection of Oral Cancer through multi-modal imaging and analysis

Report submitted in fulfillment of the requirements for the degree

of

Bachelors of Technology

In

Electrical Engineering

By

Anirudh Choudhary

05EE3008

Under the supervision of

Prof. Jayanta Pal   **Prof. Ajoy Kumar Ray**

Department of Electrical Engineering   School of Medical Science & Technology

Indian Institute of Technology, Kharagpur-721 302

INDIA

Jan 2008 – May 2009

Department of Electrical Engineering,

Indian Institute of Technology,

Kharagpur-721 302

INDIA

CERTIFICATE

This is to certify that the work presented in this thesis entitled "**Early Detection of Oral Cancer through multi-modal imaging & analysis**" submitted by **Mr. Anirudh Choudhary (Roll No: 05EE3008)** in partial fulfillment of the requirements for the award of the degree of **Bachelor of Technology (Hons)** is a bonafide record of work carried out by him under my supervision. To the best of my knowledge and belief, it has not been submitted to any other institute or university for the award of any degree or diploma.

Professor Ajoy Kumar Ray

Vice-Chancellor

Bengal Engineering and Science University

Shibpur

West Bengal, INDIA

5th May,2009

Professor Jayanta Pal

Professor, Electrical Engineering

Department of Electrical Engineering

Indian Institute of Technology

Kharagpur-721302

West Bengal, INDIA

5th May,2009

ACKNOWLEDGEMENTS

Abstract

Oral cancer is almost always preceded by certain oral pre-cancerous lesions and conditions. Important oral pre-cancerous lesions include leukoplakia and erythroplakia while oral sub mucous fibrosis (OSF) is the most common and potent pre-cancerous condition. Histopathological pre-cancer in OSF depicts changes in both sub-epithelial & epithelial mucosa. In current scenario much work has been done on evaluating the features of the epithelium layer but the features of the subepithelium layer haven't been explored. This thesis aims at quantitative analysis of textural & histopathological features of subepithelium layer in precancerous lesion using image processing techniques. It aims at quantitative analysis of the layer using Gabor & Wavelet based texture analysis to represent the layer using a entropy based feature vector. Tree structured wavelet transform is applied to construct an entropy map of the subepithelium layer & robust cell segmentation based on marker based watersheds is employed for cell density measurement. Implementation of presented algorithms is done in MATLAB. The results support the capability of wavelet based decomposition & Gabor Transforms to capture the change in texture of subepithelium with respect to the progression of cancer.

Contents

Chapter5

## Chapter 6

CHAPTER1

Introduction

Cancers of the oral cavity are increasing at an alarming rate across the globe and it is reported that each year approximately 310000 new oral cancer cases are detected. Higher incidence of Oral Cancer is mainly due to late diagnosis of potential precancerous lesions and conditions [1]. Oral Leucoplakia (OLK) is a precancerous lesion whereas Oral Submucous Fibrosis (OSF) is an insidious chronic progressive precancerous condition with a high degree of malignant potentiality [18]. A large number of these cases transform into Oral Cancer. Though aetiology of these pathosis is not well understood but maximum correlation has been observed with different addictive oral habits in combination with malnutrition and genetic predispositions. The development of oral cancer seems to begin in many cases with exposure of the mucosal surfaces of the upper aerodigestive tract to topical carcinogens, predominantly alcohol and tobacco. In some persons exposed to these carcinogens or co-carcinogens, premalignant and malignant lesions develop in a multi-step process within the mucosa.

Oral submucous fibrosis is characterized by epithelial atrophy & fibrosis of the subepithelial connective tissue, resulting in the stiffness of oral mucosa. Histopathology shows thinned out mucosa and thickened avascular subepithelial connective tissue in most of the OSF cases. Accurate histopathological diagnosis (i.e. determination of the nature of a disease by means of microscopic examination) is useful at several levels. For instance it is important to be able to differentiate a) normal from abnormal tissues, b) between premalignant and malignant, c) between those premalignant lesions that are likely to transform and those which are not so,

and d) to differentiate the malignant lesions that are likely to have a good response to a particular treatment from those which are not.

The most effective way to control oral cancer is to combine early diagnosis and timely and appropriate treatment. Because more than 90% of all oral cancers are squamous cell carcinomas, the vast majority of oral cancers will be diagnosed from lesions on the mucosal surfaces. Presently, in diagnosis of OLK, OSF and their grading with respect to progression and malignant potentiality, mainly qualitative histopathological techniques & biopsies are applied.

Fig 1.1 Microscopic structure of cell image & the components of Oral Mucosa
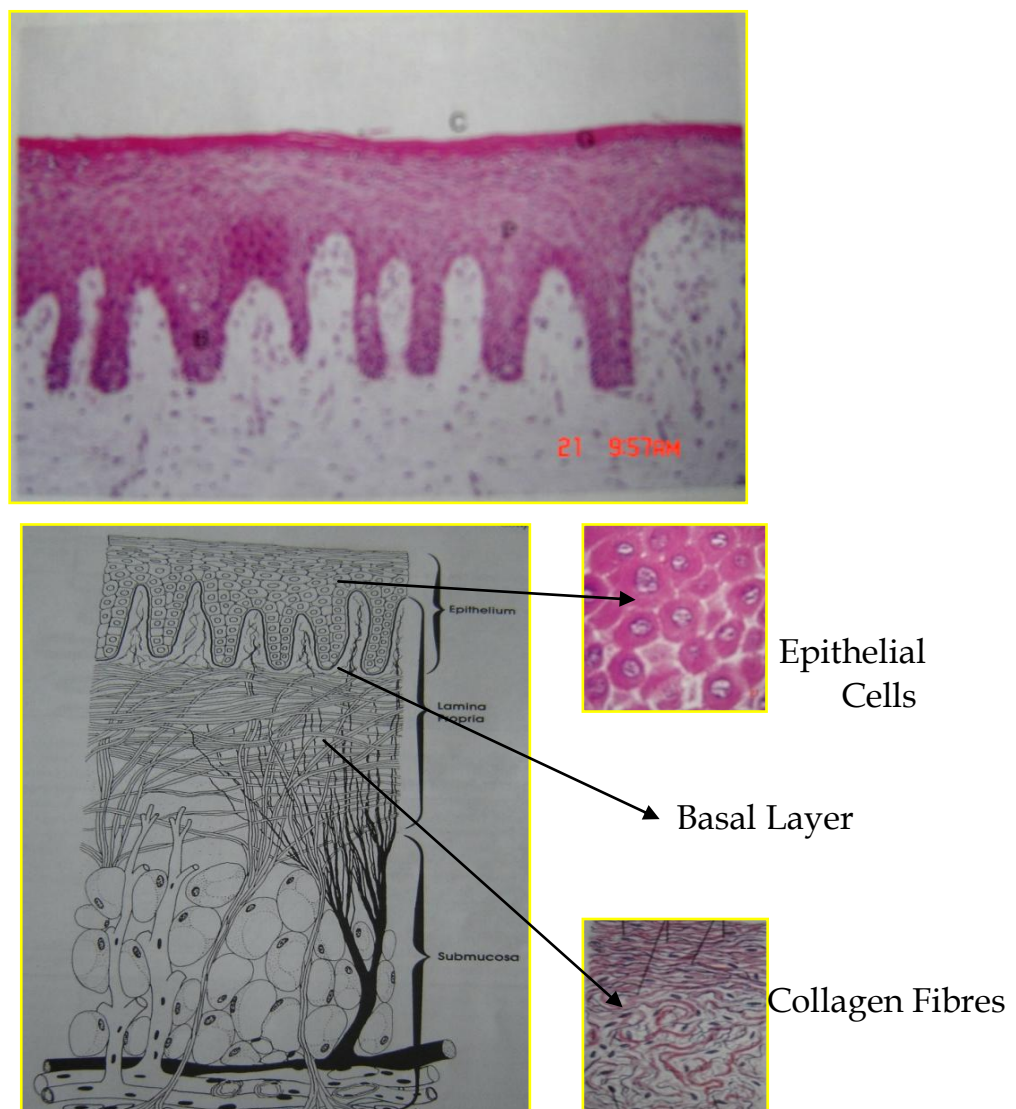
No specific quantitative histopathological evaluation procedure based on texture is available to assess and analyze the vital histopathological changes in the cell arrangement in oral mucosa to explain these pathological conditions precisely. In this thesis, the textural features of the subepithelial structure of OSF and normal oral epithelium with respect to the cell density & entropy of the subepithelial layer are extracted. Since the Oral Submucous Fibrosis starts from the basal layer, hence the subepithelial layer will be the first layer to go any changes in cell orientation, color as well as cell density & hence its analysis can be used for early diagnosis & detection of oral lesion. Texture is a representative of the morphological features of this layer & capturing the texture information of this layer is important for histological analysis.

In this analysis, to extract the texture & cell features, image processing technique has been adopted using Gabor Filters, Tree structured wavelet transform, and watershed based cell segmentation. Chapter 2 compare the various entropy based thresholding algorithms applied on medical datasets so that the best possible thresholding method will be used in subsequent cases whenever needed. Chapter 3 has presented an overview of wavelet based analysis of the test images. More specifically, first part of the chapter focuses on Gabor filter-based texture classification methodology. Second part has described tree structured wavelet based technique to capture the texture information of the test images. Chapter 4 deals with the seeded watershed based adaptive cell segmentation. Finally, conclusions are drawn in Chapter 5.

CHAPTER 2

Entropy based Thresholding

Segmentation involves separating an image into regions (or their contours) corresponding to objects. The simplest property that pixels in a region can share is intensity. So, a natural way to segment such regions is through thresholding, the separation of light and dark regions. Thresholding is an important technique in medical images & is used in combination with a variety of methods like region growing, watershed segmentation & texture based segmentation. Thresholding creates binary images from grey-level ones by turning all pixels below some threshold to zero and all pixels about that threshold to one.

Of particular interest among thresholding techniques is an information theoretic approach that is based on the concept of entropy introduced by Shannon in information theory. The principle of entropy is to use uncertainty as a measure to describe the information contained in a source. The maximum information is achieved when no a priori knowledge is available, in which case, it results in maximum uncertainty. In other words, if there is no preference among samples resulting from an experiment the best solution is to treat all samples equally. The entropy based image thresholding methods proposed by Pun, Kapur & Fuzzy theory consider an image histogram as a probability distribution, and then select an optimal threshold value that yields the maximum entropy. Relative entropy has also been used as an alternative thresholding criterion. The minimum cross entropy method proposed by Li & Lee[8] computes the error in such a way so that the selected threshold will minimize the error in pixel classification. In this chapter we do a comparative study between entropy-based thresholding methods and relative entropy-based thresholding methods for single as well as multi-thresholding. We

also incorporate Genetic Algorithm into multi-thresholding methods to optimize computation & time.

Entropy of an image is a statistical measure of randomness which defines the average amount of information stored in a pixel in the image and as its magnitude increases the uncertainty and hence the information associated with the source increases. It describes the business of an image i.e. the amount of information which must be coded by a compression algorithm. It was first brought to communication theory by Shannon and it is used to measure the efficiency of the information transferred.

Mathematical definition of entropy H:

$$H = -\sum_{i=1}^{n} p(x_i) log_2 p(x_i)$$

where $p(x_i)$ is statistical probability density of occurring of event I (in case of image this is the probability of intensity level i).

2.1 Minimum Cross Entropy Based Thresholding

Let F ={$f_1, f_2, \dots, f_N$} and G ={$g_1, g_2, \dots, g_N$} be two probability distributions on the same set. The cross entropy between F and G is an information theoretic distance between the two distributions and it is defined by

$$D(F,G) = \sum_{i=1}^{N} f_i \, log \frac{f_i}{g_i}$$

The minimum cross entropy thresholding (MCET) algorithm [8] selects the threshold by minimizing the cross entropy between the original image and its thresholded version. Let I be the original image and h(i), i =1,2, ... ,L, be the corresponding histogram with L being the number of gray levels. Then the thresholded image, denoted by I_t, using t as the threshold value is constructed by

$$I_t(x,y) = \begin{cases} \mu(1,t) & , \ I(x,y) < t \\ \mu(t,L+1), & I(x,y) \geq t \end{cases} \qquad \mu(a,b) = \sum_{i=a}^{b-1} ih(i) / \sum_{i=a}^{b-1} h(i)$$

The cross entropy is then calculated by

$$D(t) = \sum_{i=1}^{i-1} ih(i) \log\left(\frac{i}{\mu(1,t)}\right) + \sum_{i=1}^{L} ih(i) \log\left(\frac{i}{\mu(t,L+1)}\right)$$

The MCET determines the optimal threshold t by minimizing the cross entropy

$$t^* = \arg min_t\{D(t)\}$$

2.2 Pun's entropy based thresholding

Thresholding a picture is equivalent to splitting the a priori entropy H of the histogram in two parts, $H_w = \alpha H$ and $H_b = (1-\alpha)H$. Pun[7] proposes the use of an anisotropy coefficient α in thresholding which represents the ratio of the average quantity of information that we attribute a priori to white and black points & is given as:

$$\alpha = \left(\sum_{i=0}^{t} p_i. log_2(p_i)\right) / \left(\sum_{i=0}^{n} p_i. log_2(p_i)\right)$$

Where t is threshold and α determines the geometric shape of the histogram.

The optimal threshold t is determined as

$$\sum_{i=0}^{t} p_i = \frac{1}{2} + \left|\frac{1}{2} - \alpha\right| = \begin{cases} 1 - \alpha & \text{if } \alpha \le 1/2 \\ \alpha & \text{if } \alpha > 1/2 \end{cases}$$

For k level threshold, k-1 thresholds are selected as 1, 2, 3……, k-1 and similarly α-1 anisotropic coefficients are selected.

2.3 Kapur's entropy based thresholding

In this method [6] two probability distributions are considered, one for the object and other for the background which are derived from the original gray level distribution of the image as follows.

$$\frac{p_0}{P_B}, \frac{p_1}{P_B}, \dots\dots, \frac{p_s}{P_B} \qquad\qquad \frac{p_{s+1}}{1-P_B}, \frac{p_{s+2}}{1-P_B}, \dots\dots, \frac{p_{n-1}}{1-P_B}$$

where s is the threshold, p_i is statistical probability, P_B is probability of pixel with gray level less than or equal to threshold s.

The entropy of the object and the background of image will be given as follows:

$$H_B^{(s)} = - \sum_{i=0}^{s} \frac{p_i}{P_B} log_2 \left(\frac{p_i}{P_B}\right)$$

$$H_W^{(s)} = - \sum_{i=0}^{s} \frac{p_i}{1-P_B} log_2 \left(\frac{p_i}{1-P_B}\right)$$

Then the optimal threshold t* is defined as the gray level which maximizes $H_B^{(t^*)} + H_W^{(t^*)}$, that is,

$$t^* = ArgMax_{t \epsilon G}\{H_B^{(t^*)} + H_W^{(t^*)}\}$$

2.4 Fuzzy Entropy based thresholding

Fuzzy entropy based thresholding [17] considers image as a fuzzy set on which it applies entropy maximization. If we segment the image I into C partitions, i.e we need to construct a series of C sets such that

1. $A_i \subset I, A_i \neq \emptyset \; ; i = 1,2,\ldots.., C$

2. $A_i \cap A_j = \emptyset, if \; i \neq j$

3. $\bigcup_{i=1}^{C} A_i = I$

If we regard the image as a fuzzy set, the hard classification would be replaced by fuzzy classification, which means each set A has a membership function $\mu_i \; (p_{xy})$ subject to the constraint that

$$\sum_{i=1}^{C} \mu_i \; (p_{xy}) = 1 \; for \; \forall \; x, y$$

$$p_{xy} \in A_i \leftrightarrow \mu_i \; (p_{xy}) > \mu_j \; (p_{xy}), \qquad j = 1,2,3, \ldots \ldots \ldots, C$$

Correspondingly we propose the fuzzy entropy as

$$H(I) = \sum_{i=1}^{C} H(A_i)$$

where $H(A_i) = -\sum_{l=0}^{L-1} \frac{h(l)}{\mu_i(l)P'(A_i)} log \frac{h(l)}{\mu_i(l)P'(A_i)}$

where $P'(A_i) = \sum_{i=0}^{L-1} \frac{h(l)}{\mu_i(l)}$

The fuzzy entropy represents the distribution of quotients of the grayscale histogram and fuzzy membership. Fuzzy entropy based thresholding aims to

maximize $H(I)$ which implies that the coefficients will follow equal distribution & the histogram h(l) is exactly following the fuzzy membership $\mu_i(l)$. The fuzzy entropy based method is able to restrain noises without much loss of image details, because the impact of noises is taken into account in the definition of fuzzy membership. For finding n thresholds in an image, we initialize n thresholds randomly & calculate average grayscale value, membership function & entropy of each all pixels lying between thresholds t_{i-1} & t_i. Then compute the total fuzzy entropy of the image & follow this procedure iteratively unless the fuzzy entropy is maximum.

2.5 Optimization using Genetic algorithm

A **genetic algorithm (GA)** is a search technique used in computing to find exact or approximate solutions to optimization and search problems. It varies a set of parameters and evaluates the quality or "fitness" of the results of a computation as the parameters are changed or evolved .It randomly generates a population of individuals and these individuals are mutated and crossovered to generate the child population with better fitness thereby maximizing the overall fitness of population.

2.5.1 *Operators of Genetic Algorithm*

1. Reproduction: - This is the selection technique in which the functions are selected according to their fitness. There are various methods used for reproduction like Roulette wheel selection, Tournament selection, Elitism etc.

2. Crossover: - It involves exchanging elements through mating of two parent chromosomes to create one or more child chromosomes. The crossover is implemented by randomly choosing point(s) in the string called the crossover point(s) and exchanging the segments.

3. Mutation: - It creates new individuals by modifying one or more bit of string. The operator increases the variability of the population and prevents premature convergence to a poor local minimum.

2.5.2 *GA based image thresholding algorithm*

Each threshold variable is stored in an 8 bit chromosome and the number of thresholds, no. of population to be generated and population are taken as inputs from the user.

Fitness: Fitness is determined by calculating the entropy of each individual and the one with maximum entropy is the fittest one.

Reproduction: Reproduction is done using roulette wheel and elitist selection.

Crossover: Crossover point is selected randomly depending on the crossover probability. In our algorithm we used multiple point crossovers by selecting 2 crossover points randomly. Crossover probability is normally kept high (0.9).

Mutation: Mutation is done only if a random number selected between 0 and 1 has a value less than 0.01(mutation probability).Mutation probability is generally kept low.

Termination Criterion: When the program executes to the maximum generation (termination condition) or the elitist fitness does not change for n generations (stable condition), the program terminates. The termination condition is set to 1000 generations while the stable condition is set to 50 generations.

The general architecture of genetic algorithm based threshold determining technique is shown in the figure given below:
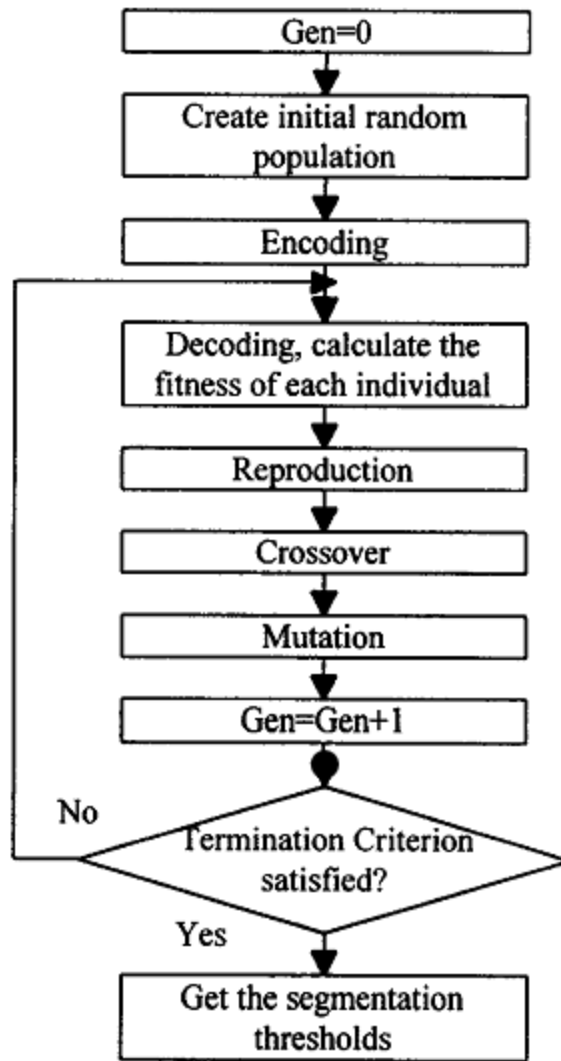
Fig [2.1] Flowchart for thresholding using Genetic Algorithm

2.6 Results

The 4 entropy based thresholding methods discussed above were applied on medical datasets with the number of thresholds equals to 3.The results of one of the datasets are shown in Fig 2.2 .It has been found from the results that the order of classification error of the 4 thresholding methods is

Pun's method> Kapur's method > Fuzzy Entropy based method > Cross entropy minimization method
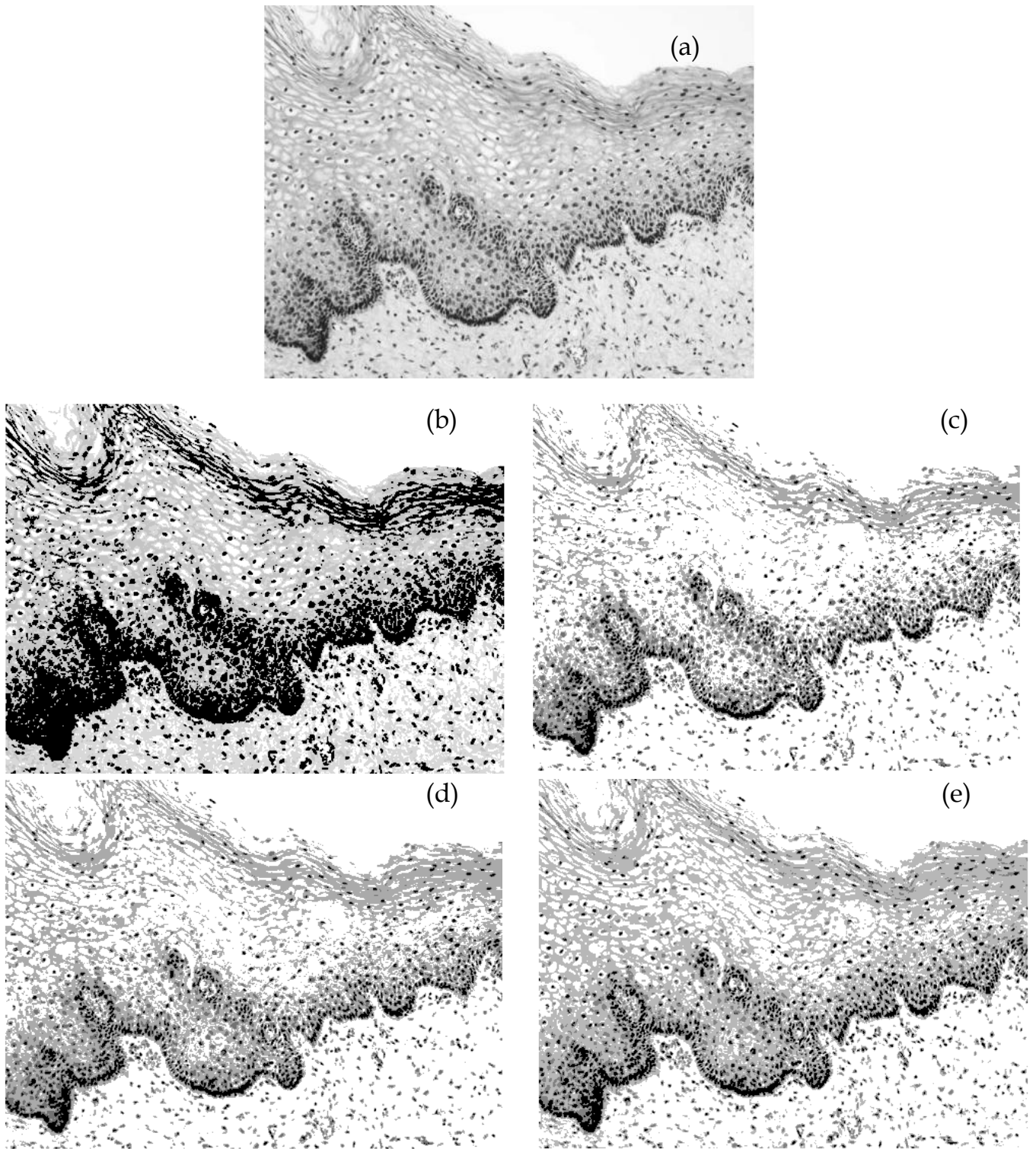
Fig[2.2] (a) Original Image (b) Pun's entropy based thresholding (c)Kapur's entropy based Thresholding(d)Fuzzy entropy based thresholding (e) Cross entropy minimization based thresholding.

CHAPTER 3

Texture Analysis

Texture is an important property of surfaces which characterizes the nature of the surface. An important task in image processing and machine vision is the task of segmenting regions of different texture in an image. However, images of real objects often do not exhibit regions of uniform intensities. For example, the image of epithelium layer is not uniform but contains variations of densities of cells which form certain repeated patterns called visual texture. The patterns can be the result of physical surface properties such as roughness or oriented strands which often have a tactile quality, or they could be the result of reflectance differences such as the color on a surface.

A typical definition of texture in the literature is "a spatial arrangement of local (gray-level) intensity attributes which are correlated (in some way) within areas of the visual scene corresponding to surface regions". It can be seen from Fig 3.1 that the cells of the sub-epithelium layer exhibit texture which is different from the remaining epithelium layer due to the orientation & greater density of cells in the sub-epithelium layer.

3.1 Gabor Filters

The frequency analysis of the textured image is best done in the Fourier domain. The human visual system analyzes the textured images by decomposing the image into its frequency and orientation components. For localizing the analysis in the spatial domain we introduce spatial dependency into the Fourier analysis using the window Fourier transform. The window Fourier transform of a one-dimensional signal is defined as:

$$F_w(\mu, \epsilon) = \int_{-\infty}^{\infty} f(x)w(x - \epsilon)e^{-j2\pi ux}dx$$

When the window function $w(x - \epsilon)$ is Gaussian, the transform becomes a Gabor transform.

Daugman[16] proposed the use of Gabor filters in the modeling of the receptive fields of simple cells in the visual cortex of some mammals. Later Farrokhnia and Jain[3] used it successfully in segmentation and classification of textured images. Gabor filters have some desirable optimality & biological properties. Gabor filters have the ability to perform multi-resolution decomposition due to its localization both in spatial and spatial-frequency domain. Texture segmentation requires simultaneous measurements in both the spatial and the spatial-frequency domains. Filters with smaller bandwidths in the spatial-frequency domain are more desirable because they allow us to make finer distinctions among different textures. On the other hand, accurate localization of texture boundaries requires filters that are localized in the spatial domain. Gabor filters are well suited for this kind of problem as their effective width in the spatial domain and their bandwidth in the spatial-frequency domain are inversely related according the uncertainty principle(**ΔxΔu ≥ π/4** & **ΔyΔv≥ π/4**).

A two-dimensional Gabor function consists of a sinusoidal plane wave of a certain frequency and orientation modulated by a Gaussian envelope[16]. It is given by:

$$h(x,y) = \frac{1}{2x\sigma_x\sigma_y} exp\left\{\frac{-1}{2}\left[\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2}\right]\right\} . cos\left(2\pi\frac{x'}{\lambda} + \varphi\right)$$

where   x'= xcos(θ) + ysin(θ), y'= ycos(θ) − xsin(θ)

In this equation, λ represents the wavelength of the cosine factor, θ represents the orientation of the normal to the parallel stripes of a Gabor function in degrees, ψ is the phase offset in degrees, and γ is the spatial aspect ratio and specifies the elliptically of the support of the Gabor function, and σ_x & σ_y are the standard deviations of the Gaussian in x & y directions.

Frequency Response of Gabor Filter
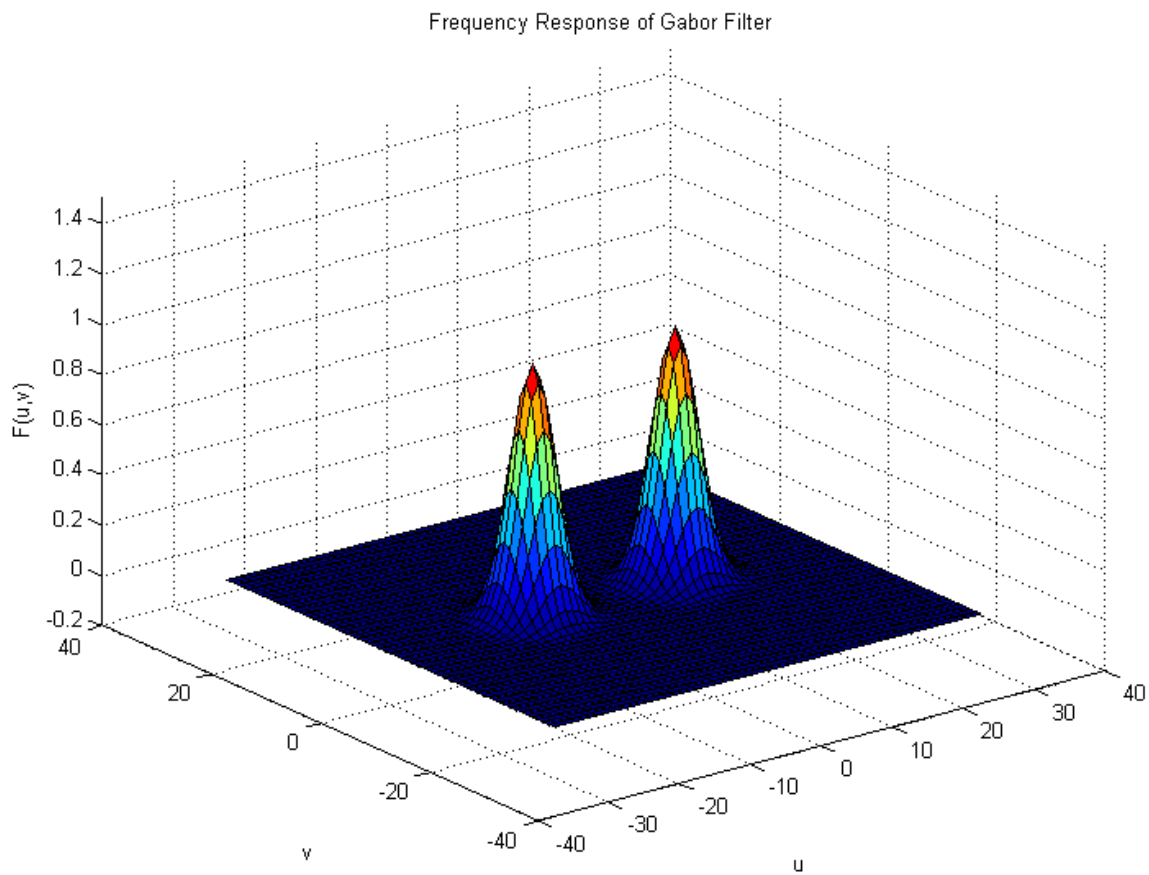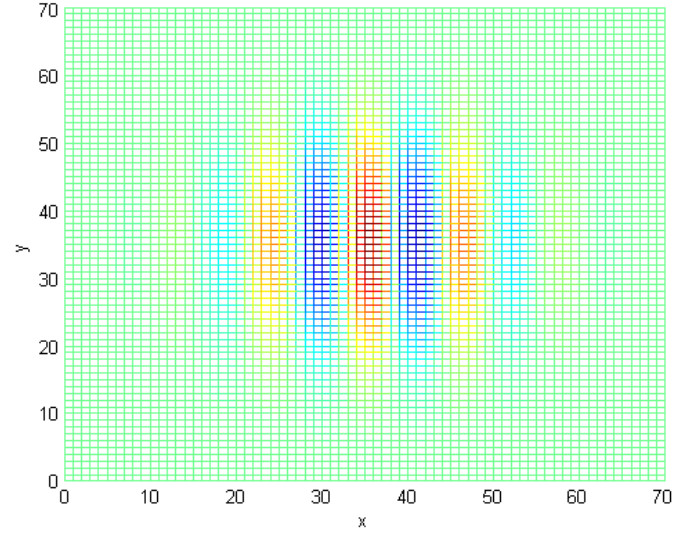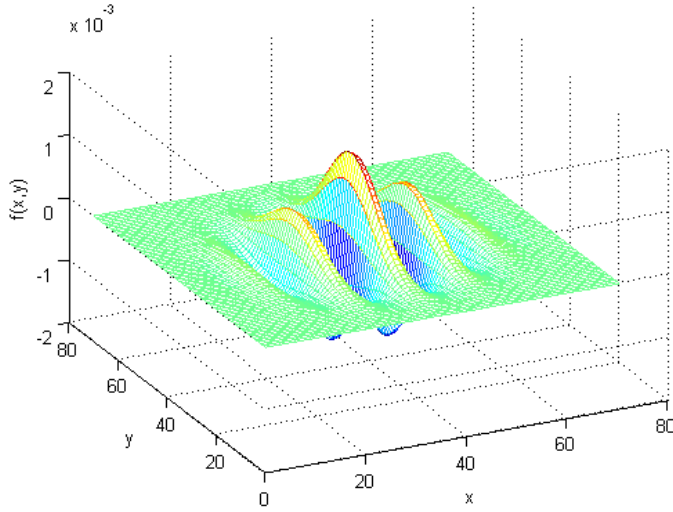
The ratio $\dfrac{\sigma}{\lambda}$ determines the spatial frequency bandwidth of simple cells and thus the number of parallel excitatory and inhibitory stripe zones which can be observed in their receptive fields. The half-response spatial frequency bandwidth b (in octaves) and the ratio $\dfrac{\sigma}{\lambda}$ are related as follows:

$$b = log_2 \frac{(\frac{\sigma}{\lambda}\pi + \sqrt{\frac{ln2}{2}})}{(\frac{\sigma}{\lambda}\pi - \sqrt{\frac{ln2}{2}})} , \quad \frac{\sigma}{\lambda} = \frac{1}{\pi}\sqrt{\frac{ln2}{2}}\frac{(2^b+1)}{(2^b-1)}$$

ψ = 0° and ψ = 90° returns the real part and the imaginary part of Gabor filter respectively.

3.2 Texture Segmentation

The process of texture segmentation using multi-channel filtering involves the following steps:

- Decomposition of the input image using a filter bank
- Feature extraction
- Clustering of pixels in the feature space

3.2.1 *Choice of filter parameters*

In our case we use orientation separation angles of 30° as it captures more texture features [4] i.e.:

$$\theta: 0°, 30°, 60°, 90°, 120°, 150°$$

The frequencies used for the filter are as recommended in [5] i.e.

$$F_L(i) = 0.25 - 2i^{\frac{-0.5}{N_c}}$$

$$F_H(i) = 0.25 + 2i^{\frac{-0.5}{N_c}}$$

where i = 1, 2, ..., $log_2(\frac{N_c}{8})$, N_c is the width of image which is a power of 2. Note that $0 < F_L(i) < 0.25$ and $0.25 \leq F_H(i) < 0.5$.

These frequencies were carefully chosen so that the center frequencies of channel filters lies close to the characteristic texture frequencies. Also care has to be taken so that the filters do not overlap in the frequency domain to avoid aliasing. In our case, we set the value of the bandwidth b of the Gabor filter to 1 octave & the total number of filters used is 72.

3.2.2 *Feature Extraction*

To extract useful information from the filter outputs, a nonlinear sigmoidal function was applied on each of the 72 filter output images. The nonlinear sigmoidal function saturates the output & acts as a blob detector. Each filter output is then smoothed using a Gaussian smoothing function having $\sigma = 3\sigma_s$ where σ_s is the scale parameter of the Gabor filter to suppress large variations in areas with similar texture. After smoothing; each pixel can be represented by a feature vector having 72 features each corresponding to different orientations & frequencies. We include the spatial coordinates of the pixels as 2 additional features to take into account the spatial adjacency information.

3.2.3 *K-means Clustering*

The final step is to cluster the pixels into a number of clusters representing the texture regions using K-means clustering method. The algorithm is as follows:

1. Initialize centroids of K-clusters randomly by taking K points among the samples**.**
2. Assign each sample to the nearest centroid based on the eucledian distance between the feature vectors of the sample & the centroid.
3. Calculate the new centroids (means) of 7-clusters.
4. If centroids are unchanged, finish. Otherwise, go to step 2**.**

The image containing the epithelium layer was segmented into clusters represented by different colors in the images shown in Figure [2.1].It should be noted that the number of clusters varies from one dataset to another but mostly a value of 6 or 7 is found to yield the desired result.
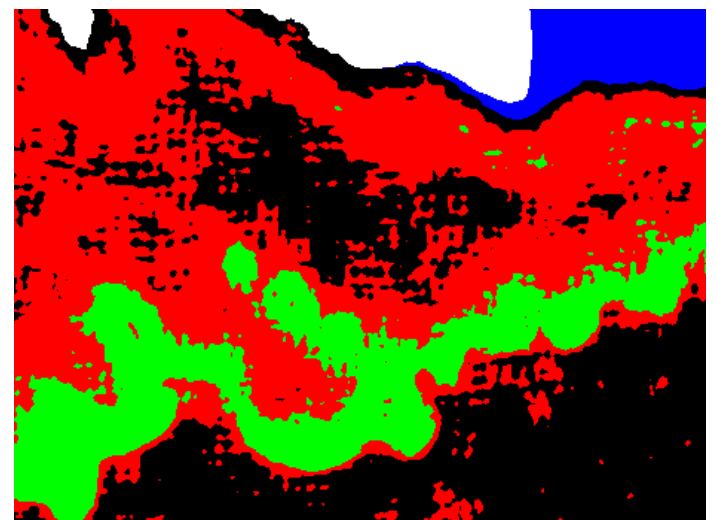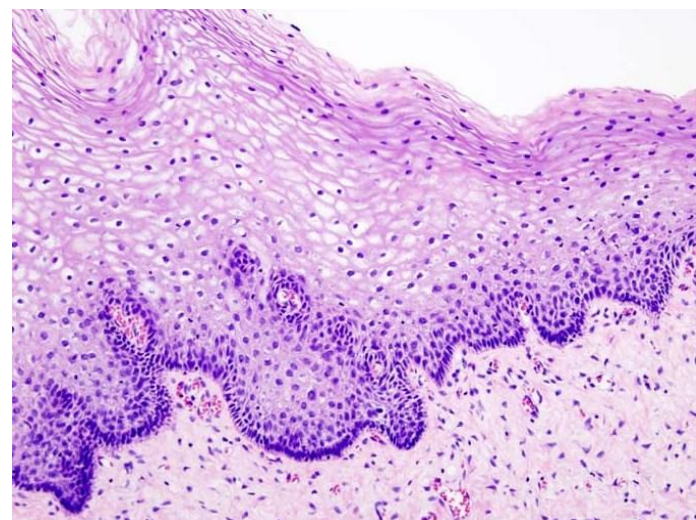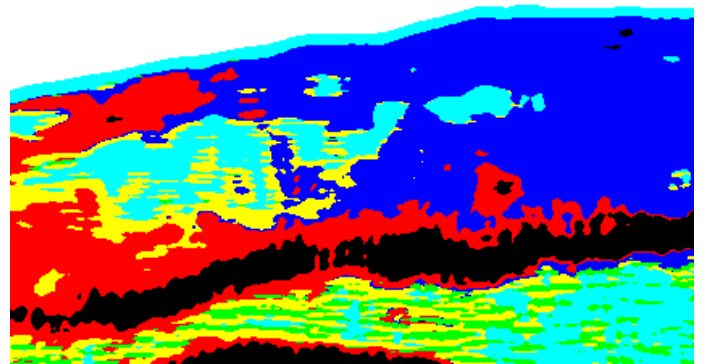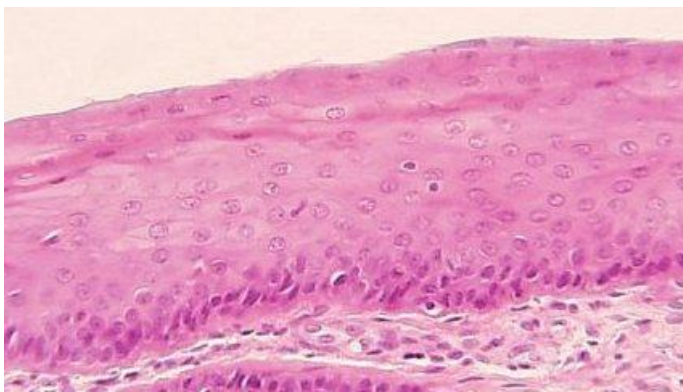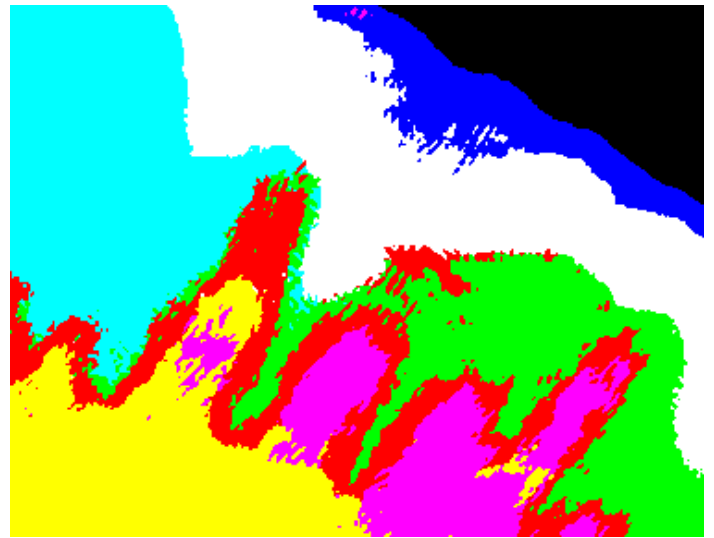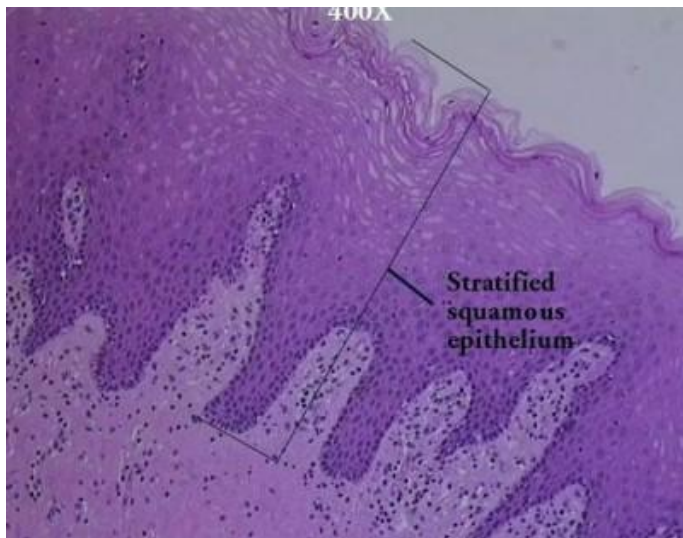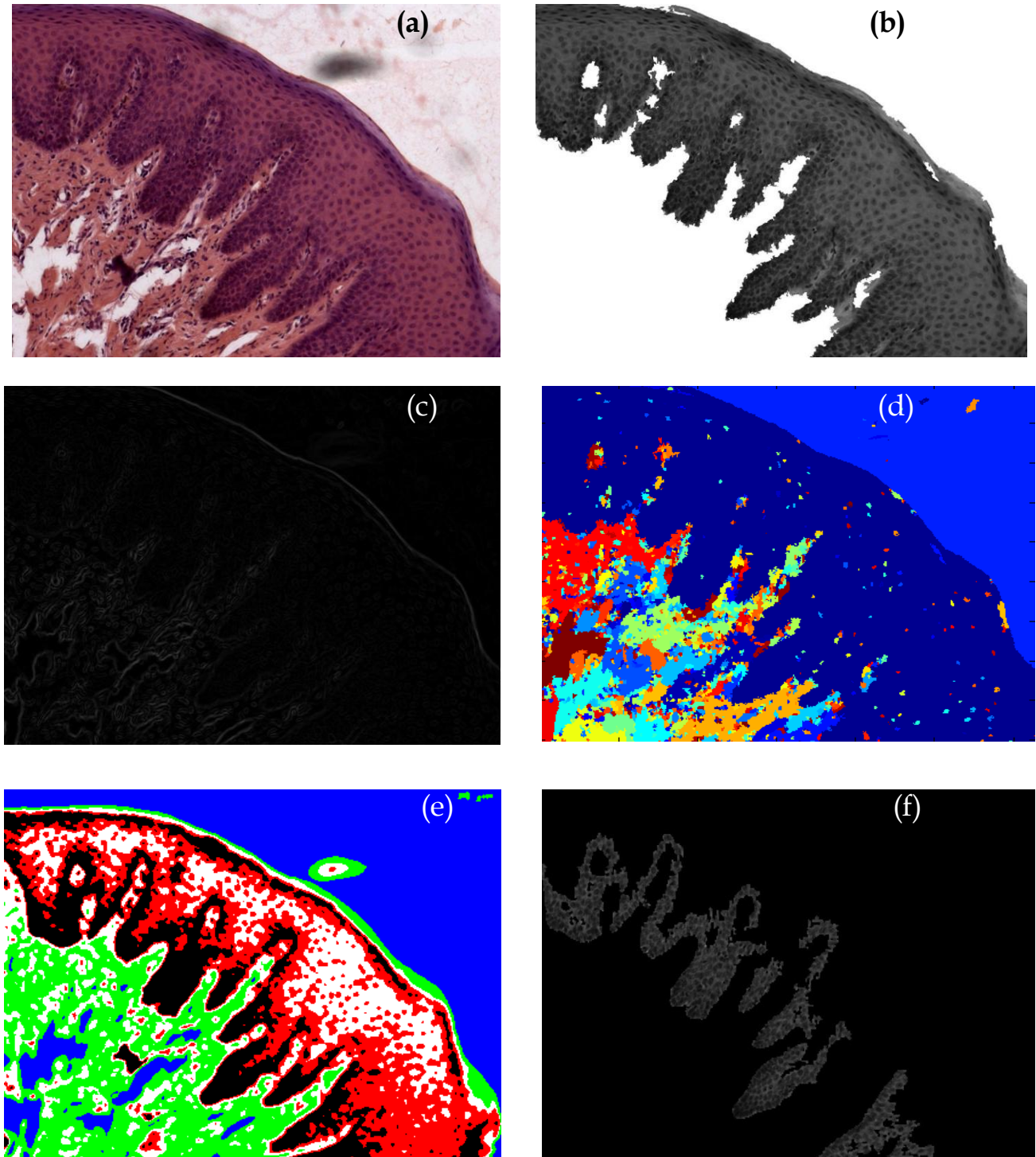
Figure[3.1]: Texture based segmentation of oral mucosa microscopic images

3.2.4 Segmentation of the Sub-Epithelium Layer

The proximity of the sub epithelium layer to the upper edge of the epithelium layer is utilized for segmenting out the region of interest. The epithelium layer is segmented using the Gradient Magnitude based Watershed Segmentation.

Taking the microscopic image (Fig 3.2(a)) as the input, a Gradient Magnitude based Recursive Gaussian Filter is applied on the image with standard deviation σ=2.5 to suppress the variations within the epithelium layer and to highlight the boundaries between the epithelium layer & the connective tissue. Watershed based segmentation is applied on the filtered image (Fig 3.2(c)) with the level parameter of 0.09 & threshold parameter of 0.07 thus segmenting the image into different classes. The classes are arranged in the order of their decreasing areas & for each class, the average intensity value & variance is calculated. The class with the minimum intensity among the first 5 classes corresponds to the epithelium layer (Fig 3.2(d)) & is converted to a mask. The outer (upper & lower) boundaries of this mask are computed & the length of each of the boundaries is computed. Also the variance on each side of the mask is computed & since the connective tissue towards the sub-epithelium layer is more uniform, the boundary on that side is segmented. For all the points on that boundary, a 10x10 window is taken around each pixel & the number of different intensity pixels lying within the mask in that window in the K-means clustered image (Fig 3.2(e)) is calculated. Since mostly the sub-epithelium layer lies near the upper boundary, the maximum count is of the color of the sub-epithelium layer in the clustered image. Hence all the classes near the boundary belonging to that color are segmented from the result of the K-means clustering & dilation using a disc of radius 5 is applied on the classes leading to the segmentation of the subepithelium layer (Fig 3.2(f)).

Figure[3.2] (a) Microscopic Image of Oral Mucosa (b) Epithelium Layer segmented using Gradient Based Watershed Segmentation (c) Result after applying Gradient Magnitude based Gaussian Filter (d) Watershed Transform of the Image (e) Texture Segmentation of original image (f) Segmented Sub Epithelium Layer.

3.3 Analysis of Sub-Epithelium Layer

For each image, 10 samples of 64x64 dimension are extracted from the sub-epithelium layer using 10 seed points & wavelet based texture analysis is performed using Tree Structured Wavelet Transform.

3.3.1 *Wavelet Transform*

Wavelets are a class of functions used to localize a given function in both space and scaling. A family of wavelets can be constructed from a function $\psi(x)$ known as a mother wavelet, which is confined in a finite interval. Daughter wavelets $\psi^{a,b}(x)$ are then formed by translation(b) and contraction(a).Wavelets are superior over traditional Fourier methods in analyzing physical situations where the image contains discontinuities and sharp spikes.

An individual wavelet is defined by

$$\psi^{a,b}(x) = |a|^{-\frac{1}{2}}\psi\left(\frac{x-b}{a}\right) \qquad \text{.... (3.1)}$$

In wavelet analysis the use of a fully scalable modulated window solves the signal-cutting problem. The window is shifted along the signal and for every position the spectrum is calculated. Then this process is repeated many times with a slightly shorter (or longer) window for every new cycle. In the end the result will be a collection of time-frequency representations of the signal, all with different resolutions. Because of this collection of representations we can speak of a multiresolution analysis. In the case of wavelets we normally do not speak about time-frequency representations but about time-scale representations. Hence, wavelet transform captures both frequency and location information.

A wavelet transform decomposes a signal f(x) using the wavelets $\psi^{a,b}(x)$ defined in (1) which form a family of real orthonormal bases .The wavelet coefficients are calculated via

$$c_{a,b} = \int_{-\infty}^{\infty} f(x)\,\psi^{a,b}(x)dx$$

& f(x) can be recovered from the wavelet coefficients using the formula

$$f(x) = \sum_{m,n} c_{a,b}\,\psi^{a,b}(x)$$

When wavelet transform is done in discrete steps, the discrete wavelet transform is obtained, for which there exists an efficient filtering implementation in the real space. Every wavelet corresponds with a high and low pass filter. The 2D discrete wavelet transform decomposes the image using high & low pass filters which are applied along horizontal & vertical directions resulting in 4 subimages. Every image can be subsampled by a factor of 2 since half of the frequencies in the subimage have been removed & hence half the samples can be discarded according to Nyquist's rule. This leads to a representation with an equal amount of pixels as the original image. The corresponding 2D filter coefficients can be represented by $h_{LL}(k,l) = h(k)h(l), h_{HL}(k,l) = g(k)h(l), h_{LH}(k,l) = h(k)g(l), h_{HH}(k,l) = g(k)g(l)$. Where h(k) & g(k) are impulse responses of the lowpass & highpass filters respectively. In the traditional pyramid type wavelet transform, this is repeated iteratively for the low pass subimages. However since in oral mucosa microscopic images, significant information of texture often appears in the middle frequency channels, further decomposition just in the lower frequency region such as for the conventional wavelet transform doesn't help much for distinguishing between

malignant & non-malignant cases. Hence tree structured wavelet transform is employed.

3.3.2 *Tree Structured Wavelet Decomposition*

In the tree structured wavelet transform [9] the decomposition is not only applied to the low frequency subsignals($h_{LL}(k,l)$) recursively.Instead,it can be applied to the output of any filter $h_{LL}(k,l)$, $h_{HL}(k,l)$, $h_{LH}(k,l)$, $h_{HH}(k,l)$ thus resulting in the formation of a quadtree structure. However since it is usually computationally expensive & also unnecessary to decompose all subimages, hence entropy of the subimage is used as a criterion to decide whether a decomposition is needed for that image. Consider a grayscale image X with histogram represented as F(x) = $\{X_1, X_2, X_3 \dots, X_n\}$ where n equals the number of intensity levels in the image (n=256 for grayscale image).Then the probability of each intensity level is defined as

$$p(x_i) = \frac{Number\ of\ pixels\ with\ intensity\ level\ i}{Total\ number\ of\ pixels\ in\ the\ image}$$

. The entropy is calculated using Shannon's entropy defined as

$$e = - \sum_{i=1}^{n} p(x_i) log_2 p(x_i)$$

The idea behind using the entropy is that it quantifies, in the sense of an expected value, the information contained in a message, usually in units such as bits. In our case we use Daubechies 4 tap wavelet (db4)(Fig 3.2) for performing the decomposition. However it may be noted that the choice of wavelets doesn't make much difference & the result is similar even if Haar wavelet is used.
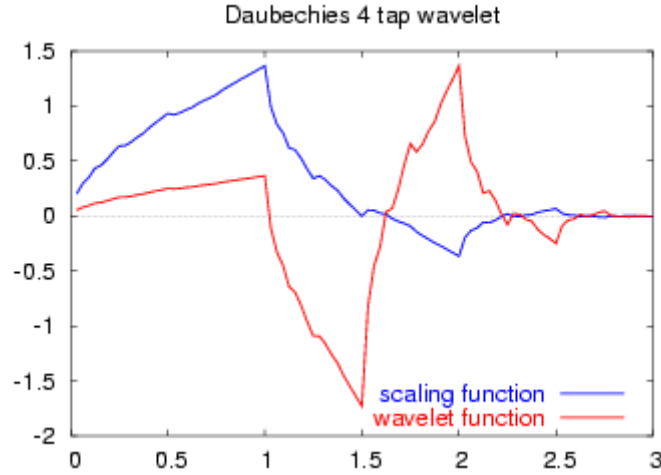
Fig 3.2

The algorithm for tree structured wavelet transform is given as follows

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

(b) Calculate the energy of each decomposed image (children node).

(c) If the entropy of a subimage is significantly smaller than others, we stop the decomposition in this region since it doesn't contain useful information. This step can be achieved by comparing the entropy with the largest entropy value in the same scale among 4 subimages. If $e < Ce_{max}$, stop decomposing this region where C lies between 0 & 1.

(d) If the entropy of a subimage is significantly larger and the size of the image is greater than 16x16 pixels, apply the above decomposition procedure to the subimage. The size of the image is important since if the decomposed channel has a very narrow size, the entropy value may vary widely & feature may not be robust.

The entropy map of the decomposition can be represented by a quadtree structure as shown in Fig 3.3(a) where A, B, C & D represent the 4 subimages. After construction of the tree, we calculate the entropy at its leaves & obtain an entropy function defined on the spatial/frequency domain known as energy map. The total entropies along all the paths are calculated & arranged in decreasing order. First 5 dominant frequency channels with largest entropy values are considered for representing the texture of the subepithelium sample image. As it can be seen from the results in table 3.1, subepithelium in normal state has h_{LL} as the dominant frequency channel while the subepithelium in precancerous condition has h_{HL} as the dominant frequency channel. Moreover the entropy map for the normal layer is found to be same in most of the cases while the entropy map for the pre-cancerous layer varies in different sample images.
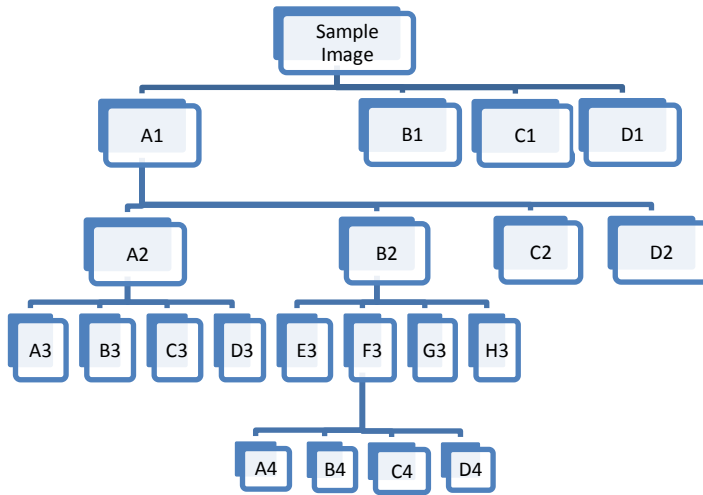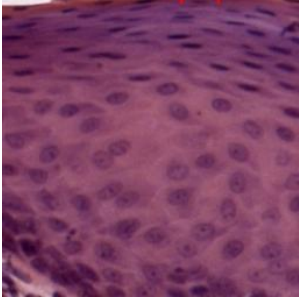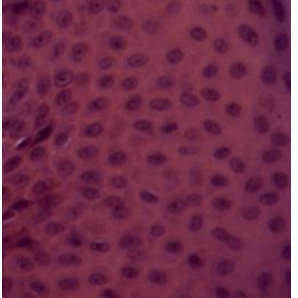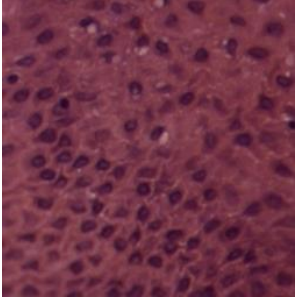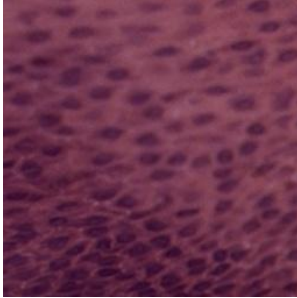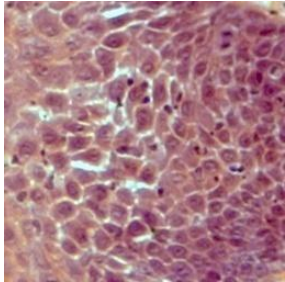

(a)
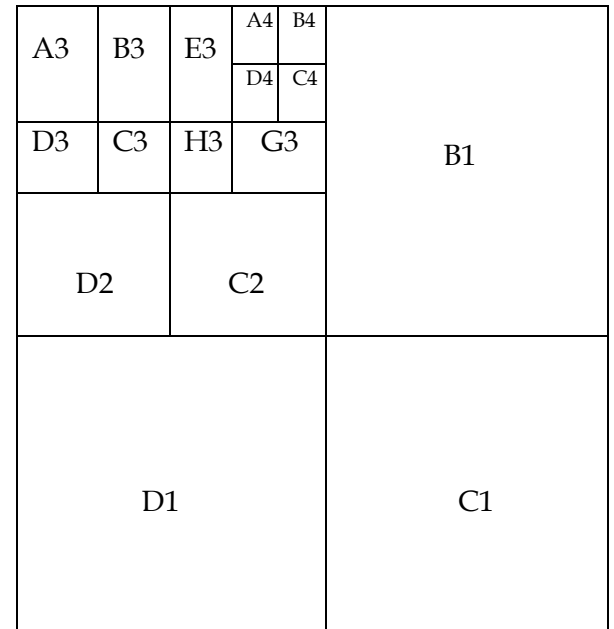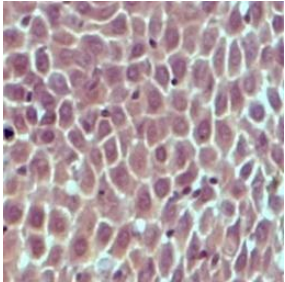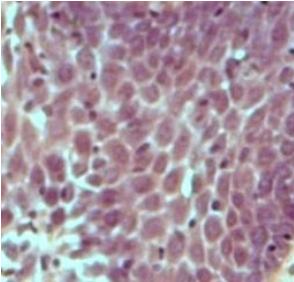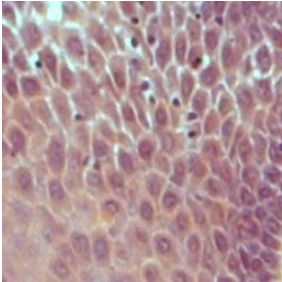
(b)

Fig 3.3(a) Quadtree Representation of Oral microscopic image (b) Channel Decomposition

3.3.3 *Entropy Map of Normal & OSF samples*

Table 3.1: 8 samples of size 256x256 pixels are taken & their entropy map after applying tree structured wavelet decompositions with C=0.15 is shown in the following.

Normal Sub-Epithelium layer			
5 Most Dominant Entropy Channels			
ABBA	AABB	AABA	AABB
ABBB	AADD	AABC	AABA
ABBD	AABA	AADD	AABC
ABBC	AABD	AABB	AABD
ABAA	AADA	AABD	AADD
Malignant Sub-Epithelium layer			
5 Most Dominant Entropy Channels			
ABBB	DDAA	ADDA	DDAA
ABBD	DDAD	ADDB	DDDA
ABCD	DDDD	ACCC	DDAD
ADDA	DDDA	ADDA	DDDD
ABBC	DDAB	ADCB	DDAB

CHAPTER 4

Subepithelium cell segmentation

Segmentation of cells is crucial for analysis of histological images. The Oral Mucosa image segmentation is one of the most difficult segmentation among different types of cell images. This is because of its diversity in the structures and the overlapped cell clusters. The major features of malignancy are related with the nuclei of the cells. Normally, nucleus shape is small and almost round as shown in Fig.4.2 (a). Its intensity is darker than cytoplasm. Malignant or pre-cancerous cells have larger and darker nuclei and tend to cling together in clusters. It is therefore essential to operate a segmentation of the image, to isolate the cells from the rest of the image. The watershed transform[11] has proven to be a powerful and fast technique for both contour detection and region-based segmentation. It is simple and intuitive, can be parallelized, and always produces an entire partition of the image. Watershed segmentation often results in over-segmentation due to intensity variations within both objects and background. Seeded watersheds, where the number of segmentation boundaries is limited to a set of input seeding regions, have been used for segmentation of cell nuclei in images produced by fluorescence microscopy [13]. Fully automatic seeding often results in more than one seed per object, or objects containing no seed at all. To improve the segmentation, the segmentation and tracking can be combined. A popular method that combines the segmentation and tracking is active contours (snakes), previously used for cell segmentation/tracking in e.g. [15]. A significant drawback with the traditional snake is its inability to deal with splitting cells.

Appropriate merging techniques have been devised based on the cell features like circularity, binding energy etc. for correcting the oversegmentation done by

watershed segmentation. The initialization step of the process is probably the most critical step of this kind of transform. A good set of markers will avoid severe over-segmentation. This chapter presents a segmentation technique based on seeded watersheds in following sections.

4.1 Initial Seeding

Adaptive Histogram Equalization is applied on the image for intensity enhancement. As the image intensities of the cells show a range of values partly overlapping with the background intensities, intensity thresholding will not separate cells from background in a satisfactory way. A more effective method for separating objects from background is using variance measures. Object seeds at the initial step are found using the extended h-maxima transform[12] that filters out local maxima using a contrast criterion. All maxima with heights smaller than a threshold level h are suppressed. A low h will result in many seeds, often more than one seed per cell, while a high h will leave some cells without a seed. We take a lower value of h=15 since the merging step below compensates for oversegmentation. All foreground seeds are uniquely labeled using connected component labeling. The seeds obtained from the h-maxima transform are then combined with the thresholded image to obtain the final seed points.

4.2 Watershed Segmentation

The object seed information is used as input for watershed segmentation. Seeded watershed segmentation is applied to the inverse of the original intensity image. The dark edges of the cells are thus interpreted as ridges, and the brighter cells and background as shallow valleys in the watershed landscape. Seeds represent holes in the landscape, and as the landscape is submerged in water, the water will start to flow into the minima of the landscape, creating one catchment basin associated with

each local minima. As the water rises, water from neighboring catchment basins will meet. At every point where two catchment basins meet, a dam, or watershed is built. These watersheds are the segmentation of the image. The watershed image is combined with the thresholded image of the cells to obtain final segmentation. The initial result of seeded watershed segmentation of Fig 4.2(a) using the seeds from Fig 4.2(c) is shown in Fig 4.2(d)

4.3 Merging of fragments

Watershed segmentation often results in over-segmentation. The segmentation is therefore followed by merging steps. This result may contain oversegmented (fragmented) nuclei components arising from image artifacts that must be merged. The guiding factor for making these merging decisions is a 2D mathematical model of the nuclei based on a combination of morphometric features, which includes area, convexity, shape factor, and bending energy.

4.4 Features for Object Modeling

Area: The area (size) of the object, A, is the total number of pixels in the object.

Convexity: The convexity, denoted C, of an object is the ratio of the object area to the area of the convex hull of the object. The convex hull of an object is formed by the well-known Jarvis March algorithm (54). The convexity is typically close to 1 for circular and elliptical objects and less than 1 for concave objects.

Shape factor: Let 'P' denote the perimeter & 'A' the area of the object. The shape factor, denoted U, is defined as

$$U = \frac{P^2}{4\pi A}$$

Bending energy: The boundary pixels of the 2D object are denoted (x_0, y_0), (x_1, y_1), ... (x_{n-1}, y_{n-1}). The bending energy, E, is defined as

$$E = \frac{1}{n} \left(\sum_{i=1}^{n-1} (\theta_{i+1} - \theta_i) \right)^2$$

$$\text{where } \theta_i = tan^{-1} \left(\frac{y_{i+1} - y_i}{x_{i+1} - x_i} \right)$$

The reason to use 2D bending energy is its simplicity and lower computation cost. With the above-defined features, the feature vector for an object is simply $\mathbf{f} = (V, C, U, E)$. This feature set was found to be effective yet minimal for these data.

The score of a fragment is formulated as:

$$S = (A - A') + C - U - E$$

where A' denotes the approximate area of the nuclei calculated in advance. The proposed score captures the intuition that, compared with the fragments of nuclei, a correctly formed nucleus tends to have (a) an appropriate area A, (b) high convexity C, (c) low shape factor U , and (d) low bending energy E.

The fragments are combined only if the final score of the combined structure is greater than the score of each of the individual fragments.
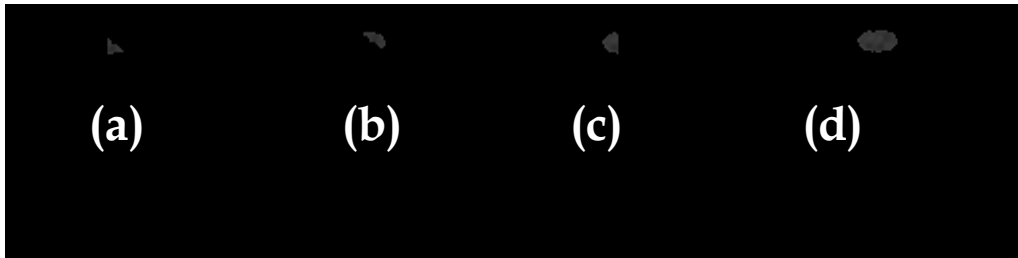

Fig 4.1 3 Fragments formed due to oversegmentation are combined to form a cell
(a) Score=-1.7423 (b) Score=-4.05 (c) Score=-9.40 (d) Score=-1.2643

4.5 Results

(a)   (b)   (c)   (d)   (e)

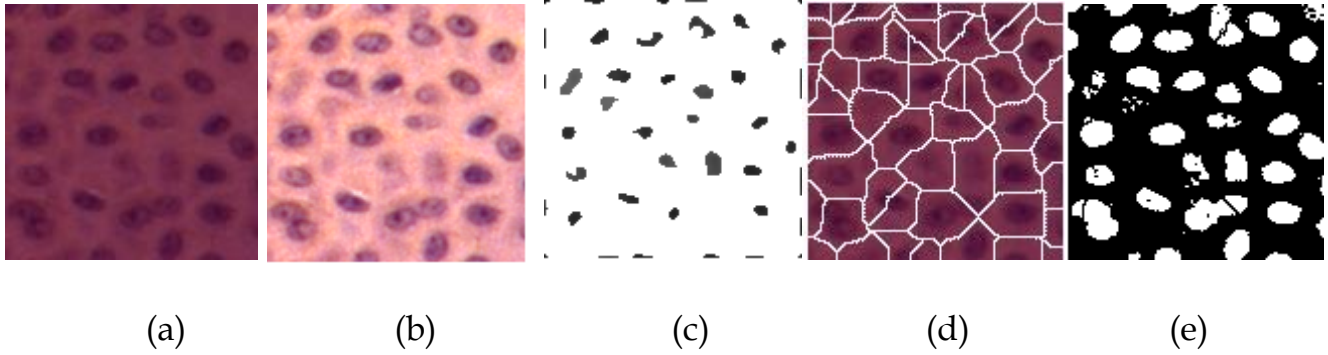Fig 4.2 (a) Subepithelium layer sample (b) Histogram equalization (c) Initial seeds for watershed (d) Result of watershed transform (e)Finally segmented cells

CHAPTER 5

Conclusion

It can be observed from the comparision of entropy based thresholding methods that relative entropy minimization thresholding gives the result most similar to the original image. Hence in all further thresholding operations in our method, we have used cross entropy minimization based thresholding. Initially it was thought of using the entropy based thresholding for segmentation of the subepithelium layer but since the layer is discontinuous and thresholding separates the cells as isolated regions, Gabor filter based texture segmentation was used for subepithelium segmentation. The proposed Gabor filter is capable of segmenting multi texture image & the boundary of each texture is well defined & spatially accurate in segmented image. However the K-means clustering is not very reliable method for merging feature vectors together & the method has to be run 5-6 times to obtain best possible segmentation. Also the number of cluster may vary in certain cases. A possible solution to this problem is use of Fuzzy C means clustering

The tree structured wavelet decomposition has been found to yield similar entropy map in case of normal oral mucosa & the dominant channel is the horizontal low frequency & vertical low frequency channels. In case of malignant mucosa, the entropy map is seen to vary among different datasets & the dominant channel is the horizontal high frequency & vertical low frequency channel. This may be accounted to the increased number of cells with different orientations & sizes in malignant condition as well as the variation in the intensity of different cells within the layer. The watershed based cell segmentation has been found to depend on the seed point initialization. If seed points are properly initialized, the watershed results in perfect segmentation & identifies the weak boundaries between 2 merged cells.

Bibliography

[1]Burkhardt, A., Advanced methods in the evaluation of premalignant lesions and carcinoma of the oral mucousa, Journal of Oral Pathology, Vol. 14, p.p. 751-

[2]R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison-Wesley, First ISE Reprint, 1998.

[3]A. K. Jain, F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp.1167-1186, 1991

[4]D. Clausi, M. Ed Jernigan, "Designing Gabor filters for optimal texture separability, " *Pattern Recognition,* vol. 33, pp. 1835-1849, 2000.

[5]Jianguo Zhang, Tieniu Tan, Li Ma, "Invariant texture segmentation via circular gabor filter", Proceedings of the 16th IAPR International Conference on Pattern Recognition (ICPR), Vol II, pp. 901-904, 2002.

[6] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, A new method for gray-level picture thresholding using the entropy of the histogram, Computer Vision Graphics Image Processing 29,1985, p.p. 273-285.

[7]T. Pun, Entropic thresholding: A new approach, Computer Vision Graphics Image Processing 16, 1981,p.p. 210-239.

[8]Li, C.H. and Lee, C.K., 1993. Minimum cross entropy thresholding. Elsevier Pattern Recognition.

[9]Tianhorng Chang & C-C Jay Kuo: Texture Analysis & Classification with Tree-structured Wavelet Transform, IEEE Transactions on Image Processing,Vol-2,No.4

[10] Gang Lin, Monica K. Chawla, Kathy Olson, John F. Guzowski, Carol A. Barnes, and Badrinath Roysam : Hierarchical, Model-Based Merging of Multiple Fragments for Improved Three-Dimensional Segmentation of Nuclei,Wiley,2004

[11]L. Vincent and P. Soille: Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE Trans. Pattern Anal. Machine Intell., 13(6) 583-598 (1991)

[12]P. Soille: Morphological Image Analysis: Principles and Applications. Springer Verlag (1999)

[13]G. Landini and E. Othman: Estimation of tissue layer level by sequential morphological reconstruction, Journal of Microscopy, 209(2) 118-125 (2003)

[14]C. Wahlby, I.-M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson: Combining intensity, edge, and shape information for 2D and 3D segmentation of cell nuclei in tissue sections, Journal of Microscopy, 215(1) 67-76 (2004)

[15] F. Leymarie and M. Levine: Tracking deformable objects in the plane using an active contour model, IEEE Trans. Pattern Anal. Machine Intell, 15(6) 617-634 (1993)

[16] J.G. Daugman:"Uncertainty relations for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters", Journal of the Optical Society of America A, 1985,vol. 2, pp. 1160-1169.

[17] Dong Liu, Zhaohui Jiang , Huanqing Feng: A novel fuzzy classification entropy approach to image thresholding, Elsevier Pattern Recognition Letters 27 (2006) 1968–1975

[18]Aziz SR. Oral submucous fibrosis: an unusual disease. J N J Dent Assoc 1997;68:17–9.