THE UNIVERSITY OF CALGARY

DIGITAL IMAGE PROCESSING TECHNIQUES FOR QUANTITATIVE

ANALYSIS OF COLLAGEN FIBRIL ALIGNMENT IN LIGAMENTS

by

SUBHASIS CHAUDHURI

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF ELECTRICAL ENGINEERING

CALGARY, ALBERTA

APRIL, 1987

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, "*Digital Image Processing Techniques for Quantitative Analysis of Collagen Fibril Alignment in Ligaments*", submitted by Subhasis Chaudhuri in partial fulfillment of the requirements for the degree of Master of Science.
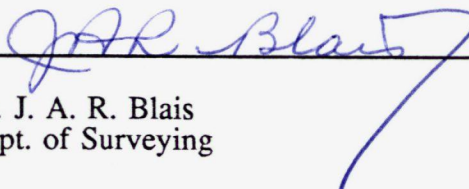
Dr. R. M. Rangayyan - Supervisor
Dept. of Electrical Engineering

Dr. C. B. Frank
Dept. of Surgery

Dr. R. A. Stein
Dept. of Electrical Engineering

Dr. J. A. R. Blais
Dept. of Surveying

Date: 20 April 1987

ii

ABSTRACT

Collagen fibers and their component fibrils make up the protenaceous "backbone" of most tissues and provide the majority of their resistance to tensile loading. Spatial orientation of collagen fibrils is an important factor in determining tissue properties. This is particularly true in ligament tissue, since ligaments must be loose enough to allow joints to move but tight enough to prevent joint surfaces from separating. A normal (healthy) ligament consists of a nearly parallel arrangement of collagen fibrils. In contradistinction fibril distributions in ligament scar are highly disordered. As the ligaments heal, fibril segments are re-aligned, providing greater axial support. In this thesis a method is presented to reproducibly quantify the collagen arrangement in normal ligaments and in ligaments at different stages of healing.

In this method an image made up of a number of line segments oriented at different angles (such as scanning electron micrographs of collagen fibers) may be decomposed into several component images by using Fourier domain directional filtering. Modeling the collagen fibers as being made up of piecewise linear fibril segments, sector filters spanning different angle bands are used to extract component images with fibrils oriented only in the chosen angle bands. A higher angular selectivity is achieved by removing a certain amount of low frequency components. Artifacts are reduced by multiplying the sector filters with a raised

cosine window to give smooth attenuation characteristics. The filtered images are then binarized and quantified in terms of area occupied by fibril segments in the specified angle bands to obtain angular distributions of collagen.

An error analysis is performed to verify the suitability of this procedure by applying it on simulated test patterns. Results obtained from test patterns and from examples of scanning electron micrographs of collagen fibers in normal and healing rabbit ligaments are presented here. Statistical measures, such as entropy, second central moment, and cross-correlation are calculated from these distributions to quantify the axial alignment of collagen in healing ligaments. It is shown that these statistical measures change monotonically with healing time, approaching the values for the normal ligament. This is the first quantitative analysis of the collagen remodeling process.

The above procedure may be applied to injured ligaments which have been treated in different ways (e.g., immobilization, exercise, etc.), and the rate of healing in each case may be determined quantitatively. This information will be useful to physicians in treating ligament injuries by helping to identify the factors which optimize collagen remodeling in the healing situation.

ACKNOWLEDGEMENTS

... to my parents

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Organization of collagen fibers in ligaments

Virtually all connective tissues in the human body are made up of variable forms of a fiber-filled matrix. A fiber, by definition, is a bundle of fibrils[1]. Fibrils consist of various sizes and shapes of chemically distinct proteins known as collagen[2], with augmentation by other fibrous materials such as elastin. There is a complex interaction between these materials and the non-fibrous "ground substance" (water, proteoglycans, other glycoproteins and glycolipids, etc.) in all tissues, giving each tissue relatively unique mechanical properties. As with any composite fiber-reinforced material, the quantity and the quality, as well as the spatial organization of reinforcing fibers in ligaments have considerable influence on their mechanical behavior[3].

Ligaments are highly organized connective tissues that stabilize joints. They normally consist of nearly parallel arrangements of slightly wavy or "crimped" collagen fibers that are attached to bone on both sides of a joint, and serve to guide the joint through its normal motions while preventing its surfaces from being separated. Injuries to ligaments are very common, resulting in this normal, highly ordered structure being replaced by a relatively disordered scar tissue. This scar tissue has many quantitative and qualitative differences from the normal

ligament[4], but the relative disorganization of its constituent collagen fibers may be among the most critical. The loose meshwork of fibrils in the scar tissue may not be able to resist tensile loads within the same limits of deformation as a normal ligament. The injured or healing joint, therefore, may be "loose" or unstable as a result of a ligament injury.

As an injured ligament heals, it is speculated that, like other tissues such as skin, this "loose" meshwork of fibrils will gradually become re-aligned by a process known as remodeling. Thus, at different stages of healing, it may be expected that the distribution of collagen fibrils in the ligament would be increasingly different from the random distribution in early scar tissue, gradually approaching the alignment of a normal ligament.

Although many qualitative and quantitative differences between normal and healing ligaments have been previously described in an animal model[5], the all important quantitative comparison of collagen alignment has not been performed. In order to describe and quantify these organizational differences, a standardized and objective method of assessment has to be developed.

1.2 The image processing problem

An image of a given tissue may have varying numbers of fibrils oriented in different angle bands, covering certain percentages of the total image area. Modeling the fibrils as being made up of piecewise linear segments oriented at various angles, we wish to break down the given image into component images,

each having line segments extracted from various fibrils, oriented within a specified angle band. This could be achieved by frequency domain directional filtering using the Fourier transform.

In this thesis, the image processing requirements for quantification of the surface orientation of collagen fibrils in normal and healing ligament tissue samples along different directions are discussed. The relative fibril-covered area within each angular sector scanned is determined. The collagen distributions so obtained are analyzed quantitatively in terms of fibril alignment at different stages of healing of ligaments by using suitable measures of dispersion and correlation.

In this project, only the surface orientation of collagen fibers seen on freeze-fractured surfaces is considered. The fibers lying in lower layers, which appear silhouetted in the image, are considered as background. It is assumed that the collagen fiber distribution at deeper layers in the ligament is similar to that at the surface.

Depending on the amount and the nature of "ground substance", the grey levels of different fibrils in the image of a tissue sample may be different. Thus, the procedure should be able to extract fibrils with varying grey levels in the component images. Further, some of the images may contain blobs of matrix material which are not directional in the true sense. It is desired that such lumps be filtered out in the component images, as they probably do not provide any significant resistance to tensile stress.

It is found in most images that there is considerable overlap between fibrils oriented at different directions. Such overlap should contribute to each of the corresponding component images. Finally, some of the fibrils in the image may follow smooth curves or twisting bends. The designed procedure should be able to break such fibrils into piecewise linear segments, or should be able to follow smooth bends within the defined range of angles.

1.3 Applications

A procedure as outlined above should find application in studies on ligament healing and growth. In newly injured tissue, for example, the distribution of collagen fibrils is highly random while in a normal ligament these fibrils are mostly aligned with its longitudinal axis. The distribution of fibrils in ligaments which have been allowed to heal for some period of time will be between these two extrema, and the extent of deviation from the normal distribution may help quantify the current stage of healing.

It has been postulated that the rate of healing of injured ligaments depends on the nature of treatment[6]. Histopathological hypotheses, such as - "exercise speeds healing of injured ligaments", or that "immobilization retards the healing process"[4] may be tested by quantitative comparison of the alignment of collagen fibrils at different healing intervals for different treatment modalities.

1.4 Thesis outline

In this thesis, an attempt is made to put forward the basic ideas of directional data analysis, the design of suitable sector or fan filters, and the scope of applications. Brevity in discussing the problems and their solutions has been maintained throughout. Nevertheless, a number of plots and figures are included to substantiate the discussions.

In the current chapter, a brief overview of collagen structure in connective tissues has been presented. The image processing problem has been defined and the desired features of the methods to be developed have been mentioned. Chapter 2 describes the preparation and handling of tissues and the acquisition of images. The concept of directionality of objects in an image is introduced in Chapter 3. Different methods of directional filtering in both spatial and frequency domains are described, and a number of practical applications of directional filters are discussed. Chapter 4 discusses different aspects of the filter design in detail. Emphasis is placed primarily on the significance of using a pre-filter. The component fibrils in the filtered images are submerged in the background, and for quantification purposes these component images need to be binarized. Chapter 5 discusses the applicability of currently available automatic thresholding schemes for this purpose.

The results of the research work are presented in Chapter 6. Error analysis has been performed using test patterns to illustrate the accuracy of the filtering method. Different statistical measures and their usefulness in quantitative interpretation of the images are the topics of Chapter 7. The gradual re-alignment of collagen

fibrils in healing ligaments is demonstrated quantitatively. Various problems encountered, artifacts arising from the filtering procedures as well as tissue preparation, and the computational requirements of the method are discussed in Chapter 8. Future research directions with the technique developed are also outlined.

CHAPTER 2

IMAGE DATA ACQUISITION

2.1 Tissue handling

The images of the ligament samples were obtained from the Department of Surgery at the University of Calgary. The animal model selected was the ruptured and nonrepaired medial collateral ligament(MCL) in the New Zealand white rabbit. Under general anesthesia and with sterile technique, the right MCL was exposed through longitudinal medial incisions in the skin and fascia. The right MCL was completely ruptured as described in references[5, 7] by passing a 3-O braided steel wire beneath it, and failing the ligament with a strong upward pull on both ends of the suture. No repair of the ligament was performed and the skin was sutured. The left MCL was unruptured and served as a normal control.

The injured MCL was allowed to heal for a scheduled period of healing with the animal being allowed normal unrestricted cage activity. The animal was then sacrificed by intravenous injection of 375mg of phenobarbitol, and the healing (right) and normal control (left) MCLs were harvested. The right and left MCLs were exposed through medial incisions in the skin and fascia. The MCLs were fixed *in situ* by dropping a fresh solution of 2.5% gluteraldehyde in 0.1 M cacodylate buffer with pH 7.4 onto their surfaces. They were then removed at their insertions, placed in 2.5% gluteraldehyde in 0.1 M cacodylate buffer with pH 7.4

7

for three hours and dehydrated in increasing concentrations of ethanol / H_2O (30%, 50%, 75% and 100%). Each fixed and dehydrated ligament was then frozen quickly in liquid nitrogen and fractured longitudinally to expose the internal collagen fiber arrangement along its length. The fractured tissue was then critically point dried, aligned, mounted and sprayed with gold / palladium. In each case, the longitudinal axis of the tissue was distinguishable at low magnification, to allow orientation of high magnification photographs relative to that axis.

2.2 Image acquisition and digitization

Specimens were viewed under a Hitachi S-450 scanning electron microscope. This model of microscope allows the specimen to be moved in both x and y axes. Initially, the mounted tissue was viewed through the monitor at a very low magnification (say, 300x or 400x) and by adjusting the directional alignment vernier in the microscope the longitudinal axis of the ligament was aligned with the vertical axis of the photograph. Once proper alignment was achieved, the magnification was increased to the desired level (typically 7K), and by adjusting x and y axes verniers in the microscope the co-ordinates of the tissue sample to be photographed were suitably chosen. In order to randomly photograph each ligament, pairs of x and y co-ordinates were obtained using a random number generator. A number of photographs (typically 10) were taken accordingly from the same tissue sample but at different locations. (It should be noted that the total area imaged is an extremely small portion of the given ligament sample, far less

than 1%. Also, an individual micrograph represents only a small fraction of the periodic "crimping" of the collagen fibers). Similarly, several photographs of each of the healing and normal control MCLs were taken under the same magnification. Some of the ligament samples were also photographed at different magnifications to test the sensitivity of the designed procedure to varying magnifications.

Each of the photographs was digitized into a 256 x 256 matrix by using a Fairchild CCD-3000 camera with a Nikkor 55mm macro lens under the same conditions of aperture and illumination, and a Colorado Video Inc. CVI 274 frame grabber attached to the VAX 11/750 research computer in the Department of Electrical Engineering. The camera uses advanced CCD (charge-coupled device) technology, and provides wide dynamic range and zero geometric distortion.

CHAPTER 3

DIRECTIONAL CHARACTERIZATION OF IMAGES

3.1 Introduction

It is often useful to measure and analyze the directionality of an image. Such measurements or analyses are usually interpreted in terms of a dominant direction of orientation and the degree of directionality. Different grey level statistics such as spatial moments and co-occurrence matrices may be used to represent the directional nature of an image. Using suitable image processing techniques, image features along certain directions may also be enhanced or suppressed. Such processing may be performed either in the space domain by using convolution masks, or in the frequency domain by using sector filters. In general, spatial operations are local, while frequency domain operations are global in nature.

In this chapter, different methods of analyzing the directionality of image features are discussed. Some applications of directional analysis are reviewed.

3.2 Spatial moments

Spatial moments of an image have been used to determine its principal axis[8]. In analyzing fiber images, this method may be helpful in finding out the dominant angle of fiber alignment.

It can be proved that the moment of inertia of an image $f(x,y)$ is the minimum when the moment is taken about the centroid $(\bar{x},\bar{y})$ of the image. The moment of inertia of the image about the line $(y - \bar{y})\cos\theta = (x - \bar{x})\sin\theta$ passing through $(\bar{x},\bar{y})$ and having a slope $\tan\theta$ is given by

$$m_\theta = \sum_x \sum_y \left[(x - \bar{x})\sin\theta - (y - \bar{y})\cos\theta\right]^2 f(x,y). \qquad (3.1)$$

In order to make m_θ independent of the choice of co-ordinates, let us select the centroid of the image as the origin. Then, $\bar{x} = 0$ and $\bar{y} = 0$, and the equation (3.1) becomes

$$m_\theta = \sum_x \sum_y (x \sin\theta - y \cos\theta)^2 f(x,y)$$

$$= m_{20}\sin^2\theta - 2m_{11}\sin\theta\cos\theta + m_{02}\cos^2\theta \qquad (3.2)$$

where m_{ij} is the $(i,j)^{th}$ order moment of the image and may be given as

$$m_{ij} = \sum_x \sum_y x^i y^j f(x,y). \qquad (3.3)$$

By definition, the moment of inertia about the principal axis is the minimum. Differentiating equation (3.2) with respect to θ and equating to zero gives

$$m_{20}\sin2\theta - 2m_{11}\cos2\theta - m_{02}\sin2\theta = 0 \qquad (3.4)$$

or

$$\tan 2\theta = \frac{2m_{11}}{(m_{20} - m_{02})}.\qquad (3.5)$$

By solving this equation we can find the slope or the direction of the principal axis.

If the input image consists of directional components along an angle ϕ only, then $\phi \approx \theta$. If there are a number of directional components at different angles, then θ represents their weighted average direction. Evidently, this method cannot detect the existence of components in various angle bands, and is thus inapplicable for the analysis of healing tissues. Also, this method cannot quantify the fibril-covered areas in the various angle bands.

3.3 Grey level co-occurrence matrices

In studies related to textural analysis of an image, some statistical measures derived from the grey level co-occurrence matrix have been used to determine the directionality of the textural primitives[8,9]. If $\delta = (\Delta x, \Delta y)$ is the spatial displacement, L is the total number of grey levels in the image and P_δ is an $L \times L$ matrix whose $(i,j)^{th}$ element represents the number of times that a pixel having grey level i occurs in position δ relative to another pixel having grey level j, then P_δ is called the grey level co-occurrence matrix of the image for the displacement δ. P_δ may be normalized with respect to the total number of point pairs in $f(x,y)$ for the displacement δ. The size of P_δ depends on the number of grey levels L present in the image, and does not depend on its size.

The matrices P_δ for various values of δ provide useful information about the spatial distribution of grey levels in the image. For example, if the diagonal concentration in P_δ is high for $\delta = (1,0)$ then the image possesses directional components along the horizontal axis. Similarly, a high concentration of elements along the main diagonal for $\delta = (0,1)$ or $(1,1)$ would indicate the presence of parallel streaks at 90° or 45° in the input image.

An estimate of the directional contrast in an image is given by

$$D = \frac{1}{1 + \sum\limits_{i=1}^{L} \sum\limits_{j=1}^{L} (i - j)^2 P_\delta(i,j)} . \qquad (3.6)$$

The higher the value of D, the greater is the number of directional components present in the image. For example, if an image consists of rectangular line segments only, then $D \approx 1$. It may be mentioned here that D represents the directionality along the angle determined by the chosen value of δ. Different values of δ have to be used to represent the directionality of components along different angles. This method suffers from a number of problems. Since δ can have only a few discrete values, analysis of images will be restricted to a few selected directions only. Further, the connectivity of line segments is not checked for larger values of δ. This method gives only approximate and relative measures of directionality.

3.4 Directional convolution masks

Convolution masks may be used in the spatial domain to enhance or suppress directional components in the processed image[8, 10, 11]. This procedure has been used for enhancement of remotely sensed images[12, 13]. A 3 x 3 convolution mask may expressed as

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

where the parameters a to d and f to i are the weights to be applied to the neighboring pixels of the central pixel corresponding to the weight e. After convolution, the $(i,j)^{th}$ (central) pixel value in the image is replaced by

$$f(i,j) = \sum_{k=1}^{3} \sum_{l=1}^{3} M(k,l) f(i+k-2, j+l-2). \tag{3.7}$$

If all the weights are equal, this convolution process is the same as mean filtering.

Directional convolution masks may be constructed by proper choice of the weighting coefficients. If M is defined to be

$$M_v = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$$

then it suppresses the components along the vertical direction. Similarly,

$$M_h = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

removes the directional components along the horizontal direction. In order to filter out the components along 45° and 135°, the following masks may be used

$$M_d = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} \quad or \quad \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & -1 \\ 0 & -1 & -1 \end{bmatrix}$$

It is obvious that this method filtering may be performed along a few selected directions only. Use of larger masks allows us to increase the directional selectivity to some extent. However, as the size of the mask grows, the amount of computation increases and the accuracy of the method decreases because of the finite number of data points available in the discrete image space. Also, contrary to our need, this kind of convolution mask is a notch (rejection) filter. Hence, this method is not suitable for the purpose of quantification of collagen alignment in ligaments.

3.5 Frequency domain directional filtering

The two-dimensional Fourier transform (FT) of an image containing a straight line is a sinc function confined to a direction orthogonal to the direction of the line[14]. If an image is composed of different lines at different directions, its FT will have components at the corresponding orthogonal directions. If we extract the

components lying within the angle band $(\theta_1 + 90^o)$ to $(\theta_2 + 90^o)$ in the transform domain and then take the inverse FT, we could expect to extract from the original image, components lying within the angle band θ_1 to θ_2 only. Thus, to decompose a fiber image, we could take its FT and apply filters that form sectors in the frequency domain spanning the desired angle bands, and apply the inverse FT to the results[15].

The properties of the Fourier transform[14,16,17] that are of relevance here are linearity and rotation. Linearity implies that the FT of an image made up by adding component images is simply the sum of the FTs of the individual components. Mathematically, it may be represented as

$$FT\,[\,g_1(x,y) + g_2(x,y)\,] = FT\,[\,g_1(x,y)\,] + FT\,[\,g_2(x,y)\,]. \qquad (3.8)$$

The rotational property states that when an image is rotated by a certain angle ψ in the spatial domain, the FT gets rotated by an equal amount in the frequency domain. Let $G(R,\theta)$ be the Fourier transform of the image $g(r,\phi)$, where the two functions are represented in polar co-ordinates. Then the rotational property may be written as

$$FT\,[\,g(r,\phi+\psi)\,] = G(R,\theta+\psi). \qquad (3.9)$$

An ideal directional filter consists of a sector filter in the frequency domain spanning an angle band θ_1 to θ_2 with the dc point in the folded spectrum as the vertex. The ideal filter may be expressed as

$$H(r,\theta) = 1 \quad \text{for } \theta_1 \leq \theta \leq \theta_2$$

$$= 0 \quad elsewhere.$$

The angular bandwidth $[\theta_1,\theta_2]$ may be chosen according to the requirements. Figure 3.1 shows an example of an ideal sector filter. Such ideal filters, however, result in artifacts in the component images[15]. The next chapter describes modifications and preprocessing required to obtain acceptable results.

It may be argued that a quantitative estimate of fibril-covered area along different angle bands in the input image could be obtained from the frequency spectrum by considering the angular distribution of spectral energy itself; that is, that the fraction of the total spectral energy lying within different sectors could give a relative distribution of fibril-covered areas along different angle bands. But this is not true, because the energy of such components will be dependent upon not only the area of the corresponding line segments, but also upon the amplitudes or grey levels of the constituent fibrils. Further, while most of the energy in the frequency spectrum is usually concentrated in the low frequency zone, the definition of angle in the Cartesian co-ordinate system (on which the discrete FT is computed) is particularly poor in the low frequency zone. Also, due to the finite width of the line segments, the spectral energy is not fully concentrated at the corresponding Fourier domain angle. As the line segments grow in width, leakage of spectral energy into other angular sectors increases. Hence, an analysis of the angular distribution of energy in the Fourier domain does not necessarily reflect the
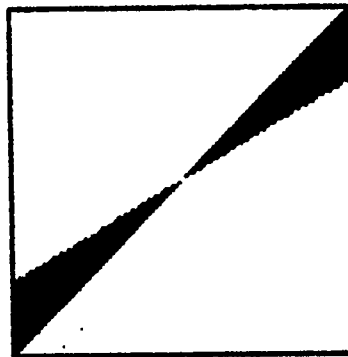
Fig. 3.1 : Example of an ideal 2-D sector filter for the angle band $30^o - 45^o$ (in Fourier space). The shaded region represents the passband with unity gain and the white region is the stopband with zero gain.

true distribution of areas of different directional components having different grey levels and varying fiber widths. The next chapter describes methods to overcome these difficulties.

3.6 Applications of directional analysis

In directional filtering, the objects (or signals) that are aligned at a particular direction are allowed to pass through the filter while the objects lying at all other orientations are suppressed. The concept of directional filtering is not very new in the literature. This principle was used more than two decades ago in extracting desired signals from high velocity noise in geophysical data processing[18]. Some of the applications of directional data analysis are briefly discussed here.

Optical Fourier analysis has been applied to evaluate the difference between normal and pathologically changed bone tissues[19]. The radial distribution of light energy in the diffractogram contains information about the sizes and the distances between the collagen fibers within the bone tissue, while the angular distribution gives the relative position of these osteological structures. These diffraction patterns differ markedly for various pathological changes in the bone tissues. Different distance measures evaluated from the spectra have been used to distinguish between normal and osteopetrotic bone tissues.

The principle of directional filtering has been used in the filtering of geophysical data[18,20] and in designing fan filters[21,22]. Similar methods have been used in wide band velocity filtering, which makes it possible to process a

seismic record-section in such a way that all seismic events with dips in a given range are preserved with no alteration over a wide frequency band, while the events with dips outside the specified range are severely attenuated[18]. However, all these methods employ the implementation of the filter transfer function in the space domain. If we wish to extract components along different angle bands, the input data will have to be processed separately for individual cases. The Fourier transform being a global process may be used very efficiently.

Optical diffractograms have also been used to filter out high velocity noise in seismic data[23-25] and for directional filtering of aeromagnetic maps[26]. When coherent light is passed through a transmissive recording of data section, the recorded signals act as an optical grating to produce a Fraunhofer diffraction pattern which is the two-dimensional FT of the section itself. With suitable lenses the diffraction pattern can be converted into an image of the original section. By obstructing portions of the pattern corresponding to particular frequencies or dips in the section, one can remove such frequencies or dips from the reconstructed image[23]. The optical processing is instantaneous and can be monitored easily. However, it needs sophisticated instrumentation and suffers due to the poor dynamic range of the photographic registration of the data section.

Directional filters find applications in image coding[27]. The low frequency components corresponding to the uniform regions in the image may be coded in the transform domain. The high frequency components may be decomposed into many directional components, each of them presumed to contain edges belonging

to a limited interval of directions. Edge information due to a line segment in an image resulting from such filtering is located at the zero crossings in the spectrum, taken in the direction orthogonal to that in the space domain. Image decoding is performed by synthesizing the directional edges from the coded zero crossing information and then superposing it on the image obtained from the low frequency components. Efficiency of this coding scheme depends on the directional nature of the encoded image.

Directional filtering has also been used for selecting features to describe textures[28]. Its application to the representation of textures is justified by the fact that directional edge information plays an important role in texture discrimination. Various features evaluated from the directional edgeness of the image can be used to classify different textures.

CHAPTER 4

FILTER DESIGN

4.1 Data windowing

Initial attempts towards directional decomposition of collagen fiber images using ideal sector filters[15] indicated the feasibility of the approach as introduced in the previous chapter. However, the results suffered from a number of artifacts (e.g., ringing artifacts, smearing at both ends of the fibril segments and interference from adjacent angle bands) due to the use of ideal filters as shown in Figure 3.1, and the absence of data windowing, among other reasons. Improved filtering and data windowing procedures[29] led to better results.

The size of the images analyzed is typically 256 x 256 pixels. For improved frequency resolution, the image is placed at the center of a 512 x 512 array with the rest of the points set to the mean grey level of the image, instead of padding with zeros. The image pixels are then multiplied with a modified Blackman-Harris window[30, 31] to reduce spectral leakage in computing the FT of the image. The modification to the Blackman-Harris window, as shown in Figure 4.1, prevents the actual image data points from getting multiplied by the window coefficients. If the image data points are directly windowed, the points lying near the periphery of the image will be multiplied by very small coefficients and fibril segments near the image boundaries will be suppressed in the filtered component images. It may be

mentioned here that the fibril segments appear amidst a reasonably uniform background, and hence the data points in the filtered images cannot be multiplied by the reciprocal of the window coefficients to take account of effects of windowing on the input data. Any attempt to do this would enhance the darker background at the periphery appreciably and may even suppress the fibril segments existing in the filtered image.

The usage of such a window is justified by experimentally verifying its performance. A set of data points representing a cosine function and shifted by a dc value was polluted by adding a zero mean random Gaussian noise. The SNR of the data sequence was 17 dB and the length of the sequence was chosen to be 256. The data sequence was placed in the center of a 512 point FT array. Figure 4.2 shows the data sequence. The power spectral density function (PSDF) for this data sequence was evaluated by using 512 point FT for three cases.

(1) In the first case, no modification to the input data was performed (equivalent to the use of a rectangular window) and the array was padded with zeros. The resulting PSDF is given in Figure 4.3.

(2) In the second case, the input data were padded with the mean value of the data sequence and then the data points were multiplied by the coefficients of the Blackman-Harris window[30]. The PSDF of the windowed data sequence is given in Figure 4.4.

(3) Lastly, the input data sequence was padded with the mean value of the sequence and only the padded points were multiplied by the modified
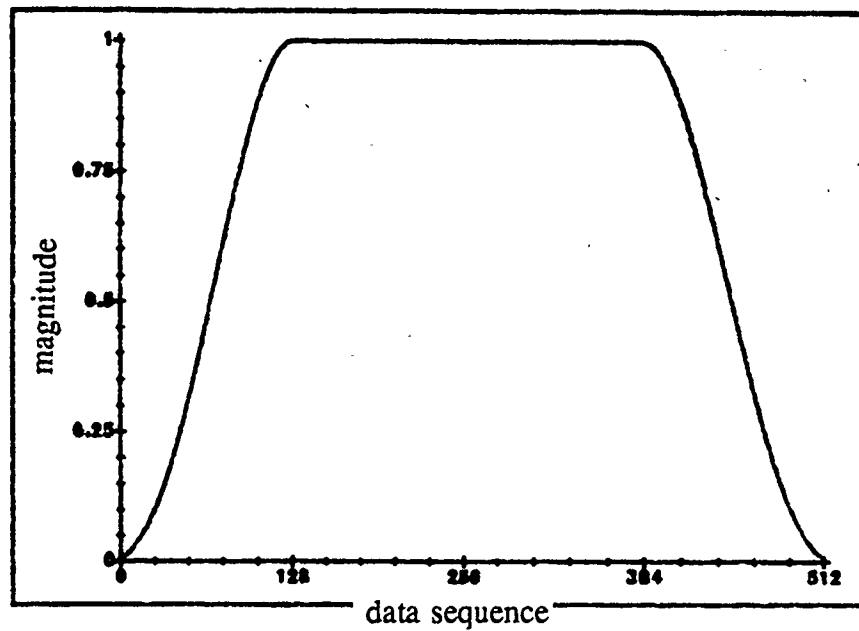
Fig. 4.1 : Modified Blackman-Harris window for 512 sample points. The flat region of 256 points at the center prevents the actual image data from being windowed. The data points are multiplied both rowwise and columnwise by the square root of this function.
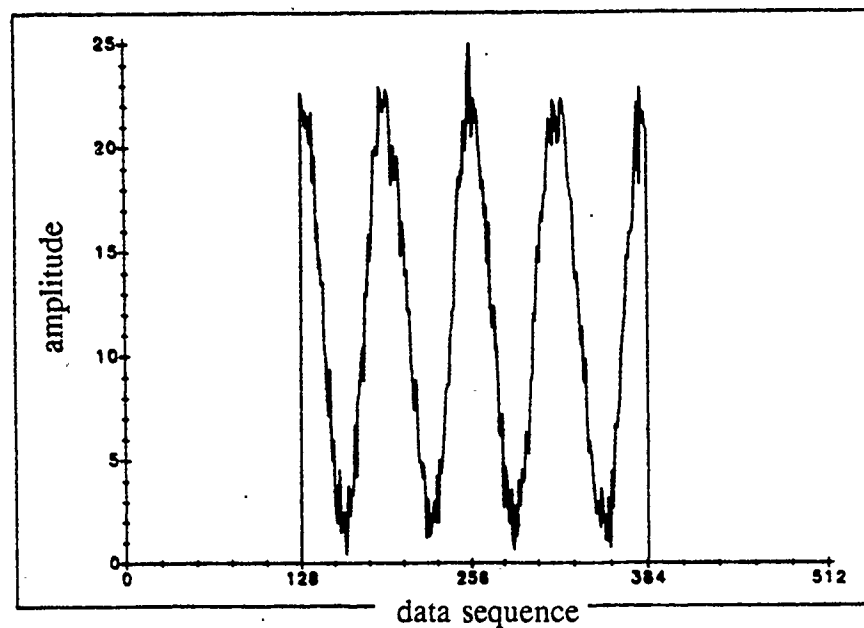
Fig. 4.2 : Example of a sinusoidal data sequence of 256 sample points with added zero mean random Gaussian noise. The SNR is 17 dB. The sequence is given a dc shift.
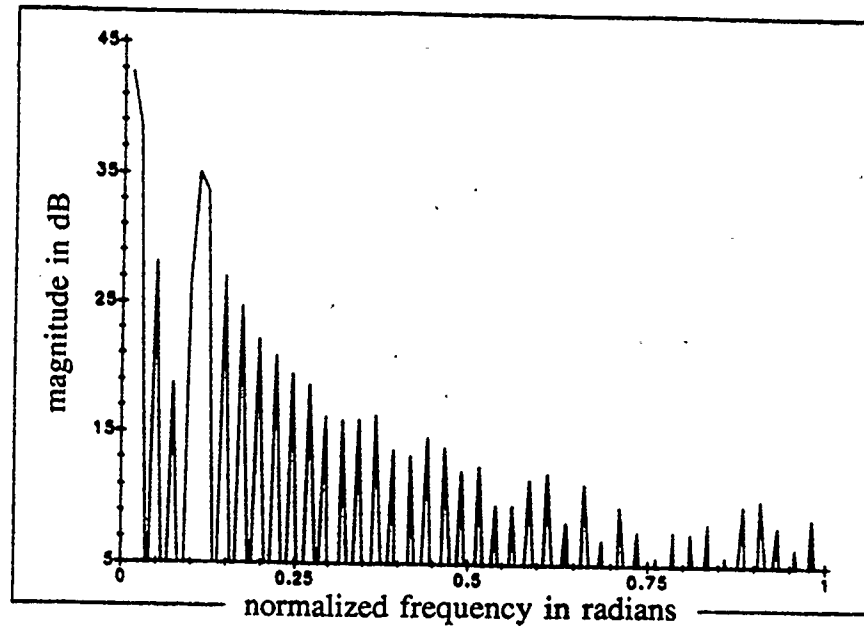
Fig. 4.3 : PSDF of the noisy data sequence using rectangular window. The signal is barely visible in the spectrum.
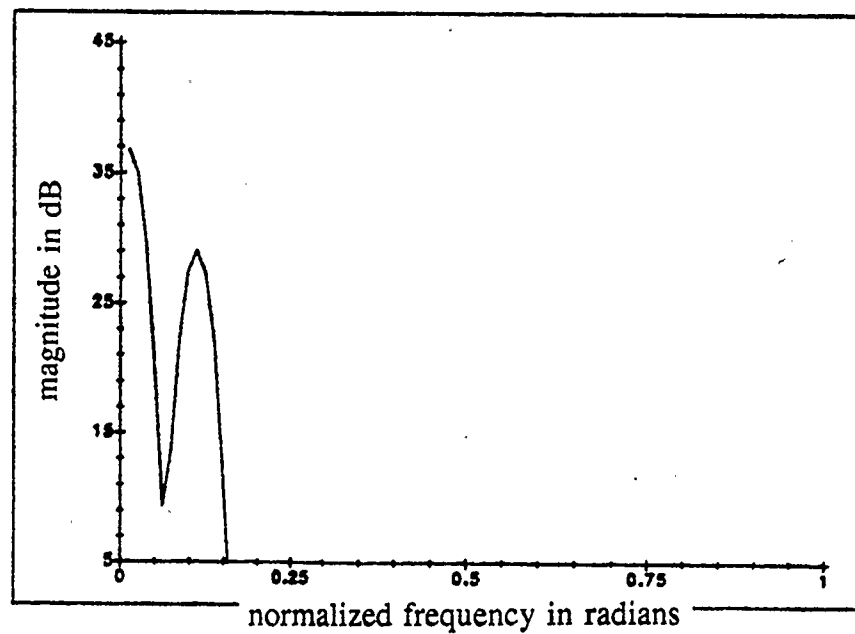
Fig. 4.4 : PSDF of the data sequence using Blackman-Harris window. Notice the high side lobe suppression in the spectrum.

Blackman-Harris window shown in Figure 4.1. The resulting PSDF is given in Figure 4.5.

It is apparent from Figures 4.3 to 4.5 that the Blackman-Harris window gives the best results as the frequency component present is highly distinguishable from the spurious noise peaks. The PSDF obtained by using the rectangular window is very noisy and the signal is almost buried in the noise. Use of the modified Blackman-Harris window suppresses the spectral noise appreciably and one can easily distinguish the signal in the spectrum from the noise peaks. Although the result obtained by using the modified window has somewhat inferior side lobe suppression characteristic compared to that obtained by using the conventional Blackman-Harris window, it retains higher spectral resolution.

The effects of using such a window on an image were studied using a real fiber image consisting of different fibril segments with varying lengths, widths, grey levels and orientations. The fiber image is given in Figure 4.6. The FT of this image using a rectangular window is given in Figure 4.7. Figure 4.8 shows the FT of the image using the modified Blackman-Harris window. As seen in Figures 4.7 and 4.8, spectral leakage is reduced by using the latter window. Hence, it may be inferred that the use of the modified Blackman-Harris window helps in reducing spectral leakage to some extent, and at the same time retains some of the advantages of not windowing the input data.

27
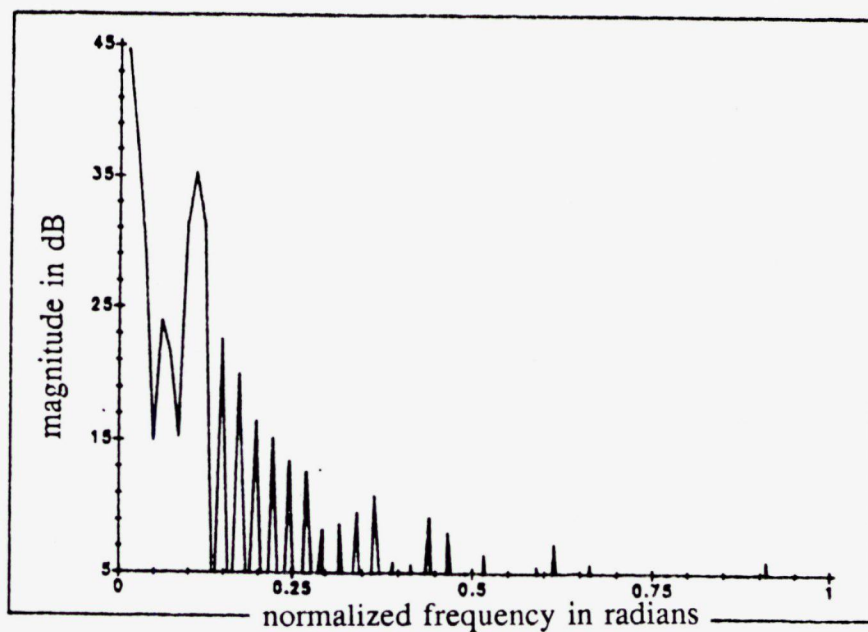
Fig. 4.5 : PSDF of the data sequence using modified Blackman-Harris window. The presence of the frequency component in the signal is highly perceptible.

Fig. 4.6 : A representative image showing collagen alignment in a normal ligament.
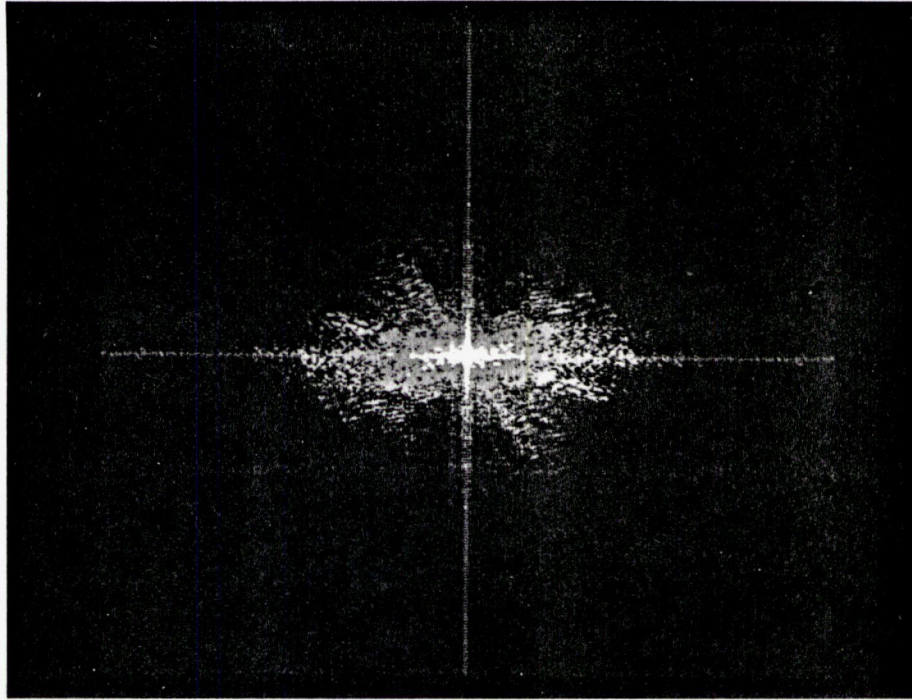
Fig. 4.7 : Fourier spectrum of the image in fig. 4.6 using a rectangular window.
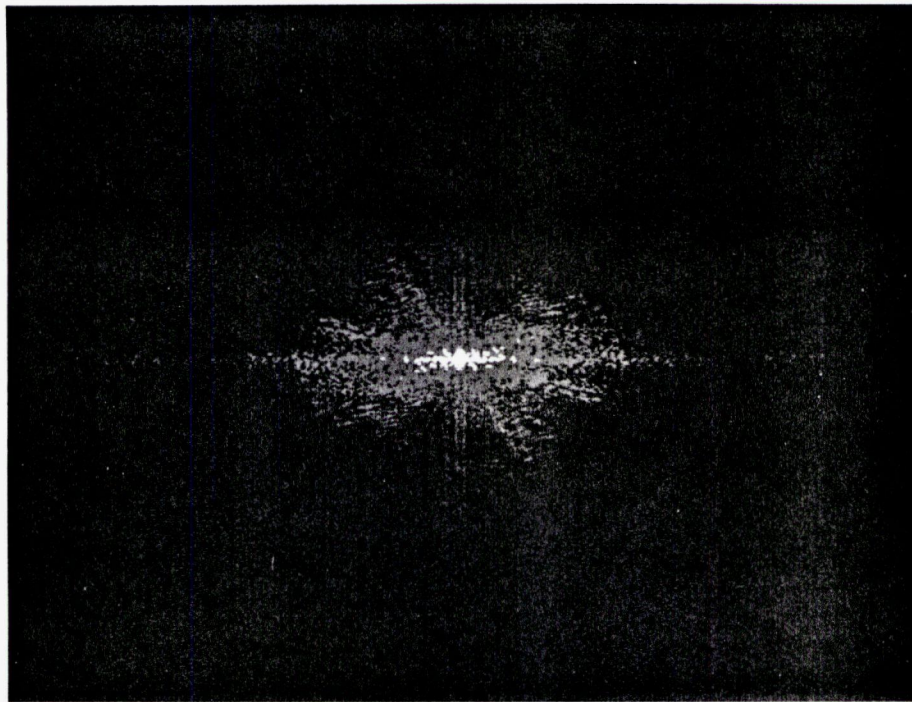
Fig. 4.8 : Fourier spectrum of the image in fig. 4.6 using the modified Blackman-Harris window.

4.2 Prefiltering of the spectrum

Let us assume that an image $f(x,y)$ consists of a rectangular line segment of length Y, width X (given $X < Y$) and grey level G only. The image may be mathematically expressed as

$$f(x,y) = G \; rect(x/X) \, rect(y/Y) \qquad (4.1)$$

where

$$rect(k/K) = 1 \quad for \, |k| \leq K/2$$

$$= 0 \quad elsewhere.$$

The power spectral density function (PSDF) of the image is given by

$$|F(u,v)|^2 = G^2 X^2 Y^2 \left| \frac{sin(\pi u X) sin(\pi v Y)}{(\pi u X)(\pi v Y)} \right|^2 . \qquad (4.2)$$

The PSDF may be expressed in polar coordinates with the change of variables $u = \rho \cos\theta$ and $v = \rho \sin\theta$ as

$$|F(\rho,\theta)|^2 = G^2 X^2 Y^2 \left| \frac{sin(\pi \rho X \cos\theta) sin(\pi \rho Y \sin\theta)}{(\pi \rho X \cos\theta)(\pi \rho Y \sin\theta)} \right|^2 . \qquad (4.3)$$

From equation (4.3) we may derive two useful functions, namely, the cumulative radial distribution of energy $\psi(R)$ and the angular distribution of energy $\phi_r(\theta)$. Using the conjugate symmetry of the FT for real input data, they may be given as

$$\psi(R) = 2 \int_0^R \int_0^\pi |F(\rho,\theta)|^2 d\theta d\rho \qquad (4.4)$$

and

$$\phi_r(\theta) = 2 \int_r^\infty |F(\rho,\theta)|^2 d\rho \quad \text{for } \theta \in [0,\pi] \qquad (4.5)$$

where r is the cutoff frequency up to which the low frequency components have been removed from the spectrum.

From equations (4.3) and (4.5) it is evident that the function $\phi_r(\theta)$ has two local maxima corresponding to $\theta = 0$ and $\pi/2$, one of them being the principal maximum and the other one being the secondary. The relative weights of these two peaks depend on the ratio Y/X. The higher the value of Y/X, the greater is the amplitude difference between the peaks and the greater is the amount of energy concentrated in a particular direction. Equation (4.5) was evaluated numerically and the function $\phi_{r=0}(\theta)$ is plotted in Figure 4.9 for different values of the ratio Y/X. From these plots, it may be inferred that the directionality of an individual line segment at a given angle depends only upon its length-to-width ratio. Also, it is important to note that even for higher values of Y/X, we have some spectral energy at all θ. This is even more true in the case of the discrete FT which suffers appreciably from spectral leakage. Hence, any attempt to extract a line segment directed at an angle β by constructing a sector filter $(90^o + \beta) \pm \delta\beta$ will lead to loss of some spectral information and may eventually bring in some artifacts in the filtered image.

Now, consider the case of an image consisting of two non-overlapping rectangular line segments of identical dimensions with width X, length Y and grey-levels G_1 and G_2 respectively, one of them being aligned along the vertical axis of the image and the other one being rotated by an angle θ_k from that axis. Using the linearity and rotational properties, the FT of this image may be expressed as a superposition of two individual distributions

$$F_2(\rho,\theta) = G_1 F_1(\rho,\theta) + G_2 F_1(\rho,\theta-\theta_k) \qquad (4.6)$$

where

$$F_1(\rho,\theta) = \frac{sin\,(\pi\rho X\cos\theta)sin\,(\pi\rho Y\sin\theta)}{(\pi\rho\cos\theta)(\pi\rho\sin\theta)}. \qquad (4.7)$$

In other words, $F_1(\rho,\theta)$ is the FT of a rectangular line segment of unit grey level in polar coordinates.

From equations (4.5) and (4.6) it is evident that the corresponding angular energy distribution function $\phi_r(\theta)$ has two principal maxima for $\theta = 0$ and $\theta = \theta_k$. A minimum lies between them at $\theta = \theta_m$. For example, if the line segments have identical grey levels (i.e., $G_1 = G_2$) then $\theta = \theta_k/2$. An index of angular separability of the two distributions (due to two line segments in the composite image) has been defined in reference[32] as

$$S = \frac{\phi_r(0)+\phi_r(\theta_k)-2\phi_r(\theta_m)}{\phi_r(0)+\phi_r(\theta_k)}. \qquad (4.8)$$

If the value of S is high, it is likely that the two line segments could be separated

in the filtered images with reduced artifacts. The value of this index depends on the relative grey levels and the dimensions of the individual line segments. Evaluation of this index involves complex mathematics and may be found in reference[32]. However, some trends of the expected behavior of the angular separability are as follows :

(1)   for a given value of θ_k, the narrower the line segments (i.e., the higher Y/X), the better is the separability, and

(2)   distributions of quasi-square segments are difficult to separate, whatever may be the angle between them.

So far nothing has been mentioned about the implication of the parameter r in equation (4.5) and it has been assumed to be equal to zero. What happens when we obliterate a circular region at the center of the spectrum? The angular energy distribution function $\phi_r(\theta)$ was calculated with different values of r for a simulated test pattern having line segments with varying lengths, widths, grey levels and orientations (given in Figure 4.10), and is plotted in Figure 4.11. From these plots it may be observed that the angular separability increases with higher values of r. But, a certain amount of spectral energy has to be retained for a meaningful reconstruction of the filtered component images, and for the artifacts due to high frequency noise to be at a relatively low level. The cumulative radial energy distribution function $\psi(R)$ is given in Figure 4.12 for the fibril image in Figure 4.6. It shows that most of the spectral energy is confined within a small radial disc, and hence care must be taken while selecting a suitable cutoff. The optimum value of
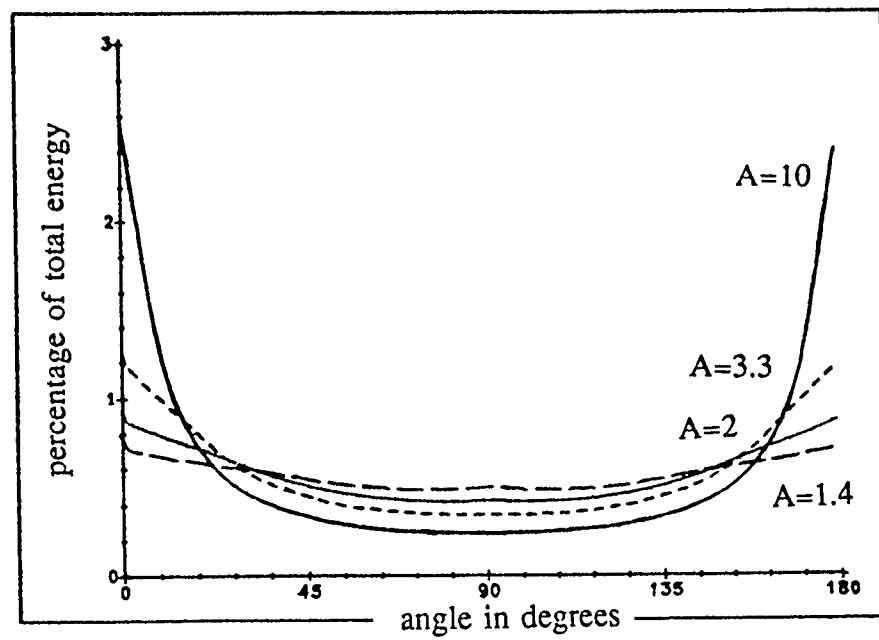
Fig. 4.9 : Angular energy distribution (theoretical) for rectangular line segments with varying length-to-width ratios (A). As this ratio decreases, there is an increased spreading of spectral components.
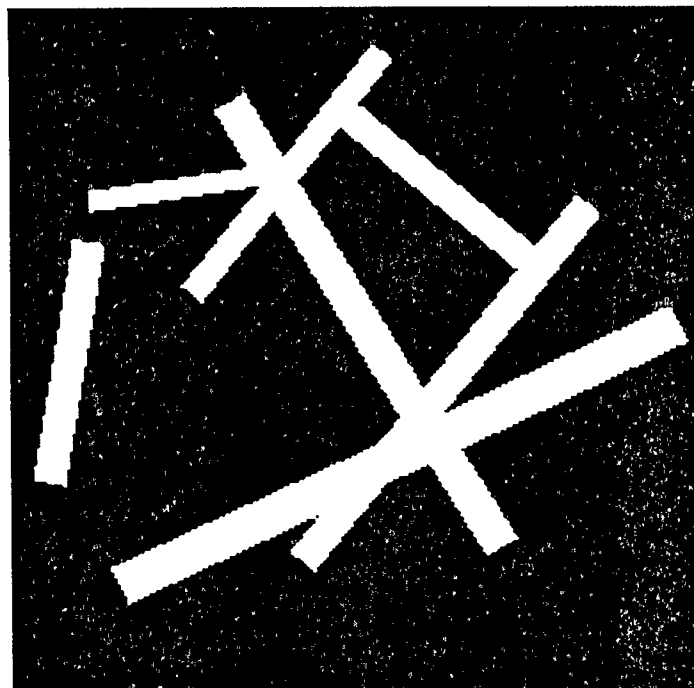
Fig. 4.10 : Example of a simulated test pattern with rectangular line segments of varying grey levels, oriented at different directions.
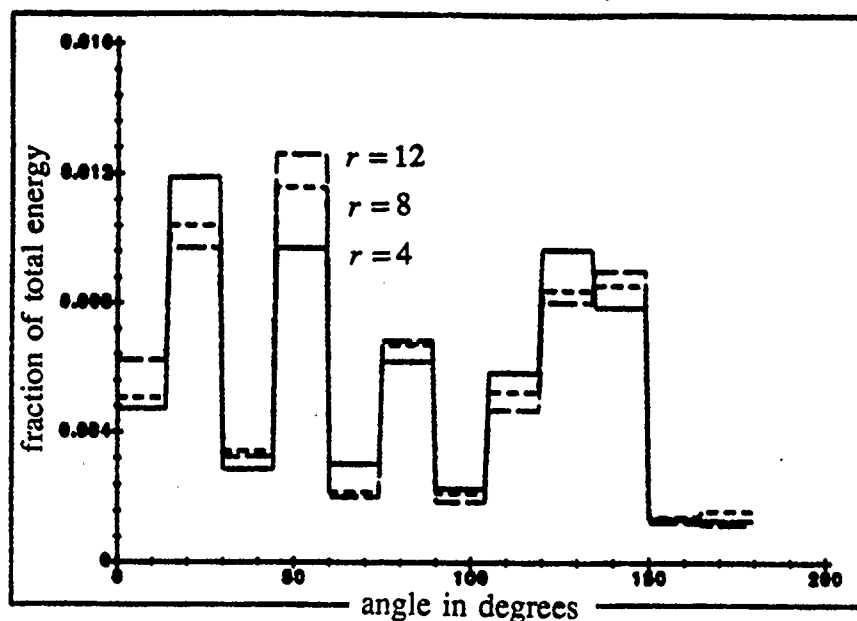
Fig. 4.11 : Plot to show the increase in angular selectivity with the removal of low frequency components for the image in fig. 4.10. The energy within different angle bands of width 15° is given in the form of a bar-graph. Here, r is the radius of the low frequency zone removed from the spectrum.
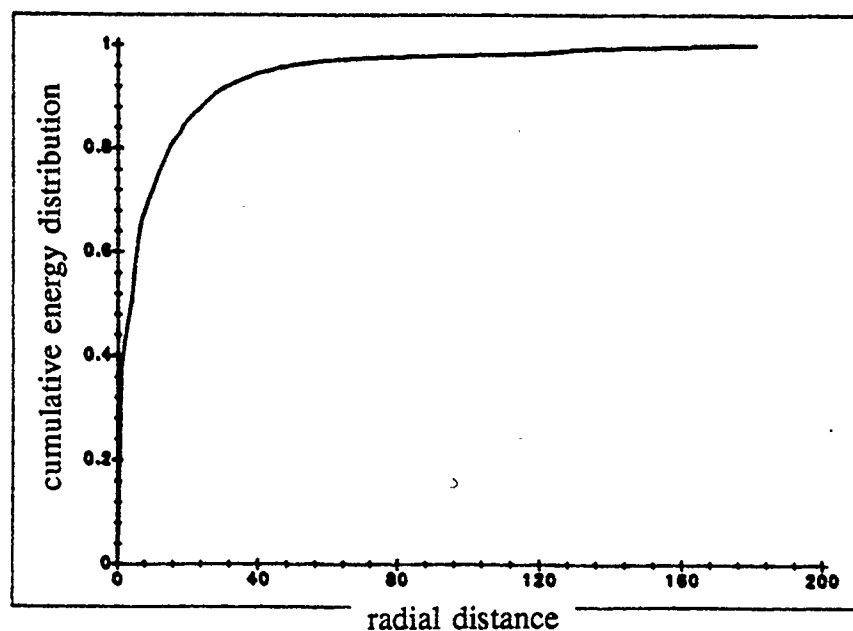
Fig. 4.12 : Cumulative radial energy distribution $\psi(R)$ for the image in fig. 4.6.

r also depends on the angular distance (θ_k) between the two distributions and on the length-to-width ratios of the line segments.

The removal of low frequency components from the spectrum of the image results in greater angular selectivity and avoids three problems: (a) overlapping of low frequency components at all angle bands (as shown in Figure 4.9), (b) interference from the components lying in the adjoining angle bands due to very coarse definition of angle in the rectangular coordinate system for the low frequency zone, and (c) contributions to the low frequency zone due to the presence of undesired quasi-circular lumps of matrix materials in a collagen fiber image. However, the choice of this cutoff frequency for a real fiber image is complicated due to the following facts :

(1) The fiber image may contain a number of line segments (fibrils) having different dimensions and grey levels.

(2) Different component images may have varying amounts of fibril-covered areas.

(3) We may have considerable overlap between fibrils oriented at different angle bands.

Low frequency components in the spectrum constitute the background information and the high frequency components correspond to the edges in the image. Removal of too much of low frequency components may result in the suppression of thick fibril bundles and the recovery of only edges in the filtered

images.

In view of the above considerations, a sixth order Butterworth highpass filter is used to remove low frequency components from the spectrum. Also, a fourth order Butterworth lowpass filter is used to reduce high frequency noise. The lower and upper cutoff frequencies f_L and f_H of the filter depend on the nature of the image being analyzed. Thus, the image, before being passed through the sector filter, is prefiltered with a bandpass filter, a radial section of which is given in Figure 4.13. It has been mentioned earlier that the removal of low frequency components tends to suppress the background information in the filtered images. In order to circumvent this problem, the lower frequency components passed by the filter shown in Figure 4.13 are weighted more than the higher frequency components by a linear function. The net filter function is thus given by

$$H(f_r) = \frac{(1 - \beta f_r)}{([1 + (f_L/f_r)^{2m}][1 + (f_r/f_H)^{2n}])^{1/2}} \qquad (4.9)$$

where -

β = slope of the weighting function,

f_r = normalized radial frequency,

m = order of Butterworth HPF = 6, and

n = order of Butterworth LPF = 4.

Optimum values of f_L, f_H and β were obtained experimentally by analyzing a number of images.
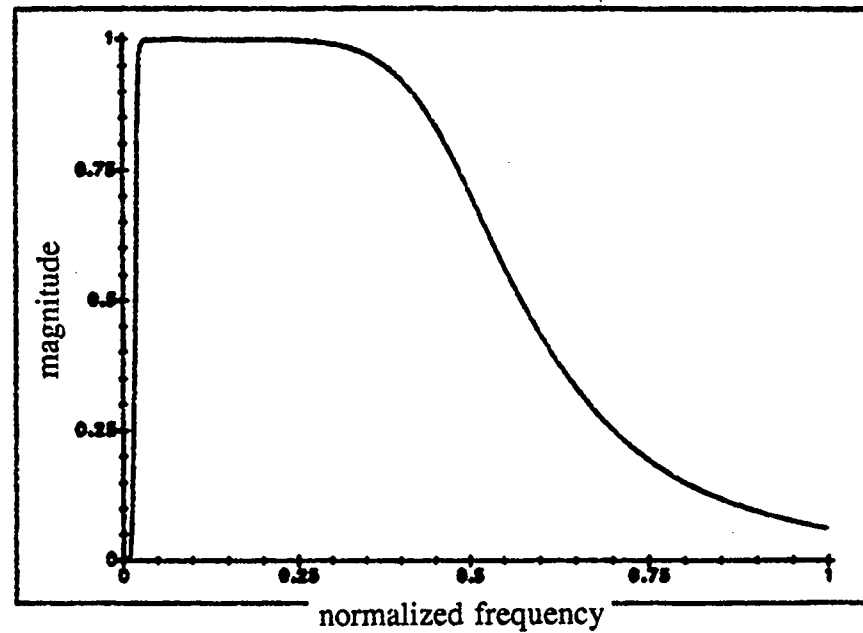
Fig. 4.13 : A radial section of the bandpass filter used, which comprises a 4th order Butterworth LPF with f_H = 0.5 and a 6th order Butterworth HPF with f_L = 0.02. The frequency scale is normalized with respect to the maximum radial frequency.

4.3 Design of the sector filter

For directional filtering, the components lying outside the desired angle band $\theta_1 < \theta < \theta_2$ are to be rejected. Use of an ideal sector filter leads to ringing and smearing artifacts[15]. In order to overcome this problem, the sector filter is multiplied by a raised cosine window[30]. The width of the cosine window is normalized with respect to the specified angle band and the function is independent of radial distance of the points. The window function may be mathematically expressed as

$$W(\theta) = \cos^{\alpha}(\frac{\theta - \theta_0}{B} \pi) ; \quad \theta \, \epsilon \, [\theta_1 , \theta_2]$$

(4.10)

$$= 0 \quad otherwise$$

where -

θ = angle subtended by the FT sample considered,

$\theta_0 = (\theta_2 + \theta_1)/2$ = center of the desired angle band, and

$B = \theta_2 - \theta_1$ = chosen angular bandwidth.

For $\alpha = 1.0$, the window coefficients are very low for the components lying near the two limits of the specified angle band, and the line segments corresponding to these angles may not appear in the filtered image at all. To alleviate this problem, a raised cosine window with $\alpha = 0.5$ is used.

For the purpose of directional decomposition of collagen fiber images, the angular bandwidth B is chosen to be 15°. The composite image is decomposed into twelve component images spanning the angle range $0^o - 180^o$. Following the

sector filtering, the inverse FT of the spectrum is computed and the filtered images are suitably thresholded and binarized (as discussed in the next chapter) to represent fibril-covered areas only. A count of the number of pixels belonging to the fiber class quantitatively represents the fibril-covered area in the chosen angle band. Such measures obtained for all angle bands spanning 0 to 180 degrees for a given image would describe the distribution of collagen fibrils completely.

CHAPTER 5

AUTOMATIC THRESHOLDING SCHEMES

5.1 Introduction

After filtering a composite image into different directional components, the fibril segments appear amidst a darker background in the corresponding filtered images. Since it is desired to evaluate the total area occupied by the fibril segments at a given direction, the filtered images have to be binarized by using suitable thresholds. The number of pixels belonging to the object class gives a measure of the desired fibril-covered area.

In the ideal case of a bright object superimposed on a dark background, the grey level histogram presents two peaks that are easily distinguishable. The choice of threshold for object detection is straight forward, being a point in the valley of the histogram. However, for most real images the thresholding process is not as simple, as the histogram provides only first order statistical information, disregarding the semantic content of the image. Further, the histograms are almost never bimodal.

In quantifying the fibril-covered areas in various component images, it is desirable to have a standard criterion to determine the suitable threshold values. Use of an automatic thresholding scheme for this type of pixel classification will make the procedure free from personal biases. This would enable a non-technical

person to use the procedure and to interpret the results.

Four major types of algorithms are available in the current literature for automatic threshold selection. In the first type, the input image is initially enhanced by manipulating each pixel using a certain neighborhood. The Laplacian or some other gradient operation is performed on the image to extract edge information[11,14,33-38]. The pixels are modified in a way that leads to a well-shaped histogram, when selection of the threshold is no longer a difficult job.

In the second type, a parametric model is developed for the histogram of the given image[8]. The histogram is approximated to be the superposition of a few Gaussian distributions in a least-square sense, and then a statistical decision theoretic method is used for classification. But these methods are tedious, and like all numerical algorithms, tend to be unstable. Hence these methods were not tried.

The third type of thresholding procedures use relaxation algorithms[8,39]. These are iterative procedures, and make use of pixel neighborhood properties in the input image. The pixel value is modified iteratively according to some *a priori* knowledge of the belongingness of a pixel to its associated neighborhood. The rate of convergence of such procedures is high. These algorithms sharpen the valley in the resulting histogram. However, in some cases, these methods could be unsuccessful.

Algorithms of the fourth type deal with the histogram of the given image only, and do not need any *a priori* knowledge about the image. Such a method remains the same for different kinds of images, and thus is probably the most

suitable scheme for unsupervised and automatic thresholding. An optimal threshold is obtained by optimizing a few parameters under certain constraints[40-44].

Segmentation based on thresholding may suffer from three types of errors[45]. If the input image is thresholded at some lower value than the actual one (not known precisely), the binarized image may contain fictitious segments which were not present in the original image. On the other hand, use of a higher threshold may remove some of the segments which were actually present in the original image. Finally, the thresholded segments get distorted for a particular threshold, when fictitious segments appear and some true segments disappear in the binarized image. This is the most difficult problem in thresholding and it may not be always possible to prevent this kind of error.

For the purpose of thresholding the component fibril images, a number of currently available thresholding algorithms belonging to different classes were tried[8, 35-38, 41, 43, 44]. Unfortunately, none of them yielded completely satisfactory results. A few of these algorithms are briefly discussed in the following sections.

5.2 Entropy considerations

Pun[42] put forward a new method of image segmentation using the entropy of the grey level histogram. Further modifications to this automatic thresholding algorithm have been proposed[43]. The modified thresholding scheme tries to maximize the information between the object and the background distributions of

grey levels in the given image.

Let k be the proposed grey level, L be the total number of grey levels in the input image, p_i be the probability of occurrence of the ith grey level, P_k be the cumulative probability of occurrence up to the kth grey level, $H(j)$ be the total entropy for the grey level distribution in the jth class ($j = 0$ *or* 1), and H_n be the entropy of the entire histogram. We may then write the following equations

$$P_k = \sum_{i=1}^{k} p_i, \tag{5.1}$$

$$H_n = - \sum_{i=1}^{L} p_i \log p_i, \tag{5.2}$$

and

$$
\begin{aligned}
H(0) &= - \sum_{i=1}^{k} \frac{p_i}{P_k} \log \frac{p_i}{P_k} \\
&= - \frac{1}{P_k} \left[\sum_{i=1}^{k} p_i \log p_i - P_k \log P_k \right] \\
&= \log P_k + \frac{H_k}{P_k}
\end{aligned}
\tag{5.3}
$$

where

$$H_k = - \sum_{i=1}^{k} p_i \log p_i. \tag{5.4}$$

Similarly,

$$H(1) = -\sum_{i=k+1}^{L} \frac{p_i}{1 - P_k} \log \frac{p_i}{1 - P_k}$$

$$= -\frac{1}{1 - P_k} \left[\sum_{i=k+1}^{L} p_i \log p_i - (1 - P_k) \log(1 - P_k) \right] \qquad (5.5)$$

$$= \log(1 - P_k) + \frac{H_n - H_k}{1 - P_k}.$$

If the function $\psi(k)$ is defined to be the sum of the entropies of the two classes $H(0)$ and $H(1)$, then from equations (5.3) and (5.5) we may write

$$\psi(k) = \log P_k (1 - P_k) + \frac{H_k}{P_k} + \frac{H_n - H_k}{1 - P_k}. \qquad (5.6)$$

The discrete value of k which maximizes $\psi(k)$ is the desired threshold value. The binarized image will have the maximum entropy or the maximum amount of information for a given image.

5.3 Moment-preserving thresholding

A new approach to automatic thresholding using the moment-preserving principle has been suggested by Tsai[44]. The threshold values are computed deterministically in such a way that the moments of an input image are preserved in the output (binarized) image. This approach may be regarded as a moment-preserving image transformation which recovers an ideal bilevel image from a blurred one.

Let p_i be the probability of occurrence of the ith grey level, L be the total number of grey levels in the image, and m_j be the jth order moment of the image. The different order grey level moments may be defined as

$$m_j = \sum_{i=1}^{L} i^j p_i. \qquad (5.7)$$

The input image can be considered as a blurred version of an ideal bilevel image consisting of only two grey levels z_0 and z_1, where $z_0 < z_1$. This method selects a threshold value such that if the image is binarized to the grey levels z_0 and z_1, then the first three moments of the image are preserved in the resulting bilevel image. Hence, if P_0 and P_1 correspond to the probabilities of the resulting classes in the binarized image, we may write the following four equations

$$P_0 + P_1 = 1,$$

$$P_0 z_0^1 + P_1 z_1^1 = m_1,$$

$$P_0 z_0^2 + P_1 z_1^2 = m_2, \qquad (5.8)$$

$$P_0 z_0^3 + P_1 z_1^3 = m_3.$$

These nonlinear equations may be solved for the four unknowns P_0, P_1, z_0 and z_1 by using the methods described in reference[44]. Once P_0 has been evaluated, the equation

$$\sum_{i=1}^{k} p_i \approx P_0. \tag{5.9}$$

may be solved for the discrete value of k, which is the desired threshold.

5.4 Probabilistic Relaxation Method

This algorithm[8] makes fuzzy or probabilistic classification decisions at every point in the image in parallel at each iteration, and then adjusts these decisions during successive iterations based on the decisions made in the preceding iteration at neighboring points.

Considering the binarization of an image, there are two output classes C_0 and C_1. The thresholding is assumed to be independent of direction. For each point A_i there are only two probabilities p_{i0} and $p_{i1} = 1 - p_{i0}$ of being classified in C_0 or in C_1. Consequently, there exist four kinds of class compatibilities c_{00}, c_{01}, c_{10} and c_{11}, where c_{ij} denotes the probability of a pixel belonging to class C_i being classified in C_j. The compatibility functions c_{ij} can have values in the range (-1,1), negative values denoting incompatibility.

The procedure starts with an initial probability estimate $p_{i1}^{(0)}$ as the normalized grey level of the pixel within the range [0,1]. For symmetrical cases $c_{00} = c_{11}$ and $c_{10} = c_{01}$. For the binarization of the fibril component images $c_{11} = 0.8$ was found to be optimal. Then $c_{01} = 1 - c_{11} > 0$. After each iteration p_{ij} is incremented or decremented according to the user-defined compatibilities. The increments at the rth iteration are defined as

$$q_{i0}^{(r)} = \sum [c_{00} \, p_{i0} + c_{01} \, p_{i1}] \qquad (5.10)$$

and

$$q_{i1}^{(r)} = \sum [c_{10} \, p_{i0} + c_{11} \, p_{i1}] \qquad (5.11)$$

where the summation is carried out over the defined neighborhood and q_{ij} is the iterative increment for the corresponding pixel i towards the class j. To ensure that after modification the value of p_{ij} does not become negative or greater than unity, it is normalized as

$$p_{ij}^{(r+1)} = \frac{p_{ij}^{(r)}(1 + q_{ij}^{(r)})}{\sum p_{ij}^{(r)}(1 + q_{ij}^{(r)})} \qquad (5.12)$$

where the summation is carried out over the neighborhood of the pixel A_i, j being class zero or one.

As the iterations progress, the resulting grey level histogram attains the shape of two isolated peaks separated by a flat valley. After the rth iteration the image may be binarized by classifying the ith pixel to class C_0 if $p_{i1} < 0.5$, or to class C_1 otherwise.

5.5 Inter-class Variance Maximization method

This is a nonparametric and unsupervised method of automatic threshold selection for image segmentation. An optimal threshold is selected by the discriminant criterion so as to maximize the separability of the resultant classes in grey level[41]. This algorithm makes use of only the zeroth and first order grey

level statistics.

Let

L = total number of grey levels,

p_i = probability of occurrence of grey level i,

k = intended grey level threshold,

$\omega_j(k)$ = probability of occurrence of class C_j for the threshold k,

μ_j = mean grey level in class C_j, and

$$\mu_T = \sum_{i=1}^{L} i\, p_i = \text{mean grey level in input histogram.}$$

Then,

$$\omega_0(k) = \sum_{i=1}^{k} p_i \quad ; \quad \omega_1(k) = \sum_{i=k+1}^{L} p_i = 1 - \omega_0(k) \tag{5.13}$$

and

$$\mu_0 = \frac{1}{\omega_0(k)} \sum_{i=1}^{k} i\, p_i = \mu(k)\, /\, \omega_0(k) \tag{5.14}$$

where

$$\mu(k) = \sum_{i=1}^{k} i\, p_i. \tag{5.15}$$

$\mu(k)$ denotes the first order moment of the histogram up to the kth grey level. Similarly,

$$\mu_1 = \frac{\mu_T - \mu(k)}{1 - \omega_0(k)}. \tag{5.16}$$

From the above relationships it is seen that

$$\omega_0(k)\mu_0 + \omega_1(k)\mu_1 = \mu_T. \tag{5.17}$$

After thresholding, the resulting image has only two grey levels. The probabilities of occurrences of C_0 and C_1 are given by $\omega_0(k)$ and $\omega_1(k)$. The variance of the data points in the binarized image is denoted by $\sigma_B^2(k)$, which may be evaluated from the following expression

$$\sigma_B^2(k) = \omega_0(k)(\mu_0 - \mu_T)^2 + \omega_1(k)(\mu_1 - \mu_T)^2. \tag{5.18}$$

Substituting the values of μ_j and $\omega_1(k)$, $\sigma_B^2(k)$ is given as

$$\sigma_B^2(k) = \frac{[\mu_T\omega_0(k) - \mu(k)]^2}{\omega_0(k)[1 - \omega_0(k)]}. \tag{5.19}$$

Now, $\sigma_B^2(k)$ is a measure of global contrast achieved by thresholding the image at grey level k. Hence, this quantity may be called inter-class variance. The total variance for a given histogram is a constant and is given by the sum of intra-class and inter-class variances. The main objective of this algorithm is to minimize the intra-class variances which will automatically ensure the maximization of inter-class variance. In other words, the aim is to minimize the probability of a point belonging to a class being assigned to the other class. Now, the optimal threshold is given by T where

$$\sigma_B^2(T) = \max[\sigma_B^2(k)]. \tag{5.20}$$

It should be mentioned here that in equation (5.19) $k = 1$ and $k = L$ correspond to two trivial solutions and should be neglected. Since p_i is non-negative and bounded, a maximum always exists within the interval $(1, L)$.

5.6 Performance of different thresholding schemes

The first two automatic thresholding schemes discussed are suitable for images with histograms having distinctive peaks and a valley in between. The filtered fibril component images usually do not have any flat valley region in the histogram, and hence these two algorithms failed in most cases.

Unlike the first two algorithms, the probabilistic relaxation method is neighborhood-dependent, and edgeness criteria may be included while thresholding. Since each point in the input image has to be dealt with individually, the computation time is naturally more than that for the inter-class variance maximization method. However, the speed may be increased appreciably by using parallel processors[8]. The method is reported to be more powerful than non-iterative methods as it refines itself after each iteration based on local context. This algorithm has been found to converge fast enough for all practical cases (typically six to seven iterations).

The initial assumption of p_{ij} as the probability of the ith pixel being classified to class j may lead to erroneous results when most of the grey levels in the given image lie on one side of the grey scale midpoint; the binarized image becomes either over-thresholded or under-thresholded. This method should only be applied

to images that have substantial grey level population on both sides of the grey level midpoint. A long trail of low counts in the histogram of the input image could render this method inapplicable.

The histograms of the filtered fibril images usually have a long trail because of the nature of the background in the component images. Hence the relaxation method does not give a good result when the fibril-covered area in the component image is low. Let G be the grey level at which the normalized histogram attains a value $1/L$ (L is the number of grey levels in the input image) for the first time starting from the lower end. To overcome the problem mentioned above, we may clip all pixels having grey levels lower than G to G. This gives good results for filtered images with component fibrils occupying smaller areas. However, it fails to analyze component images for the normal tissues along the dominant directions.

The inter-class variance maximization method (referred to as Otsu's method hereafter), in general, yields better results than any of the other three methods. This method gives good results for filtered images with a significant amount of fibril-covered area. For other cases, the obtained threshold was always found to be lower than the most suitable threshold obtained manually by trial and error. This method is the most appropriate for binarization of the component fibril images. However, to guard against errors due to the above limitations, the threshold value given by the method was used only as an initial estimate, and the actual threshold was found iteratively.

CHAPTER 6

RESULTS

6.1 Application to test patterns

The filtering procedure was initially applied to the simulated test pattern shown in Figure 6.1. The test pattern has line segments (fibrils) with various lengths, widths and grey levels at four different angles, namely at 0°, 45°, 90° and 135°. These parameters were chosen to represent characteristics of individual and grouped fibrils, and to study the performance of the filtering methods under varying conditions. The FT of this test pattern is given in Figure 6.2. The FT shows the presence of frequency components along directions orthogonal to the above mentioned directions as expected. The cutoff frequencies f_L and f_H for the initial bandpass filter were chosen to be 0.006 and 0.500, where the frequency scale has been normalized with respect to the maximum radial frequency in the spectrum. The value of β, the slope of the linear weighting function in equation (4.9), was chosen to be 0.7. Figure 6.3 shows a component image obtained using the filtering procedure described previously with the angle band $125^o - 140^o$ in the Fourier plane. Clearly, only those lines oriented at 45° (in the image plane) have been passed. The corresponding binarized image, using the threshold value given by Otsu's method of automatic threshold selection, is shown in Figure 6.4. Figures 6.5 and 6.6 show similar components (binarized) extracted from the
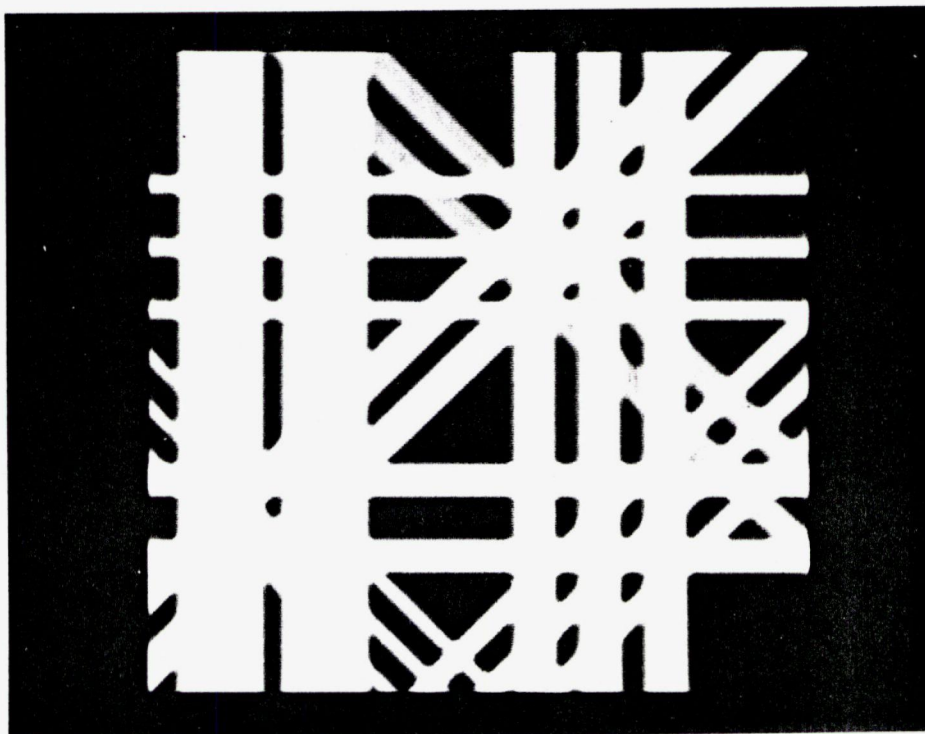
Fig. 6.1 : A synthesized composite image with fibrils (lines) oriented at four directions : 0, 45, 90, and 135 degrees. The line segments have different grey levels.
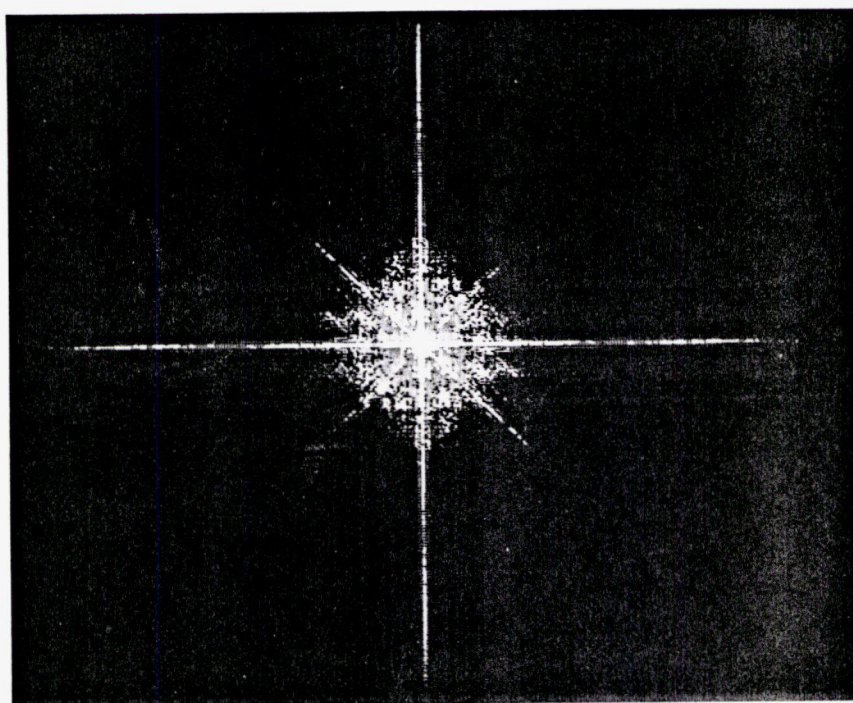
Fig. 6.2 : Fourier spectrum of the image in fig. 6.1. The presence of directional components is evident from the spectrum.
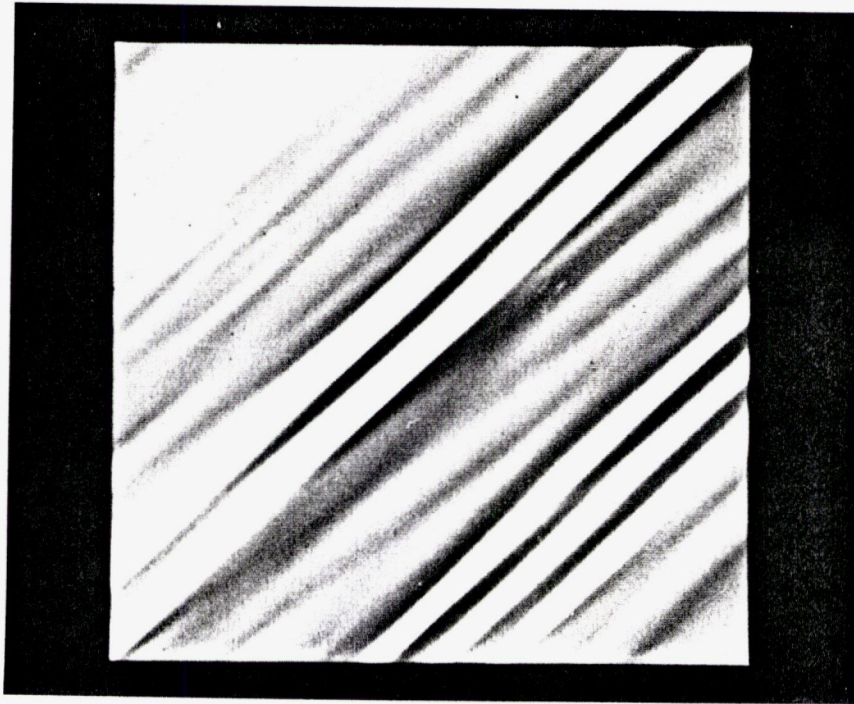
Fig. 6.3 : Result of application of the directional filter spanning the angle band $125^o - 140^o$ (in Fourier space) to the image in fig. 6.1. The filtered images have been scaled to the display range of 0-255 (integers).
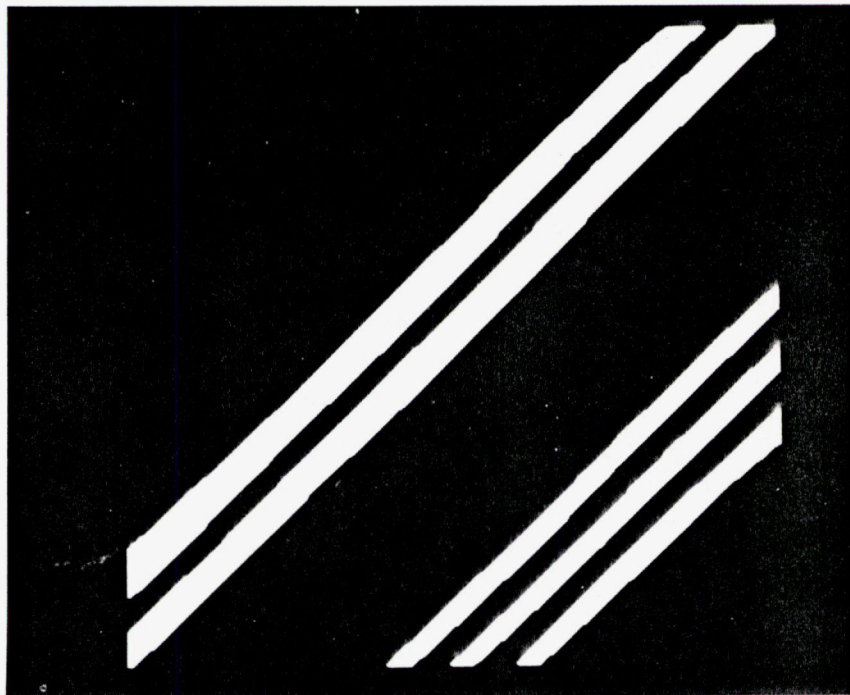
Fig. 6.4 : Thresholded and binarized component image obtained from the image in fig. 6.3.
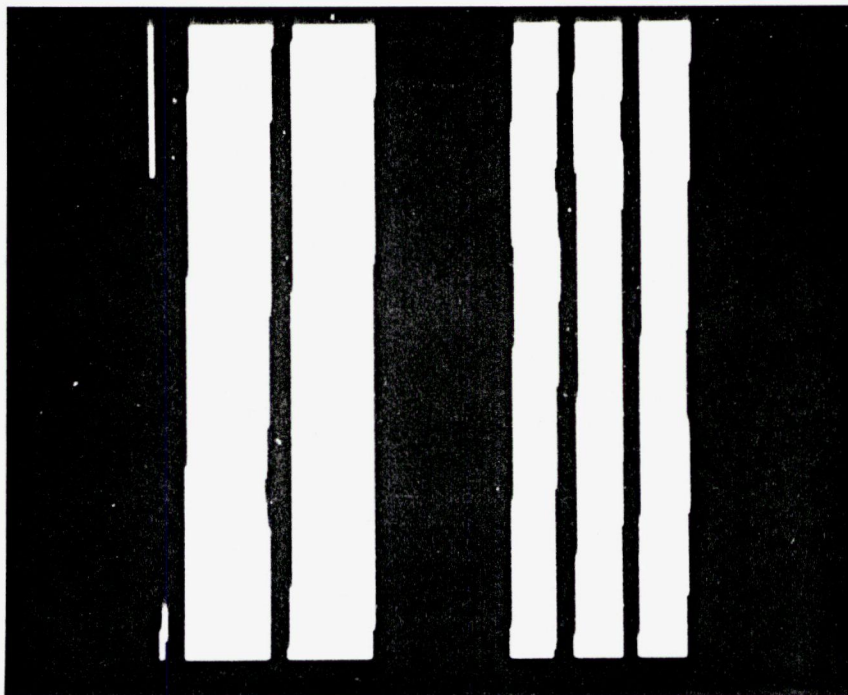
Fig. 6.5 : Thresholded and binarized component image for the angle band $80^o - 95^o$ (in image space) obtained from the image in fig. 6.1.
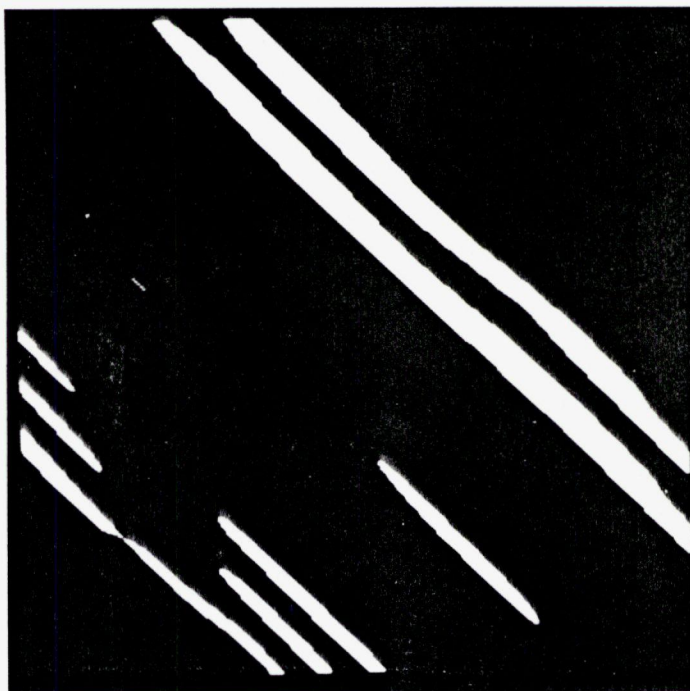
Fig. 6.6 : Thresholded and binarized component image for the angle band $125^o - 140^o$ (in image space) obtained from the image in fig. 6.1.

composite image for the angle bands $80^o - 95^o$ and $125^o - 140^o$ (in the image plane) respectively.

A close look at these component images will reveal that a region of overlap of lines oriented at different directions contributes to each direction. The 135° component image in Figure 6.6 has the maximum error as the grey level of this component is very low compared to those of the others. The quantitative distribution of fibril-covered area along different directions in the analyzed test pattern is given in Table 1. The true fibril-covered areas, which include overlapping, are also given in the table. Table 1 shows that, except for the 135° component, the errors are reasonably small.

It may be argued here that the input image may be thresholded and binarized before being analyzed to overcome the grey level dependency of the above procedures. However, the binarization process may remove some weak fibrils from the input image. Further, the FT of the binarized image may introduce a significant amount of aliasing and other artifacts.

The accuracy of the method was further substantiated by applying it to the simulated test pattern given in Figure 4.10. In this pattern the line segments do not run from one end to the other end of the image, thus allowing the effect of smearing artifacts at both ends of the segments to be visible in the component images. The filtered component images were found to have negligible effects due to smearing artifacts. The quantitative distribution of fibril-covered area along different directions in this analyzed test pattern is given in Table 2. The component

Table 1 : Distribution of fibril-covered area in the synthesized composite image in fig. 6.1.

Angle Band	True Area (pixels)	Computed Area (pixels)	Percentage Error
$35^o - 50^o$	12528	11984	-4.34
$80^o - 95^o$	29952	29046	-3.03
$125^o - 140^o$	9486	6774	-28.59
$170^o - 180^o$	10240	10549	3.02

Table 2 : Distribution of fibril-covered area in another test pattern given in fig. 4.10.

Angle Band	True Area (pixels)	Computed Area (pixels)	Percentage Error
$0^o - 15^o$	331	370	11.8
$15^o - 30^o$	2601	2674	2.8
$45^o - 60^o$	2301	2258	-1.9
$75^o - 90^o$	798	880	10.3
$105^o - 120^o$	1872	1646	-12.1
$135^o - 150^o$	812	820	1.0

corresponding to the angle band $105^o - 120^o$ has a large error due to the fact that the actual line segment is aligned exactly at 120°. The weighting function for the sector filter $W(120^o)$ in equation (4.10) being equal to zero, the defined sector excludes a substantial portion of the energy in the spectrum of the corresponding components. This leads to a poor filtered image.

6.2 Application to ligament tissue samples

Since the collagen fiber images usually have poor contrast, the images were first histogram equalized to improve their global contrast. After some experimentation, optimum values of $f_L = 0.02$ and $f_H = 0.5$ were determined for the initial bandpass filter, which were used on all the fiber images analyzed. Figure 6.7 shows the output of the directional filter after binarization for one of the dominant angle bands $(75^o - 90^o)$ in the image space for the normal ligament image shown earlier in Figure 4.6. For the purpose of illustration, one more binarized component image for the angle band $30^o - 45^o$ is given in Figure 6.8. Figure 6.8 shows that the procedure has been able to detect small and weak stray fibrils also. Proceeding in the same way, directional components at all angle bands were obtained. (The automatic thresholding method had limited success - more details are presented in the Chapter 8). The results are given in Figure 6.9 as a rose diagram[46]. This diagram, displaying relative fibril-covered areas in the various angle bands used, shows that most of the fibrils in the normal tissue lie close to the long-axis of the ligament (90°).

Fig. 6.7 : Binarized component obtained from the image given in fig. 4.6 for the angle band $75^o - 90^o$ (in image space).
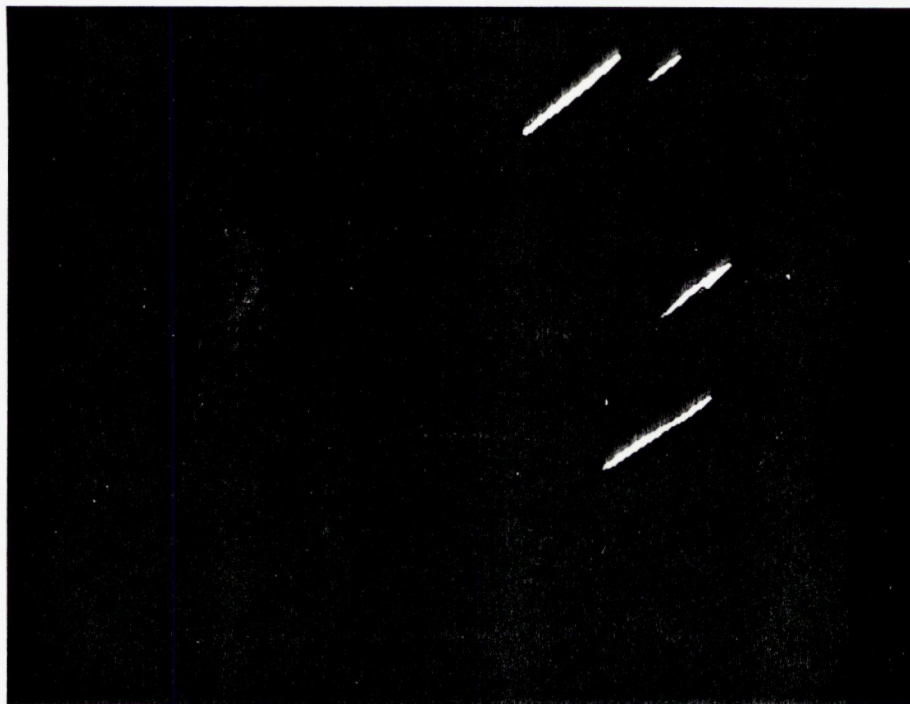

Fig. 6.8 : Thresholded and binarized component image for the angle band $30^o - 45^o$ (in image space) for the image in fig. 4.6.

Figure 6.10 shows an image of an injured ligament after one week of healing and Figure 6.11 shows its FT. It is readily seen that the collagen fibers here do not have any dominant orientation. Figure 6.12 shows the filtered output (binarized) for the angle band $75^o - 90^o$. Again, to qualitatively illustrate the efficacy of the algorithm, one more binarized component image for the angle band $30^o - 45^o$ is given in Figure 6.13. The quantitative results obtained for all the angle bands are represented in Figure 6.14 as a rose diagram. It is seen that the distribution of collagen fibrils in the healing tissue, at this interval of healing, is highly random.

Distributions as above were obtained for all images available in the normal (20 images) and one week healing (18 images) ligament groups. The rose diagrams of the average angular distribution of collagen fibrils for these two classes of ligament samples are given in Figure 6.15. The average distributions for the two classes demonstrate properties similar to those of the respective images presented in Figure 6.9 and 6.14.

Proceeding in the same way, images of ligament samples corresponding to different stages of healing (namely, three weeks, six weeks and twelve weeks) were analyzed and the average distributions of collagen fibrils for these classes were obtained. The rose diagrams for these average distributions are given in Figure 6.16. These diagrams indicate quantitatively the process of collagen re-alignment with healing. Further quantification of these distributions is discussed in the next chapter.
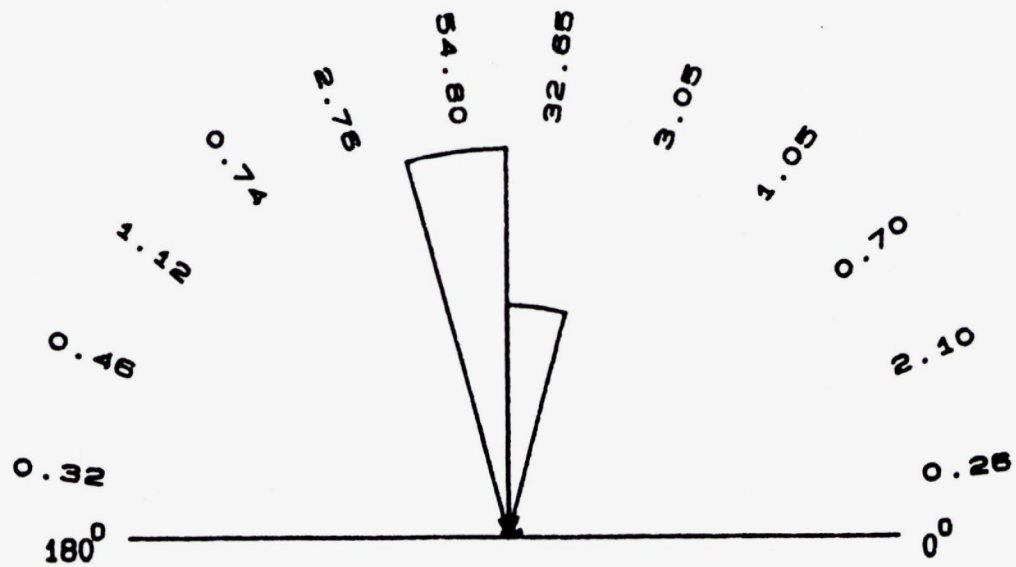
Fig. 6.9 : Rose diagram representing the relative angular distribution of fibril-covered area for the image in fig. 4.6. The axis of the ligament is at 90°.
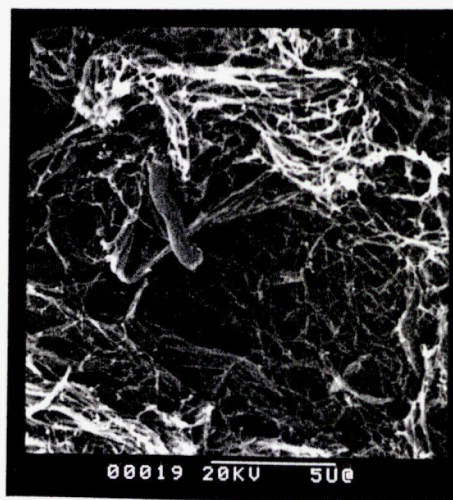
Fig. 6.10 : A representative image showing collagen alignment in a one-week healing ligament.
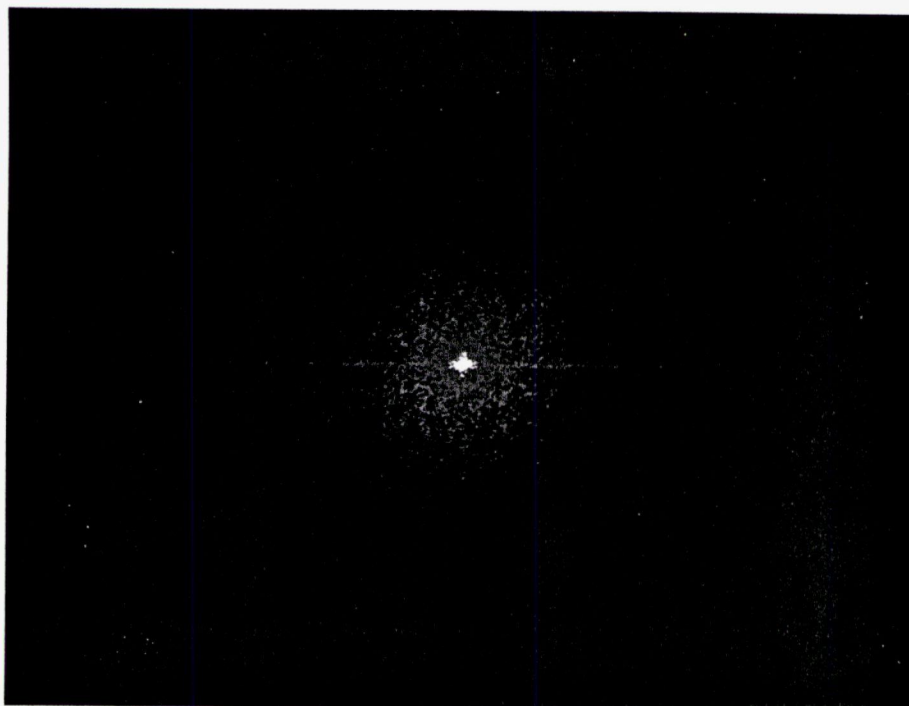
Fig. 6.11 : Fourier spectrum of the image in fig. 6.10. The energy in the spectrum is distributed in all directions, indicating that the fibrils in the tissue sample are scattered at random.
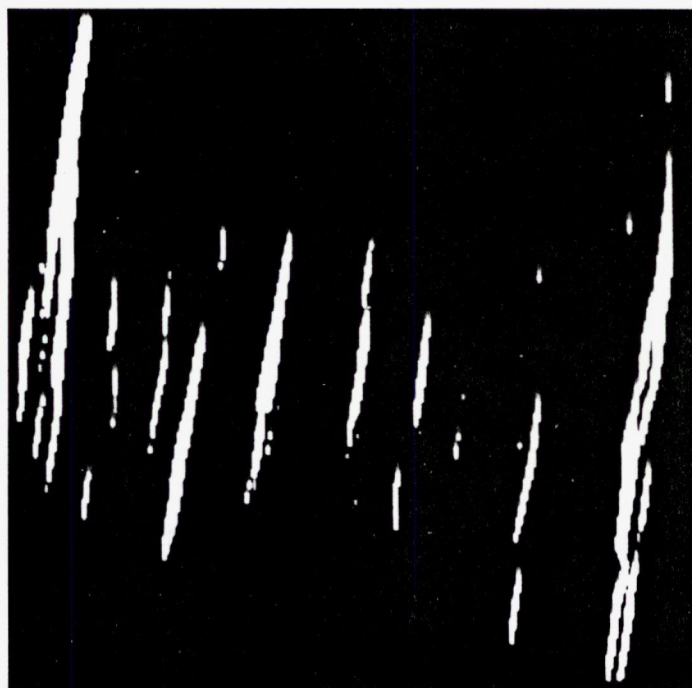
Fig. 6.12 : Binarized component image for the angle band $75^o - 90^o$ (in image space) for the image in fig. 6.10.

Fig. 6.13 : Thresholded and binarized component image obtained from the image in fig. 6.10 for the angle band $30^o - 45^o$ (in image space).
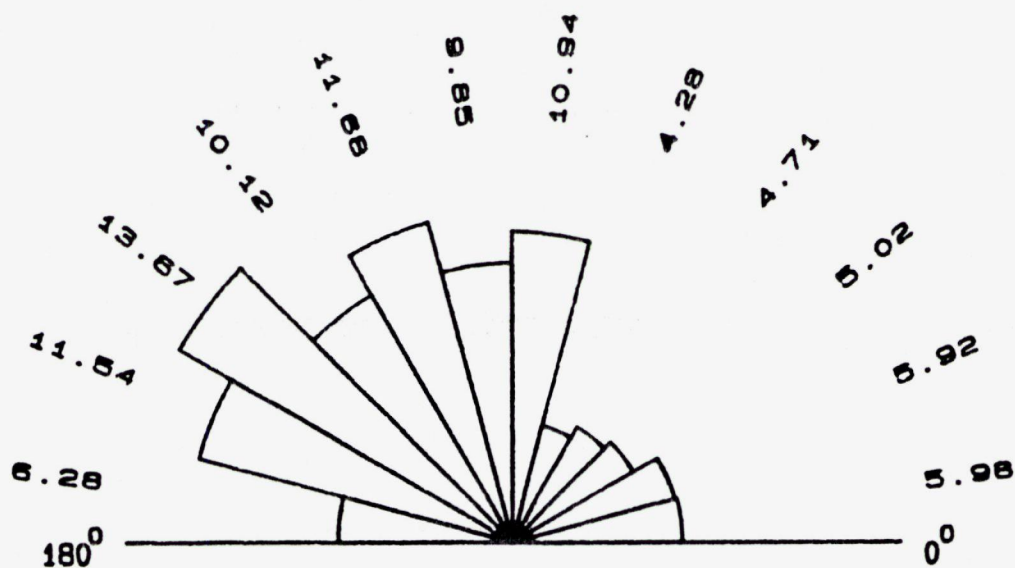
Fig. 6.14 : Rose diagram representing the relative angular distribution of fibril-covered area for the image in fig. 6.10. The axis of the ligament is at 90°.
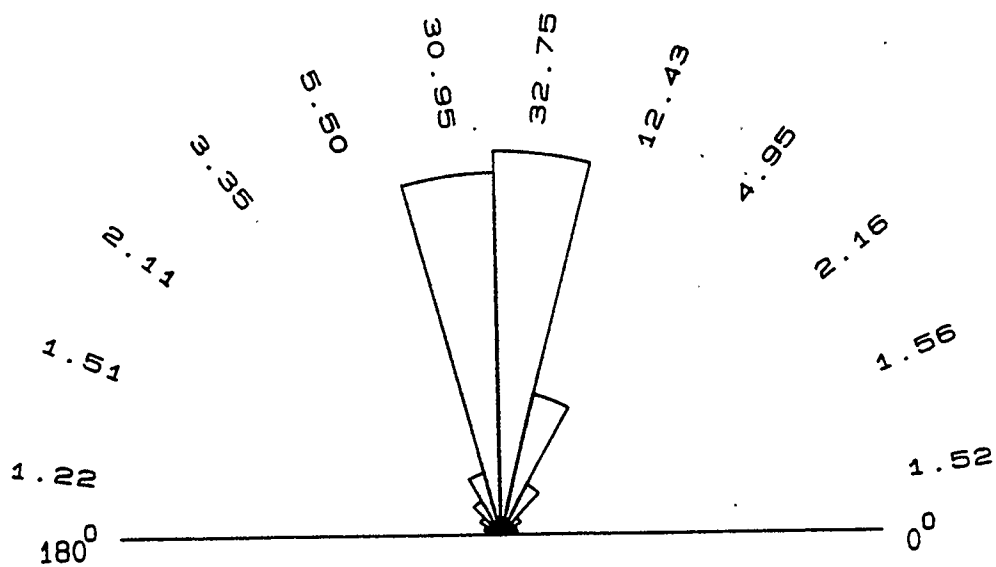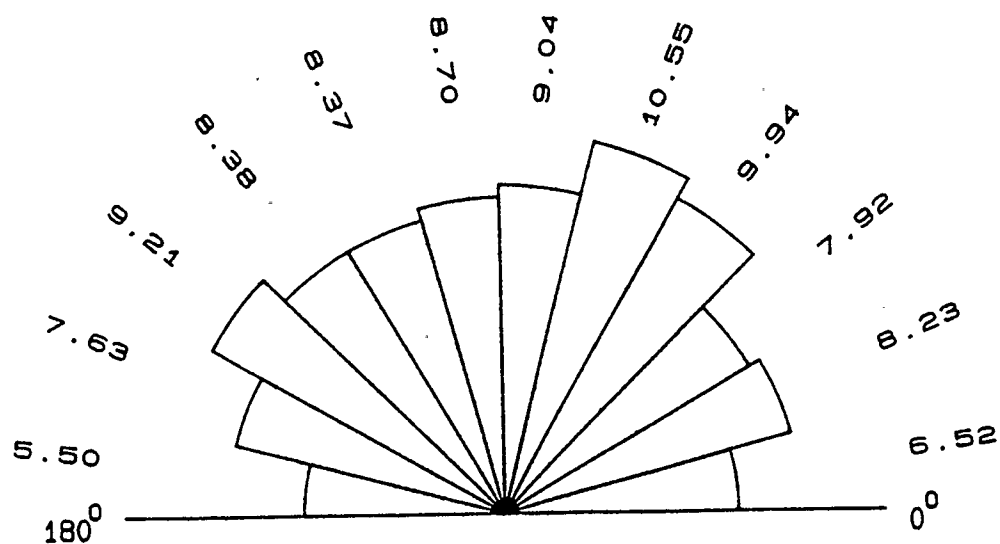
Fig. 6.15 : Rose diagrams representing the average class distribution of fibril-covered area (a) in a normal, and (b) in a one-week scar tissue.
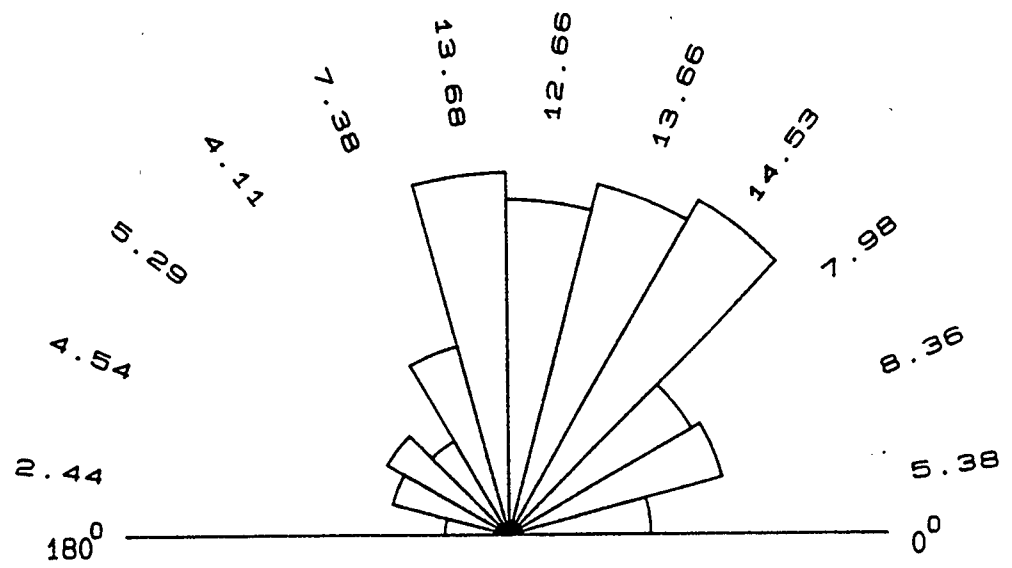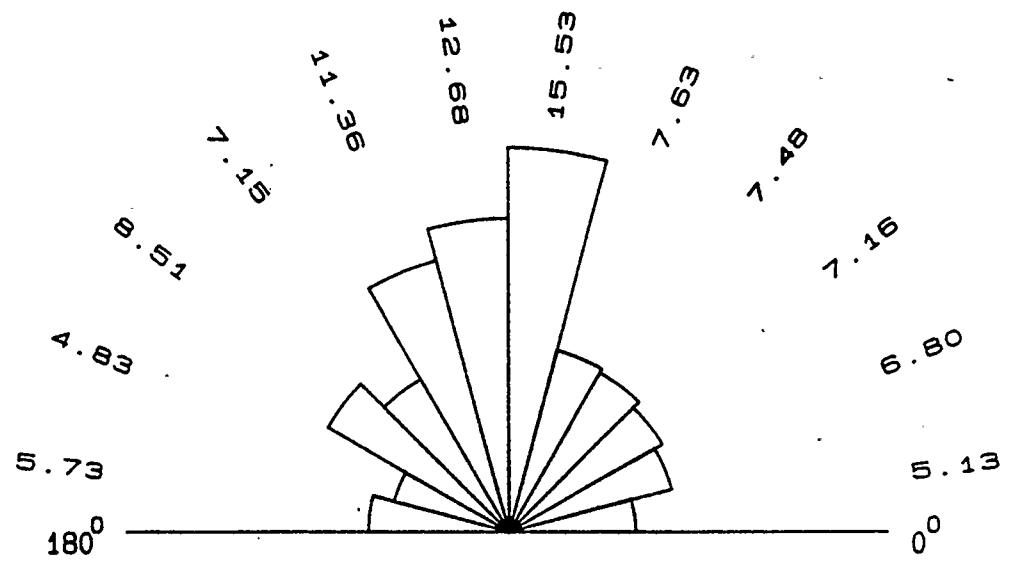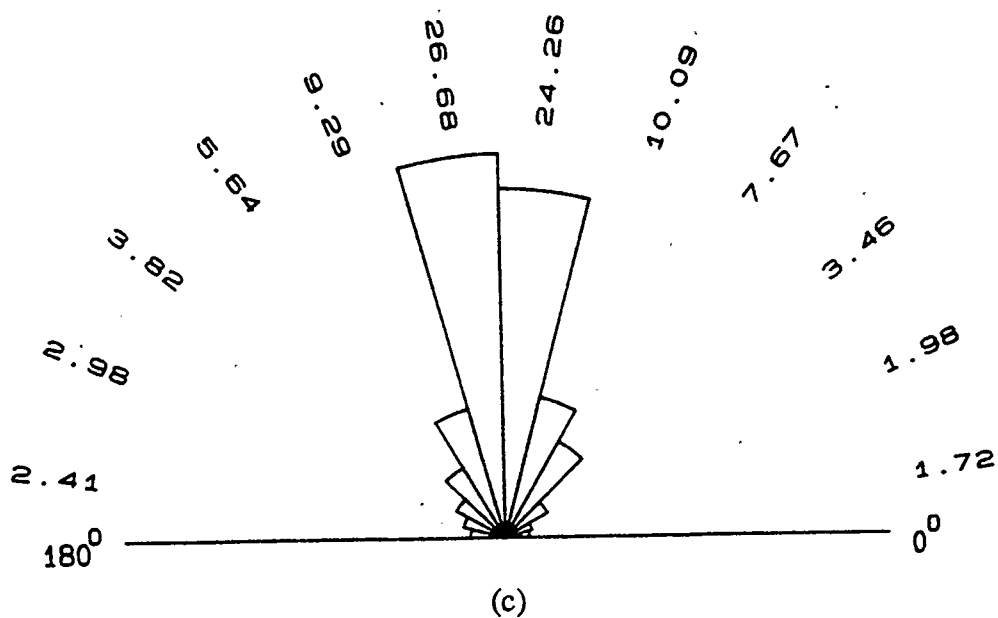
Fig. 6.16(a)

Fig. 6.16(b)

Fig. 6.16 : Rose diagrams representing the collagen remodeling process in ligaments at different stages of healing : (a) 3 weeks of healing, (b) 6 weeks of healing, and (c) 12 weeks of healing.

6.3 Reproducibility of the results

The distribution of collagen fibrils varies widely between ligaments at different healing stages. However, images taken from the same ligament sample should have similar distributions. The distribution function for an individual fiber image should show high correlation with the average distribution for the class. The correlation coefficient (R) between two distributions $x_1(i)$ and $x_2(i)$ may be defined as the normalized dot product of the distribution vectors, and may be expressed as

$$R = \frac{\sum_{i=1}^{12} x_1(i)x_2(i)}{\left[\sum_{i=1}^{12} x_1^2(i) \sum_{i=1}^{12} x_2^2(i)\right]^{1/2}} \tag{6.1}$$

where $x_k(i)$ represents the fibril-covered area for the kth image along i^{th} angle band.

To study the effect of the number of images taken from a single ligament sample on the average distribution obtained, and to study the variability of distributions for images of the same tissue, the following two experiments were performed.

First, nine randomly selected fiber images taken from the same ligament were used to form a representative group, and the average fibril distribution was obtained. The value of R was obtained for each of the images in the group with respect to the average distribution for the group. These are given in the column 1a

Table 3 : Correlation of collagen distributions of individual images with the class average distribution. Entries in columns 1a and 2a are correlations with the class averages obtained from nine images each, while those in columns 1b and 2b are with the class averages obtained from eighteen images each. The small differences between the values of R in columns a and b suggest that a group of nine images is adequate to represent the collagen alignment in ligaments.

Image	Normal ligament		Scar tissue	
number	correlation (1a)	correlation (1b)	correlation (2a)	correlation (2b)
1	0.953	0.945	0.922	0.933
2	0.932	0.915	0.930	0.938
3	0.861	0.880	0.960	0.957
4	0.869	0.852	0.934	0.917
5	0.839	0.859	0.943	0.949
6	0.956	0.969	0.895	0.880
7	0.983	0.978	0.935	0.936
8	0.971	0.962	0.972	0.971
9	0.981	0.982	0.973	0.972

of Table 3 for a normal ligament. Similar results for a scar tissue after one week of healing are given in column 2a of Table 3. It is seen that the individual distributions are highly correlated with the average distribution of the class.

Next, the group size was increased to 18 and correlation coefficients as above were calculated again. The results are given in column 1b and 2b of Table 3 for the normal and scar groups (for the same images referred to in columns 1a and 2a). The change in the values of R is negligible between column 1a and 1b, and 2a and 2b. Thus, a set of nine randomly obtained images may be considered adequate for meaningful representation of the collagen alignment in ligaments.

To study the variations in collagen alignment from one ligament sample (animal) to another of the same category, two more normal ligaments taken from two different rabbits were analyzed. The average fibril distributions for the two ligament are given in Figure 6.17. Comparing these with Figure 6.9, it is seen that the three ligaments possess similar, though not identical, fibril distributions.

6.4 Effects of magnification

The tissue samples referred to so far were taken at a magnification of 7K. This value was experimentally chosen to give an optimal compromise between the resolution of collagen fibrils and the area of the tissue being sampled. The sensitivity of the directional filtering method with respect to fibril thickness was tested by the following procedure.
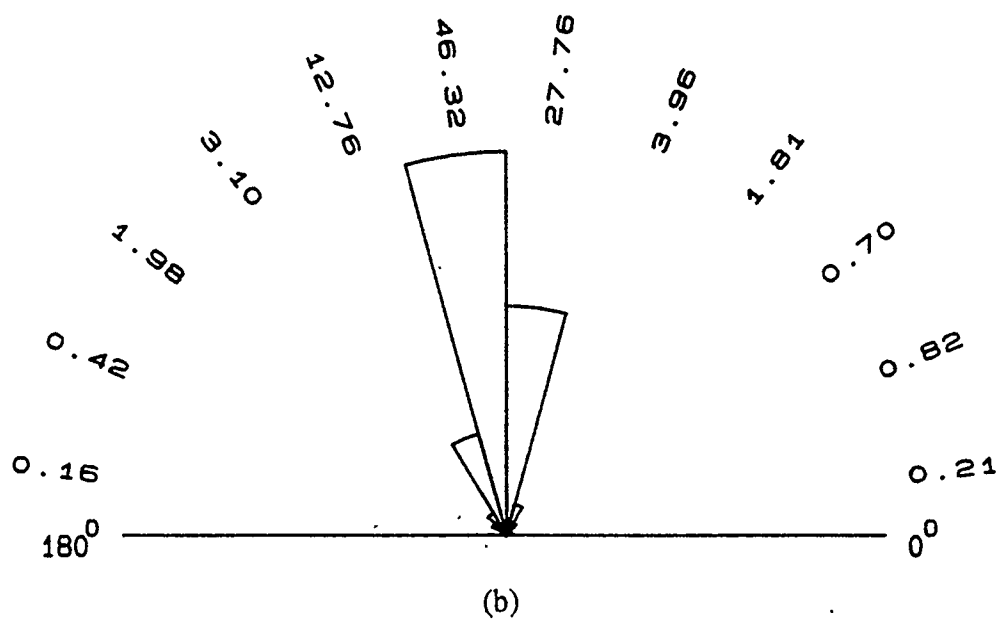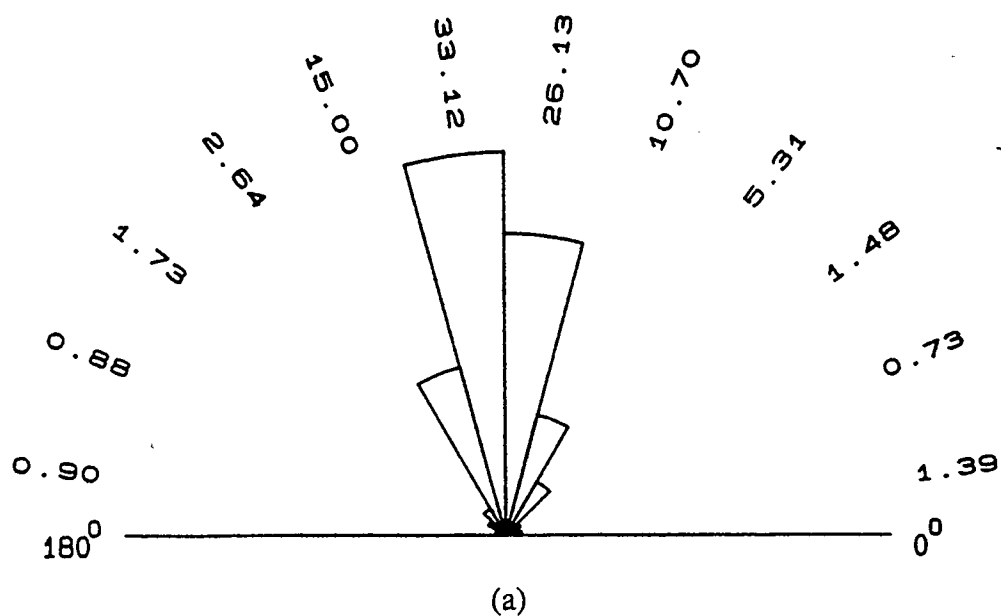
Fig. 6.17 : Rose diagrams showing the collagen alignment in normal ligaments in two other animal samples : (a) a young adult rabbit of age 8 months, and (b) an older rabbit of age 15 months.

A tissue sample image was first obtained at a magnification of 4K. This image was digitized to 506 x 474 pixels, effectively magnifying the image by a factor of about two as compared with the standard size of 256 x 256 used routinely. The digitized image was divided into four subimages of size 253 x 237 each. The magnified image was decimated by a factor of two in both horizontal and vertical directions (the resulting image covers the same area as the four subimages combined). Letting the standard digitization size of 256 x 256 represent the same magnification as that of the original micrograph, the decimated image has a magnification of 4K while the subimages have a (relative) magnification of 8K. The average distribution for the subimages was obtained, and is plotted with the distribution for the decimated image in Figure 6.18. It is seen that the two are nearly the same. The correlation (R) between these two distributions was found to be 0.99, signifying that the filtering method works at different magnifications (in the range studied) of the collagen fibrils as well.

It is evident from these experiments that the filtering method described works well for images of different medial collateral ligaments. The results obtained should help in meaningful quantification of the alignment of collagen fibrils in ligaments.
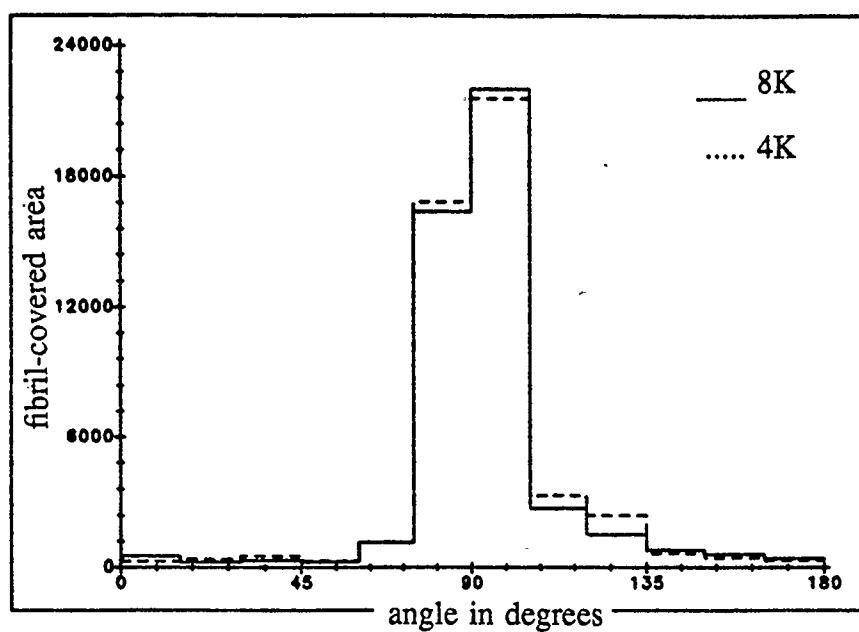
Fig. 6.18 : Plot of fibril distributions in a particular tissue sample for different magnifications. A close match between these two distributions indicates the applicability of the filtering procedure at varying magnifications.

CHAPTER 7

QUANTITATIVE ANALYSIS OF COLLAGEN ALIGNMENT

7.1 Quantification of collagen alignment distributions

The directional filtering of collagen fiber images in rabbit ligaments gives a quantitative distribution of fibril-covered areas in different angle bands. It has been demonstrated that such distributions of fibrils vary significantly with different stages of healing of injured ligaments, and that the fibrils gradually re-align themselves along the longitudinal axis as the ligaments heal. These distributions have been represented by using rose diagrams. Different statistics derived from the rose diagrams may be used to efficiently quantify the alignment of collagen fibrils at different stages of healing.

Statistical measures that are commonly used for the analysis of data points in rectangular coordinate system may lead to improper results if applied to circular distribution of data[47]. Different methods of analyzing directional data could be found in references[46-49]. In the present application, the fibrils are not directed vectors, and there is no need to differentiate between vectors at angles θ and $\theta \pm 180^o$. Thus, the fibril orientations have been restricted to a semicircular space (total angular bandwidth of 180°). According to Batschelet[49], such a distribution of angular data need not be considered as a circular one, and the distribution could be assumed to be linear. Then, the usual methods of statistical analysis of data

points in Cartesian coordinate system are applicable. Such statistical measures, as discussed in the following sections provide a meaningful insight of the alignment of collagen fibrils at different stages of healing.

7.2 Entropy of collagen distributions

The concept of entropy in information theory[50] could be very effectively applied to the rose diagrams to determine the randomness of the distributions. If x_i is the relative fibril-covered area along the ith angle band, the entropy(H) of the distribution may be defined as

$$H = -\sum_{i=1}^{12} x_i \log_2 x_i. \tag{7.1}$$

In normal ligaments, most of the fibrils lie within one or two major angle bands, whereas in scar tissues the fibrils are scattered randomly in all possible directions. The entropy of a distribution is the maximum when all the events within the distribution are equally likely. Hence, the distributions corresponding to scar tissues should have higher entropies. As a ligament heals, more and more fibrils align themselves with the longitudinal axis, and the corresponding distributions should have smaller entropies.

The entropies for all the fiber images analyzed were calculated, and are plotted as isolated points with respect to healing interval in Figure 7.1. For the purpose of illustration, the corresponding measures for fiber images belonging to

normal ligaments are plotted at $t = 0$ (before sustaining the injury). The average distribution for a set of fiber images belonging to a particular healing interval was obtained by adding and normalizing the absolute values of the fibril-covered areas in a particular angle band in all the images. The entropies for the average distributions of fibril-covered areas at different healing stages are also plotted in the same figure as small circles joined by straight lines to indicate the trend.

The measure, as expected, decreases monotonically with healing time. However, there is no significant improvement during the period of third to sixth week of healing. It may be noticed in Figure 7.1 that at earlier stages of healing, the entropy of the average distribution of a set of fiber images happens to be greater than that for individual images. This may be attributed to the fact that the individual distributions within the set of images analyzed may have a stronger concentration of fibrils in some of the angle bands, but when averaged over the entire set of images, the individual variations get smoothed out, leading to a more uniform (or random) distribution. It should be noted that the measure of entropy does not reflect directly the degree of non-alignment with the principal axis of the collagen fibrils. However, it gives a true picture of the randomness (or uniformity) of the distribution.

7.3 Angular moments for quantification of dispersion

Angular moments of the rose diagram data may be used to determine the angular dispersion of the fibrils in the image. The angular moment M_k of order k is defined as

$$M_k = \sum_{i=1}^{12} \theta_i^k x_i \qquad (7.2)$$

where θ_i represents the center of the ith angle band in degrees. Since we are interested in determining the dispersion of fibrils from their principal axis, the moments may be taken with respect to the centroidal angle $\bar{\theta}$ for the distribution. Evidently, $\bar{\theta} = M_1$. Since the second order moment is always the minimum when taken about the centroid, we choose $k = 2$ for the statistical analysis of the rose diagrams. Hence, the second central moment M_2 may be defined as

$$M_2 = \sum_{i=1}^{12} (\theta_i - \bar{\theta})^2 x_i. \qquad (7.3)$$

The second angular moments for all fiber images were calculated and are plotted as isolated points with respect to healing interval in Figure 7.2. The moments for the average distribution of fibril-covered areas at different healing stages are also plotted in the same figure as small circles and joined. This measure also, as expected, decreases monotonically with healing time.

Unlike entropy, the use of M_2 as a measure of angular dispersion has one pitfall. The fibrils may be well aligned in a normal ligament, but a few stray fibrils
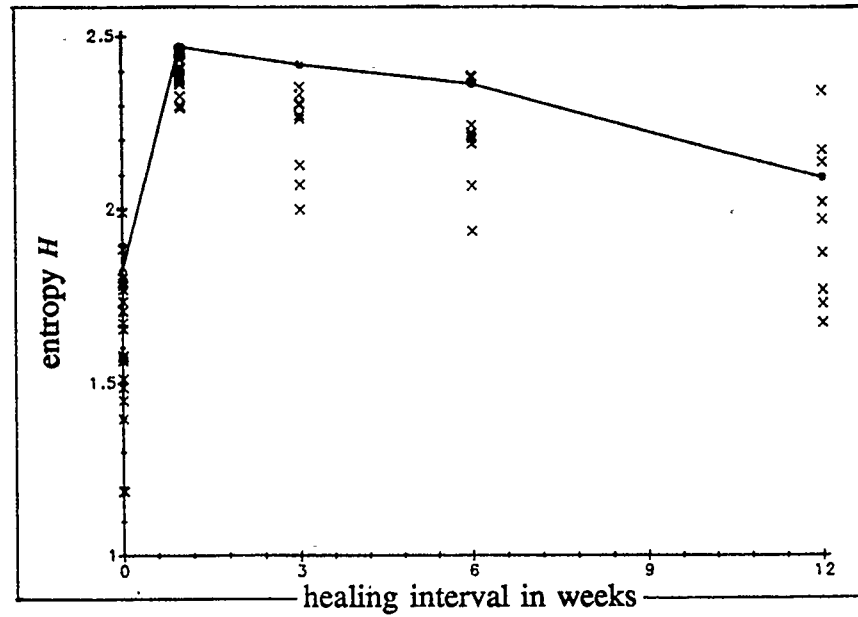
Fig. 7.1 : Entropy of the collagen distribution in ligaments at different stages of healing.
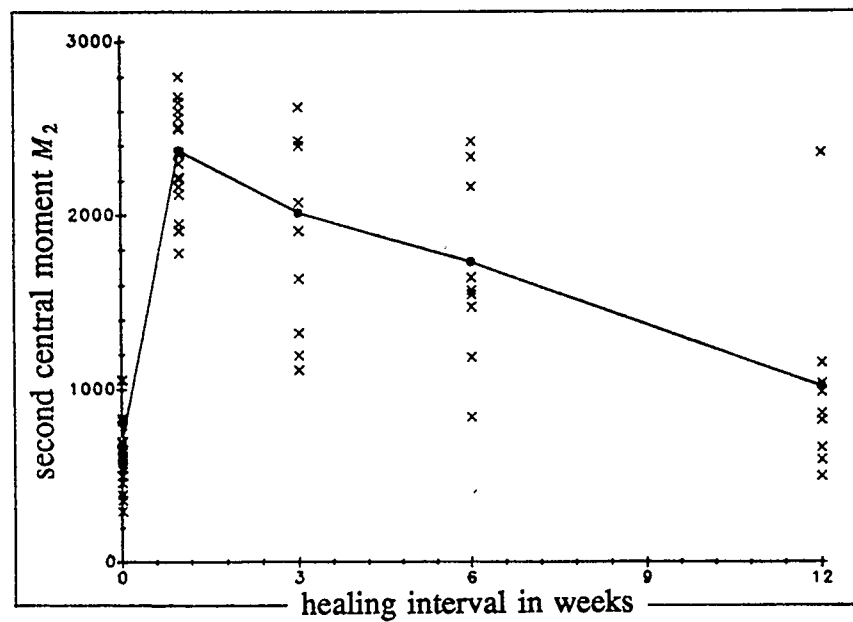
Fig. 7.2 : Quantification of collagen remodeling process in healing ligaments using the second central moment.

at a larger angular distance may result in a very high value for M_2. Theoretically, the moment is maximum when the fibrils are lying equally in the first and in the twelfth sectors only! This probably means that while the fibrils are highly aligned, the image was rotated by 90° during digitization. The entropy measures are free from such problems. Nevertheless, M_2 serves to be a very powerful tool in quantifying collagen alignment. As seen from Figure 7.2, $M_2 < 800$ corresponds to normal ligament, while $M_2 > 2000$ indicates scar tissue. Intermediate values of M_2 directly relate to the extent of healing.

7.4 Cross-correlation between class distributions

It was mentioned that fibrils in scar tissues tend to re-align themselves with healing. Hence, the collagen distribution in ligaments after a longer healing interval should have a greater correlation with the average distribution in normal ligaments.

The values of cross-correlation R as defined in equation (6.1) were calculated for fiber images belonging to different healing stages with respect to the average distribution in normal ligament, and are plotted as isolated points with respect to healing interval in Figure 7.3. The cross-correlation coefficients for the average distributions at different healing stages are also plotted in the same figure as small circles and joined. This measure, as expected, increases with the healing time.

It may be noticed from Figure 7.3 that the individual variations in cross-correlation coefficients within the one week healing tissue samples are small, indicating that the distributions at this stage of healing are scattered uniformly
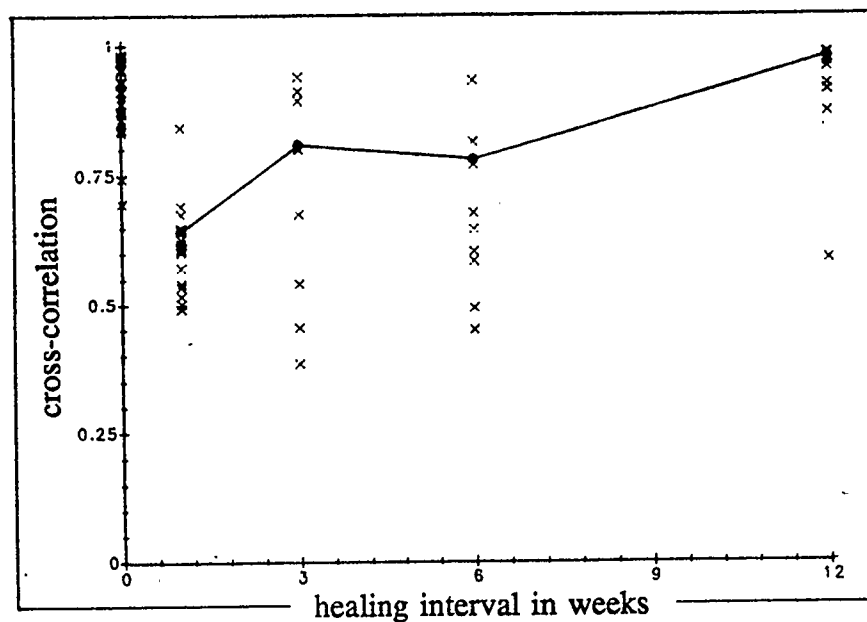
Fig. 7.3 : Cross-correlation between distributions at different stages of healing and the average normal distribution.
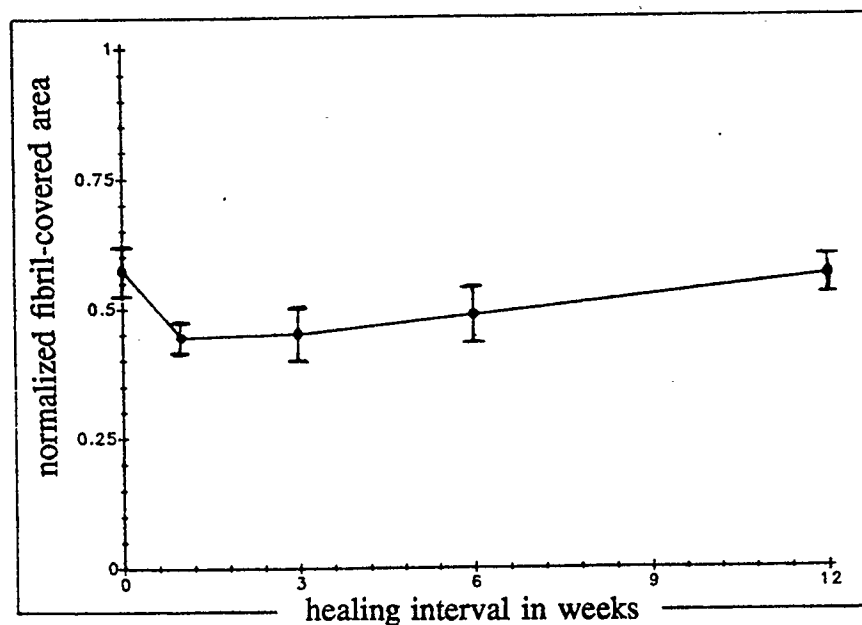
Fig. 7.4 : Plot showing an increase in total fibril-covered area in ligaments with healing. The small circles represent the mean, and the vertical lines indicate the standard deviation of fibril-covered area for all the images in the particular healing class.

along all possible directions, and that the collagen re-alignment process has not yet started. Since the healing process may be attributed to the fibroblastic infiltration from neighboring tissues[1, 6], the constituent collagen fibrils, at the initial stages of healing, may be oriented along different directions at different places in the ligament. A large variation in the values of R (but relatively smaller values of angular dispersions) for the healing intervals of three and six weeks as seen in Figures 7.2 and 7.3 supports this theory of healing. At the later stages of healing, smaller variations in the values of R in the images suggests that the fibrils tend to form a parallel arrangement.

Three different measures (viz., entropy, second central moment, and cross-correlation) have been used in this research work to quantify the collagen re-alignment process, and each of them has its own merits and demerits. Studies related to the comparison of performances of these measures have not been taken up, as we are more interested in studying the collagen remodeling process for different treatment modalities as indicated by any one of these measures.

7.5 Significance of fibril-covered area

The fibril-covered area in a tissue image may correlate with the mechanical strength that may be provided by the ligament being analyzed. For example, between two ligaments at the same healing stage, one having more fibrils per unit area will have greater total strength. The fibril-covered areas in the different angle bands were added and then divided by the area of the image (256 x 256) to obtain

the total normalized fibril-covered area for an individual image. The mean and standard deviation of the normalized fibril-covered area for each set of images representing a certain healing stage were calculated. The mean values are plotted as small circles and the standard deviations are plotted as vertical lines about their mean values in Figure 7.4. An increase in the average fibril-covered area suggests an increased concentration of collagen in healing ligaments, and lower values of standard deviation justify this interpretation.

CHAPTER 8

DISCUSSION

8.1 Artifacts due to filtering

The performance of Fourier domain directional filters was studied in detail, regarding selectivity, influence of certain image features and artifacts. Good results were obtained for angle bands of 10° or more. Use of smaller angle bands led to increasing inaccuracies in the filtered components. Because of the finite width of line segments (that make up the fibrils), the components in the frequency domain do not lie completely within the corresponding sectors. As the thickness of the fibrils increases, there is an increased spreading of the Fourier components into adjoining angle bands, leading to reduced angular selectivity. This results in poor contrast in the filtered components, as well as interference due to line segments lying in the neighboring sectors. Removal of an experimentally determined low frequency zone helped in overcoming this problem to some extent. However, if a large amount of low frequency components in the spectrum is removed, the necessary background information is suppressed, producing only the edges of the line segments. Since it is intended to recover fibril-covered area from the composite image, special care has to be taken in selecting the lower cutoff frequency for the Butterworth bandpass filter given in equation (4.9). Some improvement in the results was obtained by deemphasizing the high frequency

components.

Depending on the width of the fibril bundles, the filtered components may suffer from smearing artifacts at both ends of the segments. This is more true for the second test pattern (Figure 4.10), which has sharp edges at the ends of the line segments, because the filter sector does not include information about the edges which are in a direction orthogonal to the length of the lines. However, this effect is not pronounced for line segments having higher length-to-width ratios, as is apparent from the error analysis in Table 2. In real fiber images, thick fibril bundles usually do not end abruptly, and this effect is not a serious problem. A related artifact observed is that a small discontinuity between two fibril segments lying within the same angle band, in some cases, appeared to be continuous in the filtered image.

The use of different values of f_L in equation (4.9) for analyzing fiber images corresponding to various healing stages yielded different results. The angular separability being lower for the images of scar tissues, a higher value of f_L produced a better result. This suggests that f_L is a parameter which could be optimized for each case. However this was not done in the present study to maintain consistency and avoid bias.

8.2 Problems with thresholding

A number of automatic thresholding schemes[8, 34, 40, 41, 43, 44] were tried for binarization of the filtered images, but all met with only limited success. The

three major difficulties are :

(1) The component images do not always have good contrast, particularly when the fibrils are thick.

(2) Histograms of the filtered component images are almost always unimodal. In most cases, the histogram of the filtered image has a long trail corresponding to the background. Such a long trail of grey levels of small number of occurrences makes probabilistic relaxation methods[8] inapplicable.

(3) Fibril-covered areas in different directions differ significantly. A very low population of pixels corresponding to the fibril-covered areas makes histogram-based automatic thresholding schemes[41, 43, 44] fail.

It was found that Otsu's method for automatic threshold selection[41] was most suitable for the present application. The method worked very well for the test pattern. It worked well for real fiber images as well, when there was a significant population of fibril pixels in the filtered image. However, it failed when the filtered images had very few fibril pixels, giving lower thresholds. An interactive manual procedure was used to determine the actual thresholds for the image components, starting with the automatically determined value given by Otsu's method as an initial estimate of the threshold. In all cases, the actual thresholds used were higher than those given by Otsu's method.

Use of manual thresholds may introduce a certain amount of subjectivity, and thus the results may be affected by personal biases. Although this problem cannot

be tackled completely, every endeavor was made to keep the personal inclinations as minimal as possible.

8.3 Computational requirements

The processing of an image involves one forward FT and twelve inverse FTs. The computation of the 2-D FT of an image of size 512 x 512 takes considerable time, and hence this method is computationally expensive. Also, two complex arrays of size 512 x 512 are required to run the program, one for storing the forward FT values and the other for storing the inverse FTs. Hence, the algorithm requires a substantial central memory to avoid the problem of page-swapping.

The program was run on a VAX 11/750 computer employing UNIX operating system with only 2Mb of core memory to find out its computational requirements. The execution of the program was hampered by a large number of page faults. It took about 5 hours of CPU time (about 12 hours of real time) for the completion of the program. However, when the program was run on a Sun 3/180 computer employing the same operating system but having 16Mb of central memory, it took about 90 minutes of CPU time (about 2 hours of real time). Significant improvement was achieved by running the program on the C.D.C. Cyber - 205 Supercomputer. This system makes use of vector programming features (FORTRAN 200), and tries to vectorize the programs automatically. Further reduction in computing time was achieved by manually vectorizing some parts of the program. The program took only 8 minutes of CPU time and about 30 minutes

of real time to run on the Cyber - 205. All the results presented here were obtained by running the program on the Supercomputer.

The computational complexity of a one dimensional N-point fast Fourier transform (FFT) is proportional to $N\log_2 N$, and that of a 2-D FFT for an image of size N x N is proportional to $N^2\log_2 N$. In order to have higher spectral resolution, the FFT array size was increased from 256 x 256 to 512 x 512. This increases the memory requirements by a factor of four, and the resulting increase in computation is given by

$$k = \frac{512^2 \log_2 512}{256^2 \log_2 256} = 4.5.$$

Since the spectrum is divided into twelve equal sectors, the filtered spectra are almost 90% sparse. Thus, sparse matrix techniques may be used to reduce the memory requirements. Suitable pruning algorithms[51,52] may be also used to eliminate computations involving the zeros. However, the sector filters have varying orientations, and in order to be able to use the modified FFT algorithms, the addresses for nonzero elements have to be determined each time the FT is evaluated for different rows and columns. This will increase the software overhead significantly. Also, to take advantage of the vector processing facilities on the Supercomputer, it is necessary that the data be contiguous in memory. Hence these methods may not be useful in the present situation.

8.4 Artifacts due to tissue preparation

While there is a certain degree of three-dimensionality to the micrographs analyzed, the contribution of fibril components at varying depths to the results is affected by a number of factors. Such factors that have been standardized include the intensity of the electron beam in the microscope, depth of focus, photographic methods, and image digitization parameters. However, the final thresholding scheme applied to the filtered component images, being adaptive in nature, will affect the various component images differently depending upon their contents. This aspect cannot be controlled without introducing some bias.

The process of quantification of surface orientations of collagen fibrils may also be adversely affected by artifacts due to improper fixation and handling of tissues (i.e., fiber disruption, surface irregularities, etc.). The aims of fixation are rapid preservation of structure with minimum alteration from the living state, and protection during embedding, sectioning and subsequent treatments[53]. But, depending on the nature of the fixative, there may be some loss of "ground substance" and shrinkage of fibrous components during fixation. Also, the depth of penetration of different fixatives into the tissue samples may be different. These factors may affect the quantification procedure.

Collagen fibrils vary widely in diameter from 16nm to several hundred nanometers in different connective tissues. Proper magnification should be chosen while acquiring the images to retain the required resolution.

8.5 Scope for future research

The definition of angle in the rectangular coordinate system for low frequency components is inaccurate, leading to artifacts in the filtered images. Two other procedures that may lead to better results are :

(1) The Fourier space samples in the low frequency zone within the filter sector may be obtained by extrapolation using the samples available in the higher frequency zone where angular selectivity is better. Related procedures are commonly used in image reconstruction from projections (computed tomography) by the Fourier method[8, 54-62]. Many elaborate algorithms have been proposed for this purpose. This procedure, however, would increase the computation time appreciably.

(2) The FT may be evaluated directly on discretized polar coordinate points[59]. A sector of interest may be sampled in Fourier domain with any desired resolution, both in angle and radial distance (subject to practical computational times). However, standard FFT algorithms cannot be used for such computations.

The quantitative analysis procedure reported may be repeated under different treatment modalities as well, to study the rate of healing under various conditions. The information so gained should help clinicians design an optimum scheme for treatment of injured ligaments.

As the tissue imaging procedure is destructive, it cannot be used for diagnostic purposes to determine the nature of injury sustained or the progress of healing in human ligaments. However, if appropriate *in vivo* imaging systems (e.g., using optical fibers) were ever devised to photograph tissues, the method could be extended for clinical diagnosis applications.

REFERENCES

1. K. A. Piez and A.H. Reddi, *Extracellular Matrix Biochemistry,* Elsevier (1984).

2. M. Nimni, "Collagen: Its structure and function in normal and pathological connective tissues," *Semin. Arthritis Rheum.* **4** pp. 95-150 (1974).

3. J.C. Forrester, B.H. Zederfeldt, T.L. Hayes, and T.K. Hunt, "Tape-closed and sutured wounds: A comparison by tensiometry and scanning electron microscopy," *Br. J. Surg.* **57** pp. 729-737 (1970).

4. C. Frank, D. Amiel, S.L-Y. Woo, and W. Akeson, "Normal ligament properties and ligament healing," *Clin. Orthop.* **196** pp. 15-25 (1985).

5. C. Frank, S.L-Y. Woo, D. Amiel, F. Harwood, M. Gomez, and W. Akeson, "Medial collateral ligament healing: A multidisciplinary assessment in rabbits," *Am. J. Sports Med.* **11** pp. 379-389 (1983).

6. R. M. H. McMinn, *Tissue Repair,* Academic Press (1969).

7. C. Frank, N. Schachar, and D. Dittrich, "Natural History of Healing in the Repaired Medial Collateral Ligament," *J. Orthop. Res.* **1** pp. 179-188 (1983).

8. A. Rosenfeld and A.C. Kak, *Digital Picture Processing- Vol 1 and 2,* Academic Press, New York (1982, 2nd Ed.).

9. G. H. Granlund, "Description of Texture using the General Operator Method," *IEEE Conf. on Pattern Recognition*, p. 776 (Dec. 1980).

10. E.L. Hall, *Computer Image Processing and Recognition*, Academic Press, New York (1979).

11. W.K. Pratt, *Digital image processing*, Wiley, New York (1978).

12. G.K. Moore and F.A. Waltz, "Objective procedures for lineament enhancement and extraction," *Photogrammetric Engg. and Remote Sensing* **49** pp. 641-647 (1983).

13. H. Takahashi, "A lineament enhancement technique for active fault analysis," pp. 103-112 in *Machine Processing of Remotely Sensed Data Symposium*, (1981).

14. R.C. Gonzalez and P. Wintz, *Digital Image Processing*, Addison-Wesley, New York (1979).

15. H. Nguyen, R. M. Rangayyan, S. Walsh, and C. B. Frank, "Directional Filtering for Quantification of Collagen Alignment in Normal Ligaments," *Proc. of 12th Northeast Bioengineering Conf.*, pp. 141-145 (Mar. 13-14, 1986).

16. E.O. Brigham, *The fast Fourier transform*, Prentice-Hall, Englewood Cliffs (1974).

17. R.N. Bracewell, *The Fourier transform and its applications* , McGraw-Hill, New York (1978- 2nd ed.).

18. P. Embree, J. P. Burg, and M. M. Bacus, "Wide-band Velocity Filtering - The Pie-Slice Process," *Geophysics* **28** pp. 948-974 (1963).

19. A. Dziedzic-Goclawska, M. Rozycka, J. C. Czyba, W. Sawicki, R. Moutier, S. Lenczowski, and K. Ostrowski, "Application of the Optical Fourier Transform for Analysis of the Spatial Distribution of Collagen Fibres in Normal and Osteopetrotic Bone Tissue," *Histochemistry* **74** pp. 123-137 (1982).

20. V. Bezvoda and K. Segeth, "Directional and frequency filtering of geophysical data measured in a rectangular net," *Gerlands Beitr. Geophys. (Germany)* **90** pp. 133-146 (1981).

21. S. Treitel, J. L. Shanks, and C. W. Frasier, "Some Aspects of Fan Filtering," *Geophysics* **32** pp. 789-806 (1967).

22. L. T. Bruton, N. R. Bartley, and R. A. Stein, "The Design of Stable High-quality Two-dimensional Recursive Filters for Seismic Signal Processing," *Advances in Geophysical Data Processing* **2** pp. 223-261 (1985).

23. M. B. Dobrin, A. L. Ingalls, and J. A. Long, "Velocity and Frequency Filtering of Seismic Data using Laser Light," *Geophysics* **30**(6) pp. 1144-1178 (Dec. 1965).

24. M. B. Dobrin, "Optical Processing of Underground Sound Waves in Seismic Exploration for Oil," *Applied Optics* **8**(8) pp. 1551-1557 (Aug. 1969).

25. P. L. Jackson, "Diffractive Processing of Geophysical Data," *Applied Optics* **4**(4) pp. 419-427 (Apr. 1965).

26. H. H. Arsenault, M. K. Seguin, and N. Brousseau, "Optical Filtering of Aeromagnetic Maps," *Applied Optics* **13**(5) pp. 1013-1017 (May 1974).

27. A. Ikonomopoulos and M. Kunt, "Directional filtering, zero crossing edge detection, and image coding," pp. 203-206 in *Signal Processing II: Theories and Applications (EURASIP 1983)*, ed. H.W. Schussler, Elsevier Science Publishers (1983).

28. A. Ikonomopoulos and M. Unser, "A directional filtering approach to texture discrimination," pp. 87-89 in *Proc. 7th Intl. Conf. on Pattern Recognition (Montreal)*, IEEE (1984).

29. S. Chaudhuri, H. Nguyen, R. M. Rangayyan, S. Walsh, and C. B. Frank, "A Fourier Domain Directional Filtering Method for Analysis of Collagen Alignment in Ligaments," *IEEE Trans. on Biomedical Engineering*, (In press, 1987).

30. F. J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proc. IEEE* **66**(1) pp. 51-83 (Jan. 1978).

31. J. W. Tukey, "An Introduction to the Calculations of Numerical Spectrum Analysis," pp. 25-46 in *Spectral Analysis of Time Series*, ed. B. Harris, John Wiley & Sons, Inc. (1966).

32. J. Duvernoy and K. Chalasinka-Macukow, "Processing measurements of the directional content of Fourier spectra," *Applied Optics* **20** pp. 136-144 (1981).

33. G. Leboucher and G. E. Lowitz, "What A Histogram Can Really Tell The Classifier," *Pattern Recognition* **10** pp. 351-357 (1978).

34. J. S. Weszka and A. Rosenfeld, "Histogram Modification for Threshold Selection," *IEEE Trans. on Systems, Man and Cybernetics* **SMC-9**(1) pp. 38-52 (Jan. 1979).

35. P. A. Dondes and A. Rosenfeld, "Pixel Classification Based on Gray Level and Local Busyness," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **PAMI-4**(1)(Jan. 1982).

36. A. Y. Wu, T. Hong, and A. Rosenfeld, "Threshold Selection Using Quadtrees," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **PAMI-4**(1)(Jan. 1982).

37. J. Kittler, J. Illingworth, and J. Foglein, "Threshold Selection Based on a Simple Image Statistic," *Computer Vision, Graphics and Image Processing* **30** pp. 125-147 (1985).

38. F. Deravi and S. K. Pal, "Grey level thresholding using second-order statistics," *Pattern Recognition Letters* **1** pp. 417-422 (July 1983).

39. S. Peleg, "A New Probabilistic Relaxation Scheme," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **2**(4) pp. 362-369 (July, 1980).

40. N. Ahuja and A. Rosenfeld, "A Note on the Use of Second-Order Gray-Level Statistics for Threshold Selection," *IEEE Trans. on Systems, Man and Cybernetics* **SMC-8**(12) pp. 895-898 (Dec. 1978).

41. N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. on Systems, Man and Cybernetics* **SMC-9**(1) pp. 62-66 (Jan. 1979).

42. T. Pun, "Entropic Thresholding, A New Approach," *Computer Graphics and Image Processing* **16** pp. 210-239 (1981).

43. 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, and Image Processing* **29** pp. 273-285 (1985).

44. W. Tsai, "Moment Preserving Thresholding : A New Approach," *Computer Visions, Graphics and Image Processing* **29**(3) pp. 377-393 (March, 1985).

45. R. Kohler, "A Segmentation Based on Thresholding," *Computer Graphics and Image Processing* **15** pp. 319-338 (1981).

46. K.V. Mardia, *Statistics of directional data*, Academic Press, New York (1972).

47. J. R. Curray, "The Analysis of Two-dimensional Orientation Data," *J. Geol.* **64** pp. 117-131 (1956).

48. G. S. Watson, "The Statistics of Orientation Data," *J. Geol.* **74** pp. 786-97 (1966).

49. E. Batschelet, *Statistical Methods for the Analysis of Problems in Animal Orientation and certain Biological Rhythms*, American Institute of Biological Sciences (1965).

50. F. M. Reza, *An Introduction to Information Theory,* McGraw Hill (1961).

51. D. P. Skinner, "Pruning the Decimation In-Time FFT Algorithm," *IEEE Trans. Acoust., Speech, Signal Processing* **ASSP-24** pp. 193-194 (Apr. 1976).

52. T. V. Sreenivas and P. V. S. Rao, "FFT Algorithm for Both Input and Output Pruning," *IEEE Trans. Acoust., Speech, Signal Processing* **ASSP-27** pp. 291-292 (Jun. 1979).

53. D. Hopwood, "Theoretical and practical aspects of gluteraldehyde fixation," pp. 47-84 in *Fixation in Histochemistry,* ed. P. J. Stoward, Chapman and Hall, London (1973).

54. G.T. Herman, *Image Reconstruction From Projections: The Fundamentals Of Computed Tomography,* Academic Press, New York (1980).

55. E.W. Hansen, *Image reconstruction from projections using circular harmonic expansion,* Stanford University (Ph.D. Thesis), Stanford (1979).

56. T. Inouye, "Analysis of artifact pattern generation in two-dimensional Fourier transform image reconstruction algorithm for computerized tomography," pp. 170-175 in *Proc. 1984 Intl. Symp. Noise and Clutter Rejection in Radars and Image Sensors,* ed. H. Ogura, (1984).

57. R.M. Mersereau, "Direct Fourier transform techniques in 3-D image reconstruction," *Comput. Biol. Med.* **6** pp. 247-258 (1976).

58. M.H. Buonocore, W.R. Brody, and S.L. Wood, "Polar pixel Kalman filter for limited data for limited data computed tomography (CT) image

reconstruction," pp. 109-114 in *Proc SPIE*, (1979).

59. R.M. Mersereau and A.V. Oppenheim, "Digital reconstruction of multidimensional signals from their projections," *Proc. IEEE* **62** pp. 1319-1338 (1974).

60. R.A. Crowther, D.J. DeRosier, and A. Klug, "The reconstruction of a three dimensional structure from projections and its application to electron microscopy," pp. 319-340 in *Proc. Roy. Soc. Ser.*, (1970b).

61. S.J. Wernecke and L.R. D'Addario, "Maximum entropy image reconstruction," *IEEE Trans. Computers* **26** pp. 351-364 (1977).

62. A.R. Thompson and R.N. Bracewell, "Interpolation and Fourier transformation of fringe visibilities," *The Astronomical J.* **79** pp. 11-24 (1974).