

APPLICATIONS OF MULTI-CHANNEL FILTER BANKS TO TEXTURED IMAGE
SEGMENTATION

Except where reference is made to the work of others, the work described in this thesis is my own or was done in collaboration with my advisory committee. This thesis does not include proprietary or classified information.

Craig Alton Davis

Certificate of Approval:

Stanley J. Reeves
Professor
Electrical Engineering

Thomas S. Denney Jr., Chair
Associate Professor
Electrical Engineering

A. Scottedward Hodel
Associate Professor
Electrical Engineering

Stephen L. McFarland
Acting Dean
Graduate School

APPLICATIONS OF MULTI-CHANNEL FILTER BANKS TO TEXTURED IMAGE
SEGMENTATION

Craig Alton Davis

A Thesis

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Master of Science

Auburn, Alabama
August 7, 2006

APPLICATION OF MULTI-CHANNEL FILTER BANKS TO TEXTURED IMAGE
SEGMENTATION

Craig Alton Davis

Permission is granted to Auburn University to make copies of this thesis at its discretion, upon request of individuals or institutions and at their expense. The author reserves all publication rights.

Signature of Author

Date of Graduation

VITA

Craig Alton Davis, son of Rickie Davis and Emily (Wilson) Davenport, was born on December 15, 1978, in Alexander City, Alabama. He graduated from Benjamin Russell High School in 1997. Upon graduation, he attended Central Alabama Community College in Alexander City for two years and graduated magna cum laude with an Associates of Science degree. He then entered Auburn University in September, 1999, and graduated cum laude with a Bachelor degree in Electrical Engineering in May 2003. Upon attainment of his Bachelor degree, Craig entered Graduate School at Auburn University in September 2003 and obtained a Masters Degree in Business Administration in December 2004. In September 2003, he continued his Graduate Education by entering the Master of Science program in Electrical Engineering at Auburn University.

THESIS ABSTRACT

APPLICATIONS OF MULTI-CHANNEL FILTER BANKS TO TEXTURED IMAGE SEGMENTATION

Craig Alton Davis

Master of Science, August 7, 2006
(M.B.A., Auburn University, 2004)
(B.S. Auburn University, 2003)

121 Typed Pages

Directed by Thomas S. Denney Jr.

Multi-channel filtering provides a robust solution to the texture discrimination/segmentation problem for digital image processing. Specifically, the use of the Gabor filter for channel characterization is explored. Tag deformation analysis is used to construct a new tag tracking pre-filter based on the multi-channel paradigm. The new pre-filter is experimentally shown to be superior to two other currently used pre-filters. Also, the use of Gabor filters for channel characterization is applied to tagged cardiac MRI, cine cardiac MRI and segmentation of asphalt core images. In the applications involving MRI, segmentation algorithms are developed and tested for optimal parameters. The segmentation algorithm developed for asphalt core images is intended to be used in subsequent steps involving an aggregate orientation analysis.

ACKNOWLEDGMENTS

The author would like to thank Dr. Thomas S. Denney for his guidance throughout the writing of this manuscript. Thanks are also due to family members Emily Davenport and Rickie Davis for their support during the course of this investigation.

Style manual or journal used: IEEE Standards Style Manual

Computer Software Used: Matlab, Microsoft Word, Microsoft Excel, and Adobe Illustrator

TABLE OF CONTENTS

LIST OF FIGURES.....	xi
LIST OF TABLES.....	xv
1 INTRODUCTION	1
1.2 Texture Segmentation and Multi-Channel Filtering	2
1.2.2 Gabor Filters	6
1.2.3 Multit-Channel Filter Banks	9
1.2.4 Gray Level Co-Occurrence Matrix Features.....	11
1.3 Background on Heart Anatomy and Physiology	13
1.3.1 Heart Anatomy.....	13
1.3.2 Heart Physiology.....	15
1.4 Magnetic Resonance Imaging.....	16
1.4.1 Cardiac MRI.....	18
1.4.2 Cine MRI	18
1.4.3 Tagged MRI	20
1.5 Overview of Thesis	21
2 THE C-GABOR FILTER	23
2.1 Tag Deformation Analysis.....	23
2.1.1 Shear Model	26
2.1.2 Rotation Model	28

2.1.4	Jones Deformation Model.....	30
2.1.5	Sum of Shifted Gaussians Model.....	33
2.2	C-Gabor Filter Derivation.....	36
2.3	Filter Experiments and Results	37
2.3.1	Experiments	38
2.3.2	Results.....	39
2.4	Chapter Summary	42
3	TAG LINE IDENTIFICATION FOR TAGGED CARDIAC MRI	43
3.1	Introduction.....	43
3.2	Tag Line Identification via Multi-Channel Segmentation Algorithm	44
3.2.1	Gabor Feature Images	44
3.2.2	Combining Gabor Features	49
3.3	Tag Line Identification via Peak/Valley Detection Algorithm.....	53
3.3.1	Derivation	53
3.3.2:	QuickTag Algorithm.....	55
3.4	Chapter Summary	64
4	MULTI-CHANNEL SEGMENTATION OF MYOCARDIUM IN CARDIAC MRI IMAGES	65
4.1	Introduction.....	65
4.2	Methodology	66
4.2.1	Pre-Processing.....	66
4.2.2	Channel Characterization.....	67
4.2.3	Channel Post Processing.....	70

4.2.4	Multi-Channel Bank Construction.....	71
4.2.5	Vector Classification.....	72
4.3	Experiments and Results.....	73
4.4	Chapter Summary	78
5	SEGMENTATION OF ASPHALT AGGREGATE FROM MASTIC	79
5.1	Introduction and Background	79
5.2	Experimental Procedure.....	80
5.3	Segmentation of Aggregate from Mastic.....	82
5.3.1	Pre-Processing.....	84
5.3.2	Multi-Channel Segmentation.....	86
5.3.3	Post-Processing.....	95
5.4	Aggregate Orientation Analysis.....	97
5.4.1	Aggregate Orientation Analysis Definitions.....	97
5.4.2	Aggregate Orientation Analysis Results.....	99
5.5	Chapter Summary	99
6	CONCLUSION.....	101
6.1	Thesis Summary.....	101
6.2	Directions for Future Work.....	102
	REFERENCES	104

LIST OF FIGURES

Figure 1.1: Heart anatomy	15
Figure 1.2: Pulmonary circuit representation of the circulatory system.....	17
Figure 1.3: Pictorial representation of the standardized imaging views as defined by the American Heart Association.....	19
Figure 1.4: Short axis cine cardiac MRI images. End-diastole shown left and end-systole shown right	20
Figure 1.5: Short axis tagged cardiac MRI images. End-diastole shown left and end- systole shown right	21
Figure 2.1: Myocardium model (top-left); tags(top-right); tagged myocardium model (bottom-left); and log frequency spectrum (bottom-right)	26
Figure 2.2: Sheared myocardium model (left) and resulting log frequency spectrum (right)	28
Figure 2.3: Rotated myocardium model (left) and resulting log frequency spectrum (right)	29
Figure 2.4: Myocardium with decreased tag spacing (left) and resulting log frequency spectrum (right).....	30
Figure 2.5: Original (left) and deformed (right) Jones model	32
Figure 2.6: Log magnitude plot of original (left); deformed (middle); and absolute	

difference frequency spectrum (right)	32
Figure 2.7: Frequency domain regions predicted by the deformation model to contain spectral changes due to local deformation of the myocardium	36
Figure 2.10: Mean absolute error [mm] (top) and percent false positives (bottom) of C-Gabor (left), Gabor (center); and HARP (right) filters	41
Figure 2.11: Impulse response (top) and magnitude frequency response (bottom) of C-Gabor (left; $\sigma=0.4$, $\Delta\theta=40^\circ$), standard Gabor (middle, $\sigma_y=0.6$, $\sigma_x=0.25$), and HARP (right, $b=0.5$, $a=0.4$) filters.	41
Figure 3.1: Grid-tagged short axis image in-systole (top) and magnitude spatial frequency spectrum (bottom).....	46
Figure 3.2: Gabor filter bank used to construct 6 Gabor features.....	47
Figure 3.3: Gabor Features	48
Figure 3.4: Segmentation obtained by clustering $F(l,i)$ into two groups via k-means clustering algorithm	50
Figure 3.5: Set of identified tag pixels after morphological skeletal thinning	51
Figure 3.6: Original image(left) and original image with identified candidate tag pixels (right)	51
Figure 3.7: Block diagram of multi-channel tag line identification algorithm.....	52
Figure 3.8: Original Image (left) and pre-filtered image (right) for a tag angle of 45 degrees. The endocardium and epicardium are outlined in green	54
Figure 3.9: First Derivative shown left and second derivative shown right	54
Figure 3.10: Tags are identified as areas in the image above where the red contours are surrounded by white.....	55

Figure 3.11: Example image to demonstrate proposed algorithm. Green contours outline the left ventricle. Right image shows only the area of interest outlined with the dashed white line in the left image.	56
Figure 3.12: Original Image (left) and pre-filtered image (right) for a tag angle of $\theta_t = 45$ degrees	56
Figure 3.13: $f_{fp}(\mathbf{p})$ and $f_{fn}(\mathbf{p})$ for a tag angle of 45 degrees	58
Figure 3.14: $f_{bp}(\mathbf{p})$ and $f_{bn}(\mathbf{p})$ for the example image	59
Figure 3.15: $f_t(\mathbf{p})$ shown left and $f_t(\mathbf{p})$ superimposed over original image, $f(\mathbf{p})$ shown right	60
Figure 3.16: Ω_t superimposed over original image, $f(\mathbf{p})$	61
Figure 3.17: Ω_T superimposed over $f(\mathbf{p})$	62
Figure 3.18: 2-Chamber View (left) and 4-Chamber view (right) of tags detected with the QuickTag algorithm	62
Figure 3.19: QuickTag algorithm	63
Figure 4.1: Original image (left) and intensity corrected image (right)	67
Figure 4.2: Block diagram of ith channel characterization. Note that HWR stands for half-wave rectification.	70
Figure 4.3: GUI written to assist in supervised segmentation	73
Figure 4.4: Average P_f shown top and average P_d shown bottom. White circles indicate the “optimal” parameters.	76
Figure 4.5: Original Images shown left, segmentations shown middle, and segmentations laid over original images and displayed in red shown right.	77
Figure 5.1: Typical asphalt core samples. Field core from Alabama roadway shown top	

and field from Colorado roadway shown bottom.	83
Figure 5.2: Effect of intensity equalization. Original image shown top and intensity equalized image shown bottom.....	85
Figure 5.3: Spatial frequency contour plot of Gabor filter bank used in this segmentation. White contours indicate the -3db bandwidth of each filter. The vertical axis of each plot represents the spatial frequency in the vertical dimension in units of cycles/sample and horizontal axis represents the spatial frequency in the horizontal dimension in units of cycles/sample.	89
Figure 5.4: Response images, $F_r(\mathbf{p},i)$	90
Figure 5.5: Half wave rectified images, $F_{abs}(\mathbf{p},i)$	91
Figure 5.7: Window Averaging Step, $F_a(\mathbf{p},i)$	93
Figure 5.8: Example original image (left) and segmented image (right).....	94
Figure 5.10: Example original image (left) and labeled image after region splitting routine (right).....	96
Figure 5.12: Plots of average absolute angle vs. eccentricity and vector magnitude vs. eccentricity.....	100

LIST OF TABLES

Table 3.1: Gabor filter parameters used to construct feature images	46
Table 3.2: c_p and c_n for standard tag angles	57
Table 4.1 f_i and θ_i used for channel characterization	71
Appendix 4.1	Test Studies
.....	78
Table 5.1: Filter bank parameters of $H(\mathbf{p},i)$	88

CHAPTER 1

INTRODUCTION

1.1 Introduction

Webster's Dictionary defines texture as "the visual or tactile surface characteristics and appearance of something". One might say this is a vague explanation; however, a precise mathematical definition of texture has never been formulated [1]. Such a definition eludes mathematicians due to the infinite number of textures and the wide variety of characteristics any given texture may possess. The most one can say about all textures is that they repeat in some manner over a given area. Therefore, texture is not a local property but a property of a given region. The repetition may range from completely random (natural textures) to highly structured (artificial textures).

A common image processing task involves the labeling of various areas of an image into categories based on texture, gray level, or some other image attribute. This is known as the image segmentation problem. Segmentation based on texture is known as textured image segmentation. Textured image segmentation is a thoroughly explored problem and several solutions exist; however, the focus of this thesis is on a specific solution based on the human visual system, namely, the multi-channel solution.

Medical imaging modalities such as Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) yield images wherein various regions of the images are

textured. The modalities give rise to automatic computer processing of the images, thereby saving time and money. For instance, manual processing of tagged MRI data is labor intensive and, therefore, not feasible in many clinical applications. The same can be said of cine MRI data. Since the various regions of an MRI image are often textured, the segmentation of the areas becomes a textured segmentation problem. Another application of textured image segmentation involves the processing and analysis of asphalt core images. Once again, automatic processing is required here due to the time involved to manually process even a single asphalt core image. In this thesis, the multi-channel solution to textured image segmentation is explored in depth, with application to 1) filtering of tagged cardiac MRI, 2) segmentation of tagged cardiac MRI, 3) segmentation of cine cardiac MRI, and a non-medical application, 4) segmentation of asphalt core images.

In this chapter, background material relevant to this thesis is presented. Texture segmentation and multi-channel filtering is overviewed in Section 2 with specific attention paid to the Gabor filter. Relevant information on heart anatomy and physiology that is useful in discussing cine and tagged cardiac MRI is given in Section 3. The theory of standard MRI is presented in Section 4. Also described in this section are the cine and tagged cardiac MRI modalities. Last, a brief overview of the entire thesis is given in Chapter 5.

1.2 Texture Segmentation and Multi-Channel Filtering

Image segmentation is a fundamental task in many image processing applications. Moreover, textured image segmentation is vital to many image processing applications

wherein the desired goal is to accurately segment the various textured regions of an image. The unpredictable nature of texture makes the task of segmenting textured regions nontrivial and the task of developing a method for all textures even more challenging. According to Tan and Constantinides [3] textured image segmentation can be classified into three broad categories: statistical, structural and model-based. Statistical methods are well suited for natural textures because of their random nature, and structural methods work well for synthetic textures.

After a survey of papers on the subject one will find an overwhelming number of publications on methods based on multi-channel filtering. This approach yields a robust solution to the texture segmentation problem. The multi-channel solution is derived from a model of the early stages of the human visual system (HVS). According to the model, a visual scene is processed through a dense set of quasi-independent spatial filters tuned to various frequencies, orientations, and bandwidths. Higher-level brain functions process and analyze the channel outputs for scene interpretation. Although these higher-level functions remain a mystery, their function is most likely to perform pattern classification. Image segmentation algorithms can be designed to mimic the HVS by filtering an image through a similar set of digital filters. Channel outputs are used as features to be classified by pattern classification algorithms such as k-means clustering or neural network methods.

1.2.1 Gabor Functions

Dennis Gabor first proposed Gabor functions in the paper “Theory of Communication” [4]. The functions were introduced as a carrier signal for information

because they provide a lower bound on the uncertainty relation of signals in time and frequency. The uncertainty relation links the bandwidth of a signal in time and frequency by stating that the effective coverage of a signal in each domain is complimentary, with the exact time-frequency coverage relation being signal dependent. Gabor resolved the uncertainty relation by proving that Gabor functions are optimal in the sense that no other signal has a smaller product of time duration and frequency bandwidth as measured by the effective coverage of the signal in each domain. It should be noted that the Gaussian function is, in fact, optimal in the same sense as the Gabor function. Since the Gabor function is simply a Gaussian multiplied (modulated) by a sinusoid, the optimality of the Gabor function is derived from its unmodulated form, the Gaussian.

Besides their optimality, Gabor functions are desirable in that they are easily tunable to any center frequency and bandwidth. For the 1-D case, the general family of Gabor functions is given as a 1-D Gaussian modulated by a 1-D sinusoid, or,

$$h(t) = \exp\left(-\pi(t-t_0)^2/a^2\right)\exp(-2\pi jf_0(t-t_0)),$$

with corresponding Fourier transform,

$$H(f) = \frac{1}{a} \exp\left\{-\frac{\pi}{a^2}(f+f_0)^2\right\} \exp\{-j2\pi t_0 f\},$$

where t is the time-domain variable, f is the frequency-domain variable, t_0 is the relative time shift, a is the time/frequency spread variable, f_0 is the center frequency, and j is the complex operator.

Two popular forms of the Gabor function are the symmetric and anti-symmetric Gabor function. The use of symmetric and anti-symmetric Gabor functions becomes important when constructing multi-channel filter banks for image processing. The

symmetric and anti-symmetric 1-D Gabor functions are

$$h_s(t) = \exp\{-\pi(t-t_0)^2/a^2\} \cos(2\pi f_0(t-t_0))$$

and

$$h_A(t) = \exp\{-\pi(t-t_0)^2/a^2\} \sin(2\pi f_0(t-t_0))$$

with corresponding Fourier transform,

$$H_s(f) = \frac{\exp\{-jt_0f\}}{2} \left[\exp\{-\pi(f+f_0)^2/a^2\} + \exp\{-\pi(f-f_0)^2/a^2\} \right]$$

and

$$H_A(f) = \frac{j \exp(-jt_0f)}{2} \left[\exp\{-\pi(f+f_0)^2/a^2\} - \exp\{-\pi(f-f_0)^2/a^2\} \right],$$

where $h_s(t)$ and $h_A(t)$ are symmetric and anti-symmetric about the point t_0 respectively.

Daughman [5] was the first to introduce the family of 2-D Gabor functions, where it is defined as a 2-D Gaussian modulated by a 2-D sinusoid of given frequency and orientation. 2-D symmetric and anti-symmetric Gabor functions are important to the understanding of the early stages of the HVS and the processing, analysis and segmentation of 2-D textures. Specifically, the family of 2-D Gabor function is defined as

$$h(x, y) = \exp\left(-\pi\left[(x-x_0)^2/a^2 + (y-y_0)^2/b^2\right]\right) \exp\left(-2\pi j[u_0(x-x_0) + v_0(y-y_0)]\right)$$

and

$$H(u, v) = \exp\left(-\pi\left[(u-u_0)^2/a^2 + (v-v_0)^2/b^2\right]\right) \exp\left(-2\pi j[x_0(u-u_0) + y_0(v-v_0)]\right)$$

where x and y are spatial variables, u and v are spatial frequency variables, u_0 and v_0 are spatial-frequency components in the x and y directions respectively, and a and b are the

spatial/spatial-frequency spread variables in the x and y directions respectively. The 2-D symmetric and anti-symmetric Gabor functions are

$$h_s(x, y) = \exp\left\{-\pi\left[(x-x_0)^2 a^2 + (y-y_0)^2 b^2\right]\right\} \cos(2\pi[u_0(x-x_0) + v_0(y-y_0)])$$

and

$$h_A(x, y) = \exp\left\{-\pi\left[(x-x_0)^2 a^2 + (y-y_0)^2 b^2\right]\right\} \sin(2\pi[u_0(x-x_0) + v_0(y-y_0)]),$$

with corresponding Fourier transforms

$$H_s(u, v) = \frac{\exp\{-j2\pi(x_0 u + y_0 v)\}}{2a^2 b^2} \left[\begin{array}{l} \exp\left\{-\pi\left(\frac{(u+u_0)^2}{a^2} + \frac{(v+v_0)^2}{b^2}\right)\right\} \\ + \exp\left\{-\pi\left(\frac{(u-u_0)^2}{a^2} + \frac{(v-v_0)^2}{b^2}\right)\right\} \end{array} \right]$$

and

$$H_A(u, v) = \frac{j \exp\{-j2\pi(x_0 u + y_0 v)\}}{2a^2 b^2} \left[\begin{array}{l} \exp\left\{-\pi\left(\frac{(u+u_0)^2}{a^2} + \frac{(v+v_0)^2}{b^2}\right)\right\} \\ - \exp\left\{-\pi\left(\frac{(u-u_0)^2}{a^2} + \frac{(v-v_0)^2}{b^2}\right)\right\} \end{array} \right].$$

1.2.2 Gabor Filters

Gabor filters are a class of filters derived from the Gabor family of functions and are often used in image segmentation tasks due to their 1) time-frequency bandwidth optimality, 2) easily tunable nature, and 3) relation to the HVS. Normally, the symmetric and anti-symmetric forms of the filters are employed in a multi-channel configuration with each channel characterized by a Gabor filter with a certain center frequency, orientation, and bandwidth. A rotated Cartesian coordinate frame x' and y' are used in

the filter definition. This serves to rotate the Gaussian envelope by an angle θ_0 , such that it is orthogonal to the rotated coordinate frame. The symmetric and anti-symmetric Gabor filters are defined as

$$h_s(x, y) = \exp\left\{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right]\right\} \cos(2\pi[u_0x + v_0y]) \quad (1.1)$$

and

$$h_A(x, y) = \exp\left\{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right]\right\} \sin(2\pi[u_0x + v_0y]). \quad (1.2)$$

Further, let,

$$f_0 = \sqrt{u_0^2 + v_0^2}$$

and

$$\theta_0 = \tan^{-1}\left(\frac{v_0}{u_0}\right),$$

then $h_s(x, y)$ and $h_A(x, y)$ can be re-written as

$$h_s(x, y) = \exp\left\{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right]\right\} \cos(2\pi f_0 [\cos(\theta_0)x + \sin(\theta_0)y]) \quad (1.3)$$

and

$$h_A(x, y) = \exp\left\{-\frac{1}{2}\left[\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right]\right\} \sin(2\pi f_0 [\cos(\theta_0)x + \sin(\theta_0)y]) \quad (1.4)$$

where σ_x is the standard deviation in the x -direction, σ_y is the standard deviation in the y -direction, u_0 is the sinusoidal frequency component in the x -direction, v_0 is the sinusoidal

frequency component in the y-direction, f_0 is defined as the filter *center frequency*, θ_0 is defined as the filter *orientation*, and

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta_0) & \sin(\theta_0) \\ -\sin(\theta_0) & \cos(\theta_0) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

is the Cartesian coordinate frame obtained by rigid rotation of angle θ_0 . In the frequency domain, the filters are written as

$$H_S(u, v) = \frac{2\pi^2}{\sigma_u^2 \sigma_v^2} \begin{bmatrix} \exp\left\{-2\pi^2\left(\left(\frac{u'}{\sigma_u}\right)^2 + \left(\frac{v'}{\sigma_v}\right)^2\right)\right\} + \\ \exp\left\{-2\pi^2\left(\left(\frac{u'}{\sigma_u}\right)^2 + \left(\frac{v'}{\sigma_v}\right)^2\right)\right\} \end{bmatrix} \quad (1.5)$$

$$H_A(u, v) = \frac{2\pi^2 j}{\sigma_u^2 \sigma_v^2} \begin{bmatrix} \exp\left\{-2\pi^2\left(\left(\frac{u'}{\sigma_u}\right)^2 + \left(\frac{v'}{\sigma_v}\right)^2\right)\right\} - \\ \exp\left\{-2\pi^2\left(\left(\frac{u'}{\sigma_u}\right)^2 + \left(\frac{v'}{\sigma_v}\right)^2\right)\right\} \end{bmatrix} \quad (1.6)$$

where,

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \cos(\theta_0) & \sin(\theta_0) \\ -\sin(\theta_0) & \cos(\theta_0) \end{bmatrix} \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix}$$

is the rotated Cartesian coordinate frame, $\sigma_u = 1/\sigma_x$ is the standard deviation in the u -direction, $\sigma_v = 1/\sigma_y$ is the standard deviation in the v -direction. Thus, four parameters are needed to fully define a Gabor filter; σ_x , σ_y , u_0 , and v_0 or σ_x , σ_y , f_0 , and θ_0 .

1.2.3 Multit-Channel Filter Banks

2-D Multi-channel filtering has existed for some time now as a means of 2D texture analysis [1]. The technique was borrowed from studies of the early stages of the HVS. These studies reveal that the visual scene is first subjected to a multi-channel decomposition by the simple cortical cells of the striate cortex in the retina [6]. Higher-order brain functions then take the information provided by the channels and perform image processing to reconstruct and interpret the visual scene. Little is known about the higher-order functions; however, the structure and function of the low-level channels has been thoroughly studied and documented [7]. Daughman was the first to characterize the receptive field profiles (i.e. impulse response) of these channels in both spatial dimensions and the first to define the 2D Gabor signal [8]. According to the HVS studies, each channel is locally sensitive in both the spatial and spatial-frequency domains. In other words, each simple cell will only have a strong response over a given spatial area and for a limited band of spatial frequencies and orientations [9]. Most importantly, each simple cell exhibits a symmetric or anti-symmetric “Gabor-like” response to a given stimulus.

Multi-channel filter banks are extensively used in texture classification algorithms wherein the task is to distinguish one texture from another. According to [3] the following HVS concepts are useful in constructing HVS-based multi-channel algorithms to perform image segmentation and texture analysis.

- 1: The HVS processes the visual scene through a dense set of channels of varying center frequency, orientation, and bandwidth.
- 2: Simple cortical cells are primary mechanisms that characterize the channels. The

response of each cell is either symmetric or anti-symmetric and can be described as “Gabor like”.

- 3: Each channel performs an approximately linear summation of the channel input.
- 4: Each channel performs a non-linear half-wave rectification of the channel output.
- 5: Physically adjacent symmetric and anti-symmetric simple cells act in quadrature filter pairs tuned to the same center frequency, orientation and bandwidth. A symmetric simple cell will be paired with an anti-symmetric simple cell to provide information to the brain.
- 6: Because of the half-wave rectification, another pair of physically adjacent cells with opposite polarity preserves all information.

These properties are used in [3] to develop a texture segmentation algorithm based on the HVS and are the topic of Chapter 4 of this thesis. In [10] a bank of Gabor filters is used in conjunction with an automatic vector classification algorithm to perform unsupervised texture segmentation. This publication also introduces the use of the hyperbolic tangent function for post-processing. A similar algorithm is used to perform the segmentation in Chapters 3 and 5 of this thesis, and the nonlinear hyperbolic tangent function finds use in Chapters 3, 4, and 5. Other papers using Gabor filter banks to perform texture classification/segmentation are [11] [1] [12] [13] and [14]. These papers have a common link in that they filter an image through a bank of Gabor filters and then classify the filtered images with a vector classification algorithm such as k-means clustering, fuzzy k-means clustering, or neural network clustering methods.

The multi-channel filter bank approach to texture segmentation is used in a wide variety of practical applications in many publications. A few of these applications are

listed here. [15] uses this approach to detect defects in textured material such as cloth or wood for industrial applications. [16] uses the multi-channel approach to segment various areas of remote sensing imagery such as crops, woods, and urban areas. [17] uses the multi-channel approach in the automatic diagnosis of osteoporosis in X-ray images. The extraction of tissue deformation in tagged cardiac MRI using a bank of Gabor filters is discussed in [18], Also, tagged cardiac MRI tagging lines are automatically segmented in [24] using a bank of Gabor filters.

1.2.4 Gray Level Co-Occurrence Matrix Features

Gabor functions have been shown to be effective for extracting texture features; however, another method of texture feature extraction is that of the Gray Level Co-Occurrence Matrix (GLCM). Although GLCM is not the topic of this thesis, it is discussed here due to its popularity as a texture classification device. Hamlick [20] was the first to propose the GLCM method as a means of texture feature extraction. GLCM methods are actually estimates of the second-order statistics of a texture and have been shown to be effective at texture discrimination [21].

Tuceryan and Jain [21] best describe the gray-level co-occurrence matrix, P_d , as the number of joint occurrences of the scalar value $I(r,s) = i$ and $I(t,v) = j$ where I is the input image, $((r,s),(t,v))$ are image coordinates, and (i,j) are pixel values. An equivalent definition of P_d is the joint probability of $I(r,s) = i$ and $I(t,v) = j$.

Let $I(\mathbf{p})$ be an $N_r \times N_c$ input image where \mathbf{p} is defined to be the spatial point

$$\mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}.$$

Then let \mathbf{d} be a displacement vector

$$\mathbf{d} = \begin{bmatrix} dx \\ dy \end{bmatrix}.$$

Now, the set of possible gray levels, G , in the input image, I , is defined as the set of positive integers from 1 to N_g , where N_g is then number of possible gray levels.

Then, $G = \{1, 2, \dots, i, \dots, j, \dots, N_g\}$. Therefore, $P_{\mathbf{d}}$ is defined to be an $N_g \times N_g$ matrix. The $(i, j)^{\text{th}}$ entry into this $N_g \times N_g$ matrix is the number of joint occurrences of the gray level i and the gray level j which are the vector distance \mathbf{d} apart. Mathematically, Tuceryan and Jain define $P_{\mathbf{d}}$ as

$$P_{\mathbf{d}} = \left| \{(\mathbf{p}), (\mathbf{p} + \mathbf{d}) : I(\mathbf{p}) = i, I(\mathbf{p} + \mathbf{d}) = j\} \right|,$$

where,

$$(\mathbf{p}), (\mathbf{p} + \mathbf{d}) \in \{1, 2, \dots, N_r\} \times \{1, 2, \dots, N_c\},$$

$$(\mathbf{p} + \mathbf{d}) = \begin{bmatrix} r + dx \\ s + dy \end{bmatrix},$$

and $|\cdot|$ is the cardinality of the set.

Several GLCM features have been defined by Haralick. Normally, features are calculated on a normalized version of $P_{\mathbf{d}}$ such that the sum of the elements of $P_{\mathbf{d}}$ equals one. Features defined in this manner are consistent with the joint probability definition of $P_{\mathbf{d}}$. If $P_{\mathbf{d}}$ is the GLCM matrix defined previously, then the normalized version of $P_{\mathbf{d}}$ will be called P_{nd} and is defined as

$$P_{nd} = \frac{P_{\mathbf{d}}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_{\mathbf{d}}(i, j)}.$$

In practice, feature images are obtained by calculating a GLCM feature over a sliding window across the entire image. Some of the more common GLCM features are listed below.

$$\text{Energy} = E = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P_{nd}^2(i, j)$$

$$\text{Contrast} = C_o = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} |i - j|^2 P_{nd}(i, j)$$

$$\text{Correlation} = C_r = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{(i - u_i)(j - u_j) P_{nd}(i, j)}{\sigma_i \sigma_j}$$

$$\text{Homogeneity} = H = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{P_{nd}(i, j)}{1 + |i - j|}$$

GLCM features are used extensively in processing of Synthetic Aperture Radar (SAR) imagery; however, the method is not limited to such an application. Soh and Tsatsoulis used GLCM features to map sea ice in SAR images [22]. Liew, Lim, Kwoh, and Tay used GLCM features to distinguish between urbanized areas, forests, and swamps from SAR images [23]. Lopez, Moctezuma, and Parmiggiani use a combination of GLCM features and Markov Random Fields to detect oil spills from SAR images [24].

1.3 Background on Heart Anatomy and Physiology

1.3.1 Heart Anatomy

The heart is part of the circulatory system and is the organ responsible for pumping blood throughout the body. The heart has 4 chambers; left atrium (LA), left ventricle (LV), right atrium (RA), and right ventricle (RV). The two atria are located in

the upper half of the heart, and the two ventricles are located in the lower half. A thick muscular wall known as the septum separates the LA from the RA and the LV from the RV. The LV is the high-pressure region of the circulatory system. It is the largest chamber with the thickest walls and is considered the “work horse” of the heart. The shape of the LV is often compared to the lower half of an American football with the bottom of the football called the apex and the cut-off top of the football called the base. Automatic processing of magnetic resonance images of the LV is the focus of Chapters 2, 3, and 4 of this thesis. Figure 1.1 details the left ventricle, right ventricle, left atrium, and right atrium.

Four valves regulate blood flow through the four heart chambers: mitral valve, aortic valve, tricuspid valve, and pulmonary valve. The mitral valve regulates blood traveling from the left atrium to the left ventricle. The aortic valve regulates blood flowing from the left ventricle to the aorta. The tricuspid valve controls blood flow from the right atrium to the right ventricle. The pulmonary valve controls blood flowing from the right ventricle to the pulmonary arteries. Papillary muscles, along with the chordae tendineae act like anchors for each valve allowing a one-way flow of blood. A damaged papillary muscle or malfunctioning valve can cause partial blood flow in the reverse direction (i.e. from ventricle to atrium) in a condition known as mitral valve regurgitation. Figure 1.1 shows the location of the four valves discussed here.

Heart muscle, also called myocardium, has three primary layers: endocardium, midwall, and epicardium. The endocardium is the thin inner layer of the myocardium, and the epicardium is the thin outer layer of the myocardium. The midwall is the muscle between the endocardium and the epicardium.

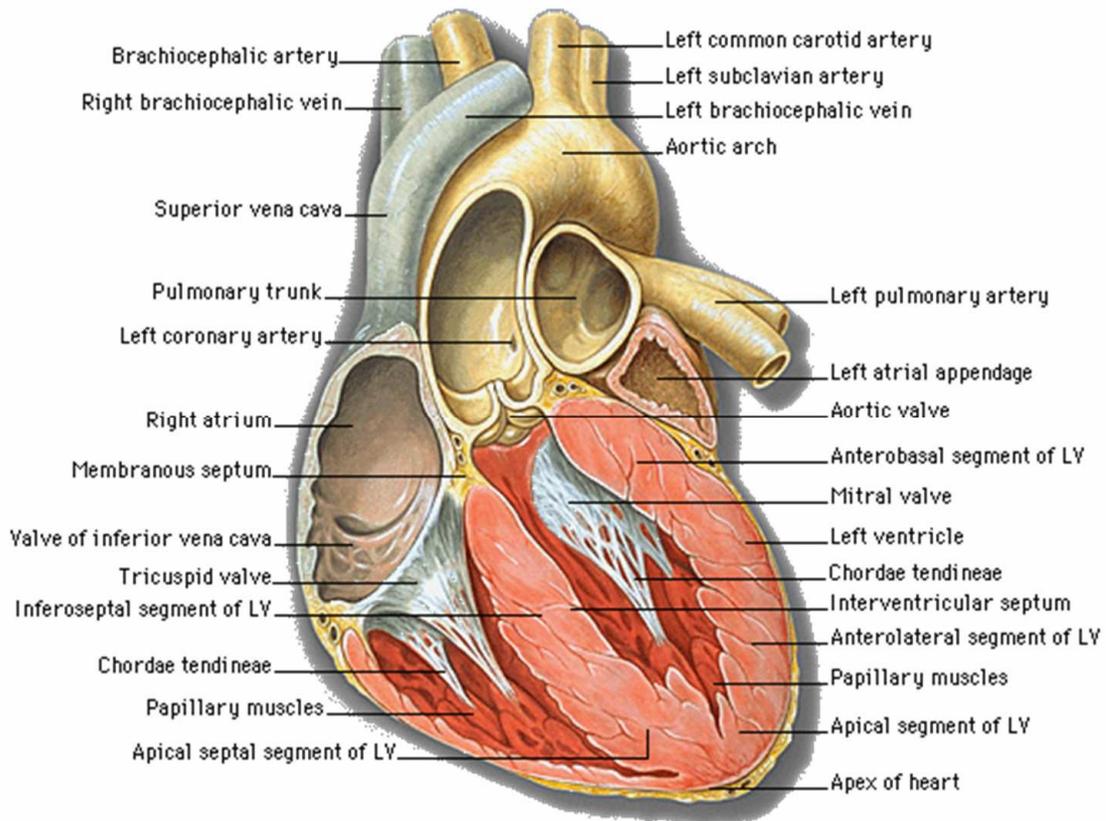


Figure 1.1: Heart anatomy

1.3.2 Heart Physiology

The function of the heart is to pump oxygen-rich blood to the many organs of the body and force oxygen-depleted blood to the lungs for a new supply of oxygen. The left chambers of the heart receives the oxygen-rich blood from the lungs and delivers it to the body through the circulatory system while the right chambers deliver oxygen-depleted blood to the lungs for re-oxygenation. The left and right ventricles act as the primary pumping mechanisms in this process with the LV the stronger of the two. Left and right atria act as blood delivery devices to the LV and RV respectively.

According to the Texas Heart Institute, the average human heart beats 100,000 times a day to deliver 2,000 gallons of life-giving, oxygen-rich blood. The heart must

keep up this relentless pace throughout the person's lifetime. The periodic contractions of the heart can be divided into a two-part cycle called the systolic cycle. Half the cycle is called systole and the other half is called diastole. Systole is the part of the cycle when the left and right ventricles contract to deliver blood to the body and lungs respectively. This is also the part of the cycle when the left and right atria fill with blood ready to be delivered to the left and right ventricles during diastole. Diastole is the part of the cycle when the left and right ventricle are expanding and filling with blood from the contracting left and right atria. The point in the systolic cycle when the ventricles are fully contracted is called end-systole. The point in the systolic cycle when the ventricles are fully dilated is called end-diastole.

Figure 1.2 shows the pulmonary circuit representation of the systolic cycle. Here, oxygen-rich blood from the lungs enters the right atrium and is delivered to the left ventricle. The left ventricle then pumps the oxygenated blood to the body where it is filtered to become oxygen-poor blood. Oxygen-poor blood is returned to the right atrium. The right atrium delivers the oxygen-poor blood to the right ventricle which then pumps the oxygen-depleted blood to the lungs for re-oxygenation and the cycle starts over.

1.4 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) has become an invaluable tool to medical professionals, providing high resolution images of the body. MRI is a non-invasive imaging modality which exploits a phenomenon known as nuclear magnetic resonance (NMR). According NMR theory, atoms with odd atomic number or mass are said to

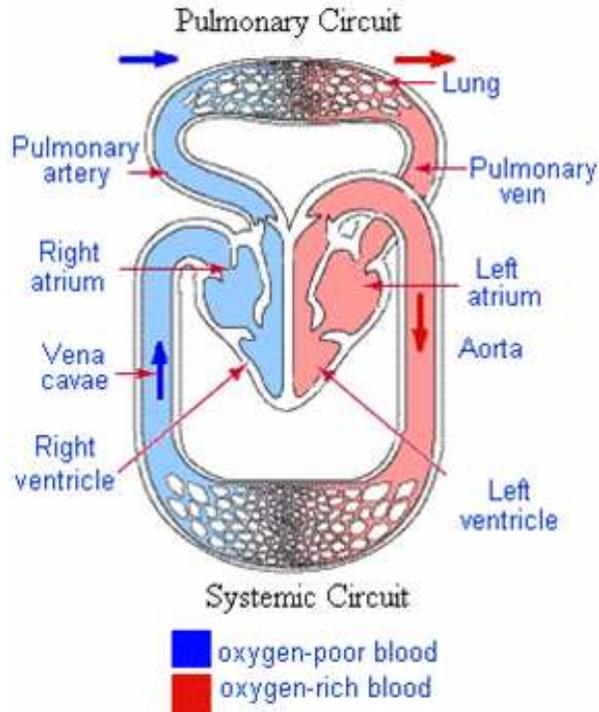


Figure 1.2: Pulmonary circuit representation of the circulatory system

possess spin. Of particular importance are hydrogen atoms because they are most prevalent in the body and are, in fact, what standard MRI measures. Each hydrogen atom can be visualized as a ball spinning on an axis. The spinning action of the atom, in turn, causes a magnetic moment vector. A collection of such spinning atoms are called a nuclear spin system. Normally, the magnetic moments of a spin system are randomly oriented. However, when exposed to an external magnetic field \mathbf{B}_0 the individual magnetic moments of the spin system align in a direction parallel to \mathbf{B}_0 creating a bulk or macroscopic magnetization vector \mathbf{M} .

By applying a second external magnetic field, \mathbf{B}_1 , perpendicular to \mathbf{B}_0 , \mathbf{M} can be tipped away from \mathbf{B}_0 . When \mathbf{B}_1 is removed, \mathbf{M} begins to precess about \mathbf{B}_0 at the Larmor frequency until it returns to its equilibrium state. The precession of \mathbf{M} , in turn, creates an

RF signal measurable by the MRI scanner. With the use of gradient coils and various pulse sequences, the MRI scanner systematically scans the body to form an image.

1.4.1 Cardiac MRI

Cardiac MRI is the use of MRI to image slices of the heart. Although MRI can image in any 3-D plane or slice, several standardized segment views have been defined by the American Heart Association for imaging of the left ventricle [2]. The views can be combined to extract 3-D information of the heart from several 2-D slices. The views are defined as short axis and long axis. Long axis views are further defined as either 4-chamber or 2-chamber views. Figure 1.3 shows a pictorial representation of the different views.

1.4.2 Cine MRI

Cine MRI is the use of MRI to produce movie-like images. This technology has recently become important in cardiac MRI. Usually, the electrocardiogram of a patient is used to synchronize image acquisition to the cardiac cycle. By imaging the heart over a period of time, physicians are able to non-invasively examine the heart over the entire systolic cycle and derive quantitative data useful in the diagnosis of certain heart diseases. Figure 1.4 shows a pair of short axis view cardiac cine images taken at end-diastole and end-systole.

REGIONAL WALL SEGMENTS

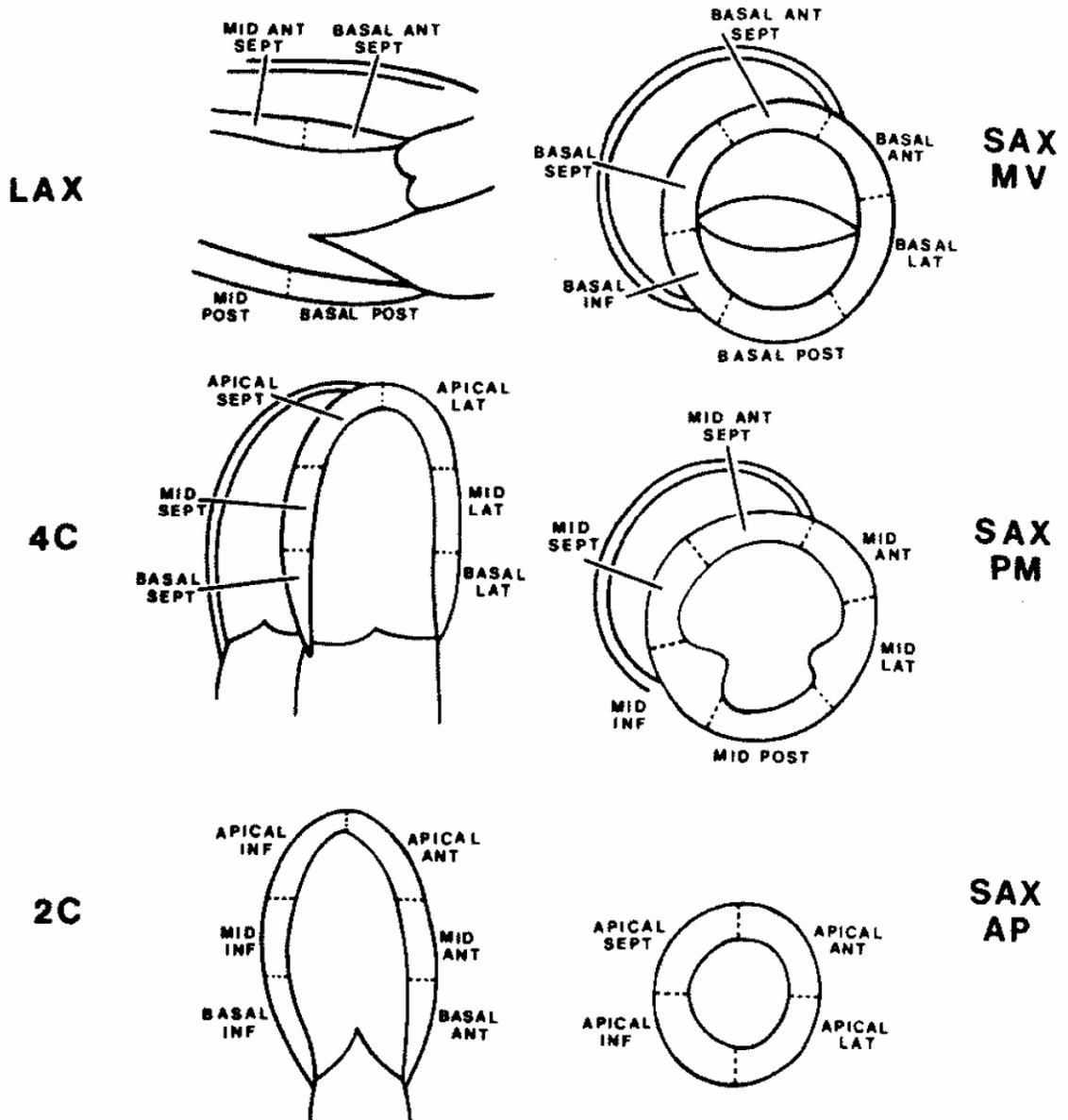


Figure 1.3: Pictorial representation of the standardized imaging views as defined by the American Heart Association

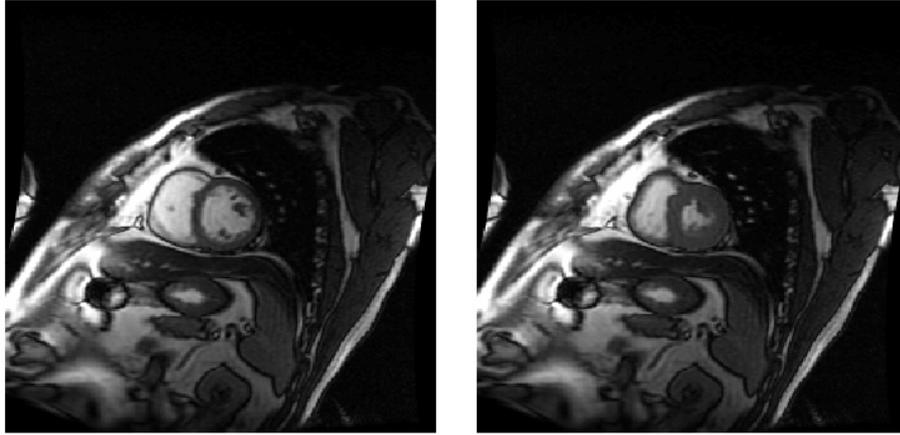


Figure 1.4: Short axis cine cardiac MRI images. End-diastole shown left and end-systole shown right

1.4.3 Tagged MRI

Tagged cardiac MRI provides a non-invasive means to perform a three-dimensional strain analysis of the heart in vivo. The modality involves placing parallel planes of decreased signal intensity in the body to be examined. The planes are called tags. Tags will deform with the tissue on which they are placed to reveal the tissue deformation. An image plane is viewed perpendicular to the tags; therefore, the tags appear as dark stripes in the image. This can be likened to drawing a black line on an inflated balloon and then stretching the balloon to examine the deformation of the balloon surface. A necessary step in three-dimensional strain reconstruction is to “track” the movement of the tags from end-diastole to end-systole in a process called tag tracking. Figure 1.5 shows a short axis view of a tagged cardiac MRI image at end-systole and end-diastole.

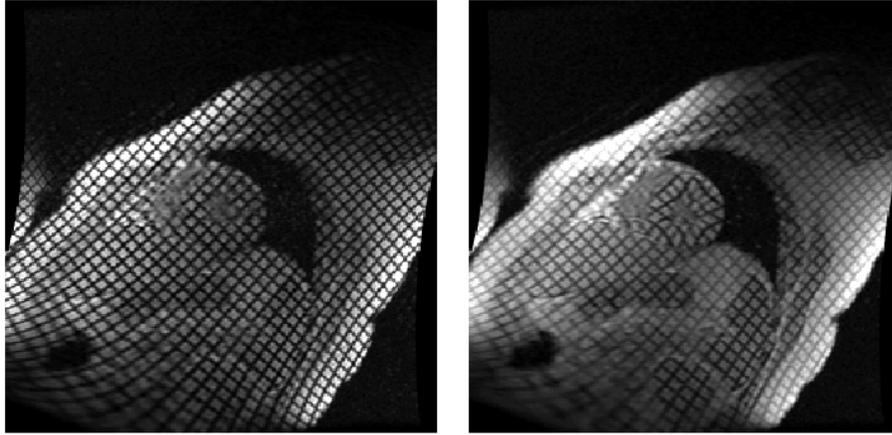


Figure 1.5: Short axis tagged cardiac MRI images. End-diastole shown left and end-systole shown right

1.5 Overview of Thesis

The topic of this thesis is the use of multi-channel filters in the filtering and segmentation of images. Specifically, each channel is characterized by a Gabor filter, and the algorithms tested here are drawn heavily from the literature previously discussed. In Chapter 2 of this thesis, several models are developed for analysis of spectral changes due to myocardial deformation of tagged cardiac MRI images. From these models, a new tag-tracking pre-filter is designed. The design has its roots in the multi-channel paradigm. The filter is tested against two other commonly used tag tracking pre-filters and is shown to be superior in performance.

The use of multi-channel filters to automatically identify tag lines in tagged cardiac MRI in Chapter 4. A method of tag line identification based on [19] is examined in Chapter 3.2. The method is shown to be effective; however, sub-pixel refinement of identified tag pixels is not explored here. The filter developed in Chapter 2 is used as a pre-filter in Chapter 3.3. A peak detection is then performed using the pre-filtered image

to identify candidate tag pixels. The tag pixels are then refined to sub-pixel locations using quadratic interpolation.

The use of multi-channel filter banks to perform a supervised segmentation of the LV myocardium in cine cardiac MRI images is explored in Chapter 4. The methodology of an algorithm developed for this purpose is detailed in Chapter 4.2. The algorithm involves pre-processing the images with a local contrast enhancement method to alleviate the intensity inhomogeneity that is prevalent in MRI images. The workhorse of the method presented in this chapter is the Gabor-based multi-channel filter bank used to create feature images. An experiment to determine the optimal filter bank parameters and subsequently gives the results of the experiment is given in Chapter 4.3.

An automatic multi-channel segmentation to segment aggregate in asphalt core images is detailed in Chapter 5. The purpose of the automatic segmentation is to perform an aggregate orientation analysis of each asphalt core sample. A complete description of the experimental procedure is given, including asphalt core sample acquisition, image processing, and aggregate orientation analysis. The chapter is ended with some results from the asphalt study.

Last, a summary of Chapters 1 through 5 is provided in Chapter 6. This includes a brief description and what was learned from the work performed in each chapter. Chapter 6 is concluded with some suggestions for future work.

CHAPTER 2

THE C-GABOR FILTER

A tag line identification algorithm which involves pre-filtering an end-systolic image and performing peak detection to highlight candidate tag pixels is presented in Chapter 3. Tag tracking is achieved by statistical classification of candidate tag pixels [25]. Analytical models used to motivate the design of three pre-filters and quantitatively evaluate their relative performance by testing them in conjunction with the tag line identification algorithm of Chapter 3 are presented here in Chapter 2.

In Section 2.1, the frequency-domain effects of tag line deformation are examined through five analytical models. A new filter based on an analysis of the models is derived in Chapter 2.2. The filter has its roots in the multi-channel paradigm and is shown to be superior to commonly used filters. The performance of the new filter of Chapter 2.2 is tested against two commonly used filters in Chapter 2.3. Last, a summary and conclusion for Chapter 2 is provided in Chapter 2.4.

2.1 Tag Deformation Analysis

Local myocardial deformations can be modeled as a combination of shear, rotation, and change in tag frequency. Five tag deformation models are now presented to study the frequency-domain effects of a local tag line deformation. The models range in

complexity and accuracy. One should note, however, that the models are not intended to provide an accurate simulation as in [26], but rather, to provide tractability to filter design. The five models are: 1) Shear Model, 2) Rotation Model, 3) Contraction/Expansion Model, 4) Jones Model, and 5) Sum of Shifted Gaussians Model.

The first three models apply a global deformation to hypothesize possible frequency-domain effects of a local deformation. The last two models apply a local deformation, adding validity to the three global deformation models. For each model, the tag pattern is modeled as

$$\psi(\mathbf{p}) = 1 + \cos(\mathbf{w} \cdot \mathbf{p}) \stackrel{\mathcal{S}_{2D}}{\Leftrightarrow} \psi(\mathbf{k}) = \delta(\mathbf{k}) + \frac{1}{2}(\delta(\mathbf{k} - \mathbf{w}) + \delta(\mathbf{k} + \mathbf{w})),$$

where

$$\mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}$$

is a 2D spatial point,

$$\mathbf{k} = \begin{bmatrix} u \\ v \end{bmatrix}$$

is a 2D frequency point

$$\mathbf{w} = 2\pi \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}$$

is the 2D tag frequency, and

$$\stackrel{\mathcal{S}_{2D}}{\Leftrightarrow}$$

represents the 2D Fourier Transform.

The untagged, undeformed image is simulated as an annulus, $f(\mathbf{p}) \Leftrightarrow F(\mathbf{k})$, of inner radius r_1 and outer radius of r_2 . The tags are applied to the myocardium by multiplying $\psi(\mathbf{p})$ by $f(\mathbf{p})$. The complete model for the tagged, undeformed image is given as

$$f_{tag}(\mathbf{p}) = f(\mathbf{p})\psi(\mathbf{p}) \Leftrightarrow F_{tag}(\mathbf{k}) = F(\mathbf{k}) ** \psi(\mathbf{k}) = F(\mathbf{k}) + \frac{1}{2}(F(\mathbf{k} - \mathbf{w}') + F(\mathbf{k} + \mathbf{w}'))$$

where $**$ is the 2D convolution operation.

Tag lines are not modeled exactly as sinusoids. Most tag lines are, in fact, much thinner than a true sinusoid and have a Gaussian-like shape near the intensity minima; however, tag lines with higher frequency components can be modeled with a Fourier series expansion of the tag pattern. One way to view the tag pattern is as a local modulator. If the function $\varphi(\mathbf{p})$ is thought of as a modulator of frequency \mathbf{w} , then the frequency domain effect of tagging the myocardium, $f(\mathbf{p})$, with the tag pattern $\varphi(\mathbf{p})$ is to create scaled copies of the spectrum of $F(\mathbf{k})$ at frequency $\pm\mathbf{w}$. Any local change in $\varphi(\mathbf{p})$ should cause a corresponding change in spectral energy. Figure 2.1 shows the untagged myocardium model $f(\mathbf{p})$, and the tags $\varphi(\mathbf{p})$ at the top-left and top-right respectively. Figure 2.1 also shows the complete model, $f_{tag}(\mathbf{p})$, in both the spatial and frequency domains at the bottom-left and bottom-right respectively.

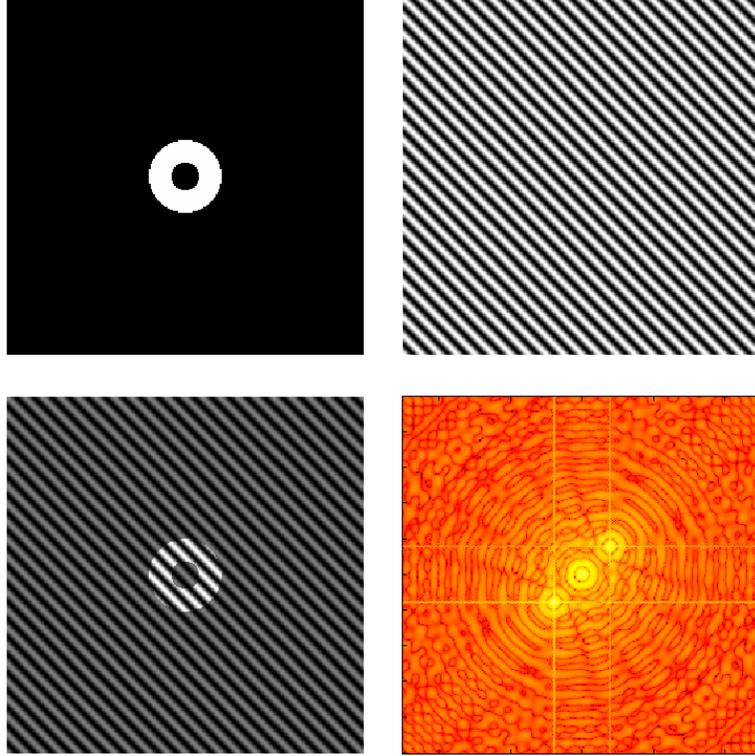


Figure 2.1: Myocardium model (top-left); tags(top-right); tagged myocardium model (bottom-left); and log frequency spectrum (bottom-right)

2.1.1 Shear Model

The frequency domain effects of a global simple shear are derived in this section.

Let

$$f_s(\mathbf{p}) = f_{tag}(\mathbf{Sp}) = f(\mathbf{Sp})\psi(\mathbf{Sp})$$

be a simple shear of the tagged image in the x -direction, where

$$\mathbf{S} = \begin{bmatrix} 1 & b \\ 0 & 1 \end{bmatrix}$$

is a simple shear matrix and b is called the shear constant and the transformation \mathbf{Sp} is applied to every point \mathbf{p} in the image. Then, by the Simple Shear Theorem of the 2D

Fourier Transform [27], the deformed tag line spectrum is given as

$$F_s(\mathbf{k}) = F(\mathbf{S}_s \mathbf{k}) + \frac{1}{2} [F(\mathbf{S}_s \mathbf{k} - \mathbf{w}) + F(\mathbf{S}_s \mathbf{k} + \mathbf{w})],$$

where

$$\mathbf{S}_s = \begin{bmatrix} 1 & 0 \\ -b & 1 \end{bmatrix}.$$

\mathbf{S}_s is also a simple shear matrix wherein the shear is perpendicular to the spatial shear \mathbf{S} .

Hence, a global simple shear of the tagged image causes a global simple shear of its spectral energy perpendicular to the spatial shear. One should note that analogous results are obtained if the simple shear is applied in the y -direction. Also, a compound shear can be viewed as a combination of simple shears. From the analysis above, one can deduce that a local myocardial shear will shear a portion of the spectral energy in a direction perpendicular to the spatial shear, thereby causing a partial spreading of spectral energy. Figure 2.2 shows a globally sheared myocardium and resulting frequency spectrum with a shear constant of $b = 0.5$ applied to the tagged myocardium model of Figure 2.1. Note the arrows showing the direction of spatial and spectral shear. As predicted, Figure 2.1 shows that a global simple spatial shear has caused a global simple spectral shear perpendicular to the global simple spatial shear.

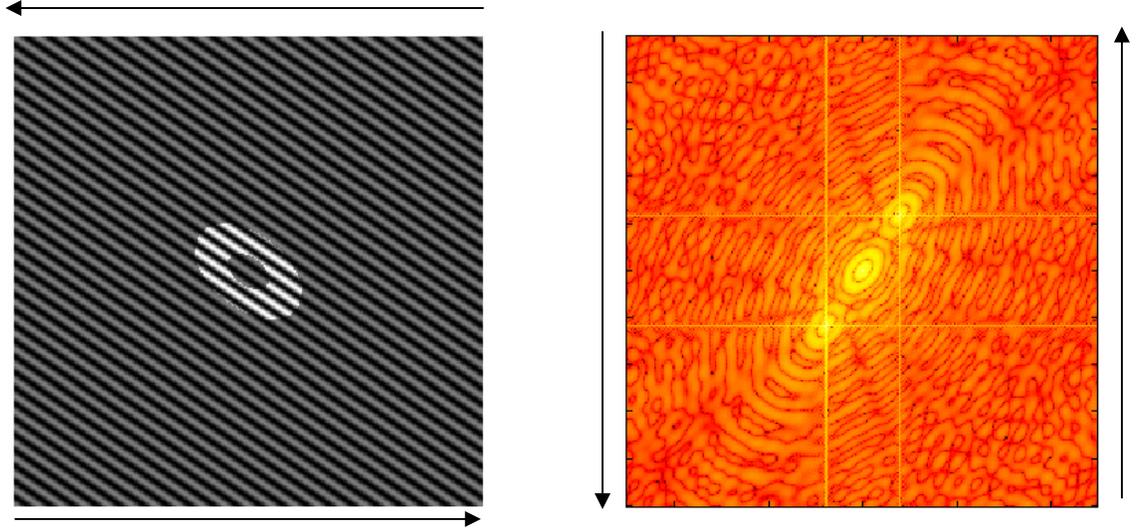


Figure 2.2: Sheared myocardium model (left) and resulting log frequency spectrum (right)

2.1.2 Rotation Model

The frequency-domain effects of a global rotation are derived in this section.

Let

$$f_R(\mathbf{p}) = f_{tag}(\mathbf{Rp}) = f(\mathbf{Rp})\psi(\mathbf{Rp})$$

be a global rotation of the tagged image, where

$$\mathbf{R} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

is the rotation matrix and θ is the angle of rotation and the transformation \mathbf{Rp} is applied to every point \mathbf{p} in the image. Then, by the Rotation Theorem of the 2D Fourier Transform [27], the rotated image spectrum is given as

$$F_R(\mathbf{k}) = F(\mathbf{Rk}) + \frac{1}{2} [F(\mathbf{Rk} - \mathbf{w}) + F(\mathbf{Rk} + \mathbf{w})].$$

Hence, a global tagged image rotation causes a corresponding global rotation of the spectral energy. From the analysis above, one can deduce that a local myocardial rotation will rotate a portion of the spectral energy in the same direction of rotation as the

spatial rotation, thereby causing a partial spreading of spectral energy. Figure 2.3 shows a globally rotated myocardium and resulting frequency spectrum with $\theta = 20$ degrees applied to the tagged myocardium model of Figure 2.1. Notice that the global rotation of spectral energy is in agreement with the spatial rotation.

2.1.3 Contraction/Expansion Model

Local changes in tag spacing are observed when the myocardium expands or contracts. Here, the frequency domain effects of global contraction or expansion are examined. Let

$$f_A(\mathbf{p}) = f(\mathbf{A}\mathbf{p})\psi(\mathbf{A}\mathbf{p})$$

where \mathbf{A} is the 2x2 contraction/expansion matrix

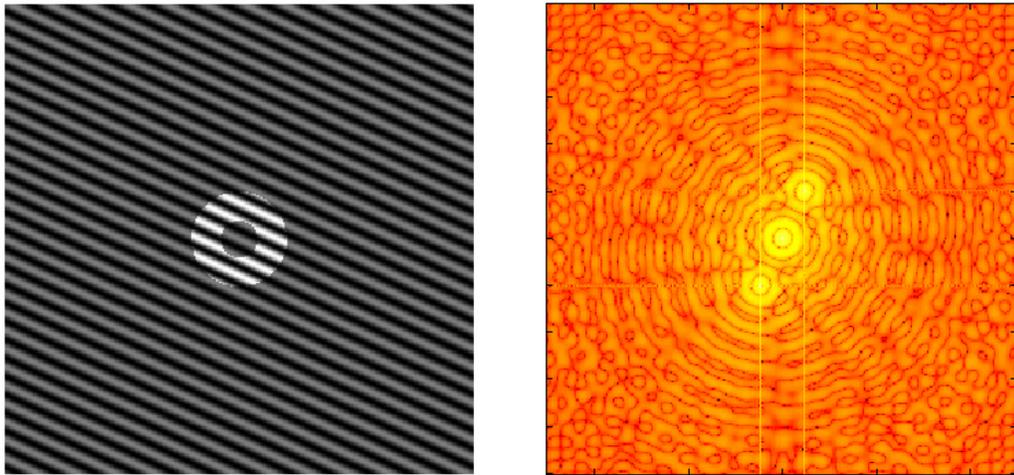


Figure 2.3: Rotated myocardium model (left) and resulting log frequency spectrum (right)

$$\mathbf{A} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

and the transformation $\mathbf{A}\mathbf{p}$ is applied to every point \mathbf{p} in the image.

Then the frequency spectrum of the image with a global contraction/expansion is given as

$$F_A(\mathbf{k}) = \frac{1}{ab} \left[F(\mathbf{A}^{-1}\mathbf{k}) + \frac{1}{2} \left[F(\mathbf{A}^{-1}\mathbf{k} - \mathbf{w}) + F(\mathbf{A}^{-1}\mathbf{k} + \mathbf{w}) \right] \right].$$

Hence, a global contraction or expansion in the spatial domain will cause a global expansion or contraction in the frequency-domain respectively. From the analysis above, one can deduce that a local contraction or expansion will cause a portion of the spectral energy to be expanded or contracted in the frequency-domain. Figure 2.4 shows the frequency domain effects of a global contraction/expansion. Note that the global contraction or expansion in the spatial-domain has caused a corresponding expansion or contraction in the frequency-domain

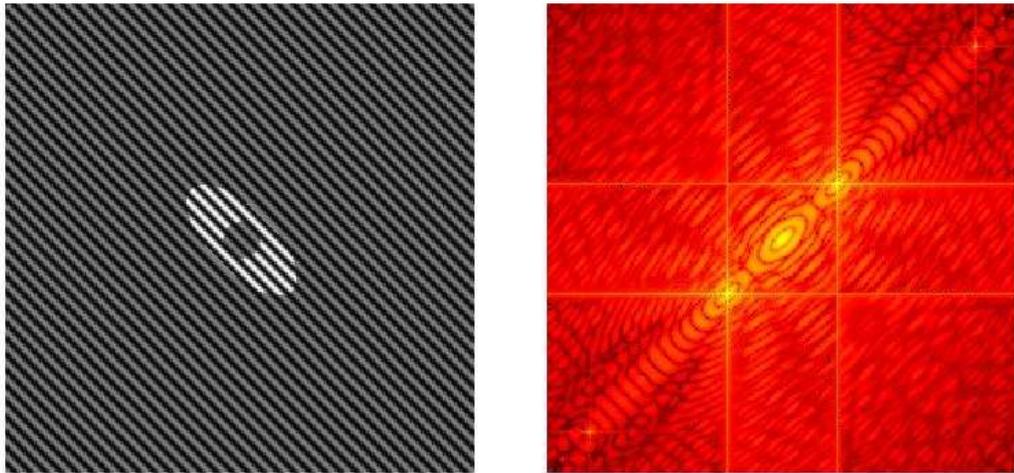


Figure 2.4: Myocardium with decreased tag spacing (left) and resulting log frequency spectrum (right)

2.1.4 Jones Deformation Model

Here, the Jones deformation model is applied to a synthetic tagged image to examine the effects of tag deformation on the frequency spectrum and to add validity to the global analyses above [28]. This idealized model of heart deformation simulates

contraction of the myocardium as an incompressible fluid moving inward toward a central sink with a velocity of $-1/\text{radius}^2$. Let

$$f_{tag}(\mathbf{p}, t) = f(\mathbf{p}, t)\psi(\mathbf{p}, t)$$

be the contracting tagged image model as a function of material point, \mathbf{p} , and time, t .

Now, re-writing $f_{tag}(\mathbf{p}, t)$ in polar coordinates as

$$f_{tag}(\mathbf{r}, t) = f(\mathbf{r}, t)\psi(\mathbf{r}, t)$$

where

$$\mathbf{r} = \begin{bmatrix} r \\ \theta \end{bmatrix}, \quad r = \sqrt{x^2 + y^2}, \quad \text{and} \quad \theta = \tan^{-1}\left(\frac{y}{x}\right)$$

the change in position with respect to time is given by the Jones model as

$$v(r, t) = \frac{d}{dt}(f_{tag}(\mathbf{r}, t)) = -\frac{\alpha}{r^2}$$

where α is a constant controlling the rate of myocardial contraction. Since, by the Jones model definition, a material point of $f_{tag}(\mathbf{r}, t)$ is invariant with respect to θ , $f_{tag}(\mathbf{r}, t)$ is

re-written as $f_{tag}(r, t)$. Finally, the equation describing the deformation is given as

$$f_{tag}(r, t + \Delta t) = f_{tag}(r, t) - \frac{\alpha}{r^2} \Delta t,$$

where Δt is the time step.

A simulation of the Jones deformation model is performed in MATLAB with the built-in ODE45 function. The spectrum of the deformed tagged image is examined for changes in spectral distribution. Figure 2.5 shows the original tagged image model (left) and the image deformed under the Jones deformation model (right). Notice the local shear, rotation and changes in tag spacing from the original to the deformed model in

Figure 2.5 (right). Figure 2.6 shows a log plot of the original (left) and deformed (middle) spectrum followed by a log plot of the absolute difference between the two (right). Notice the spreading of spectral peaks from the original to the deformed spectrum. The spectral difference image of Figure 2.6 (right) markedly highlights the radial spreading of the tag line spectrum due to deformation. Also, both Figures 2.5 and 2.6 shows that tag deformation can be thought of as a combination of local shear, rotation, and changes in tag frequency and adds credibility to the global analyses previously performed.



Figure 2.5: Original (left) and deformed (right) Jones model

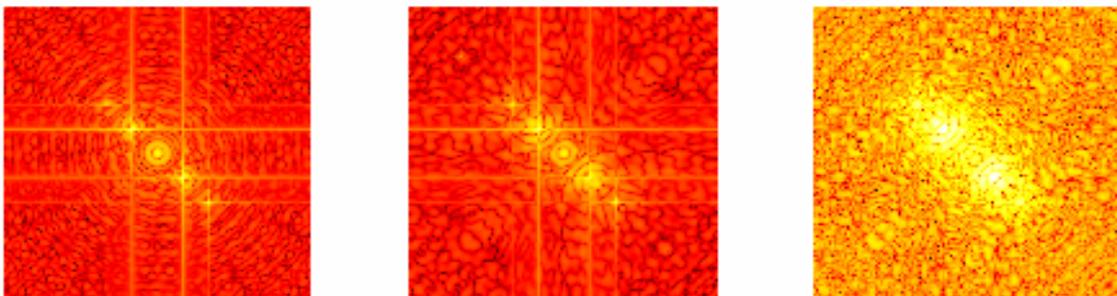


Figure 2.6: Log magnitude plot of original (left); deformed (middle); and absolute difference frequency spectrum (right)

2.1.5 Sum of Shifted Gaussians Model

A model is derived for 2D deformation in the tagged image spectrum due to cardiac deformation in this section. Here the left ventricle is modeled as a sum of shifted Gaussian functions whose local deformation is modeled as affine.

The untagged, undeformed image intensity f is modeled as a sum of stationary and deformable tissues. The stationary tissue is denoted as b , and the deformable tissue is modeled as a sum of Gaussian functions

$$f(\mathbf{p}) = b(\mathbf{p}) + \alpha \sum_{i=1}^n g(\mathbf{p} - \mathbf{p}_i),$$

where \mathbf{p} is a material point (fixed in the tissue), α is an amplitude, g is a Gaussian function with variance σ_m^2 , and \mathbf{p}_i is the center of the i -th Gaussian function. Gaussian functions are used because they were shown by Gabor (in the 1D case) and Daughman (in the 2D case) to be optimally localized in both the spatial and spatial frequency domains [4] [5]. The exact value of variance σ_m^2 is not critical because the purpose of this model is intended to motivate the frequency-domain design of a tag-tracking pre-filter. Modeling the myocardium with more Gaussian functions with smaller variance will increase the accuracy of locally affine deformation described above; however, the frequency-domain effects of the deformation will be spread over a larger area.

The tag pattern is modeled as

$$\psi(\mathbf{p}) = 1 + \cos(\mathbf{w} \cdot \mathbf{p}).$$

The tagged, undeformed image is the untagged image multiplied by the tag pattern

$$f_{tag}(\mathbf{p}) = \psi(\mathbf{p})f(\mathbf{p}) = \psi(\mathbf{p})b(\mathbf{p}) + \alpha \sum_{i=1}^n \psi(\mathbf{p})g(\mathbf{p} - \mathbf{p}_i).$$

The deformed tagged image at time t is obtained by letting $\mathbf{p}=\mathbf{p}(\mathbf{r},t)$ where \mathbf{r} is a spatial point (fixed in space) and $\mathbf{p}(\mathbf{r},t)$ is the reference map describing the motion [29], which yields

$$f_{tag}(\mathbf{r}) = f(\mathbf{p}(\mathbf{r},t)) \approx \psi(\mathbf{r})b(\mathbf{r}) + \alpha \sum_{i=1}^n \psi(\mathbf{A}_i\mathbf{r} + \mathbf{d}_i)g(\mathbf{A}_i\mathbf{r} + \mathbf{d}_i - \mathbf{p}_i)$$

Note that in the stationary tissue, $\mathbf{p} = \mathbf{r}$. In the myocardium, the deformation is approximated as locally affine

$$\mathbf{p} = \mathbf{A}_i\mathbf{r} + \mathbf{d}_i ,$$

where \mathbf{A}_i and \mathbf{d}_i define the affine deformation of the i -th Gaussian function. Since $b(\mathbf{r})$ is stationary, we can ignore the frequency domain changes due to this term. The 2D continuous Fourier transform of the second term is given by

$$\begin{aligned} & \mathfrak{F} \left\{ \alpha \sum_{i=1}^n \psi(\mathbf{A}_i\mathbf{r} + \mathbf{d}_i) g(\mathbf{A}_i\mathbf{r} + \mathbf{d}_i - \mathbf{p}_i) \right\} \quad (2.1) \\ & = \alpha \sum_{i=1}^n \frac{1}{|\mathbf{A}_i|} \exp\{j(\mathbf{d}_i\mathbf{A}_i^{-T}\mathbf{k})\mathbf{p}_i\} \times \left(\left[\begin{array}{l} \frac{1}{2} \exp\{-j(\mathbf{k} - \mathbf{w})\mathbf{p}_i\} G(\mathbf{k} - \mathbf{w}) \\ + \exp\{-j\mathbf{k}\mathbf{p}_i\} G(\mathbf{k}) \\ + \frac{1}{2} \exp\{-j(\mathbf{k} + \mathbf{w})\mathbf{p}_i\} G(\mathbf{k} + \mathbf{w}) \end{array} \right]_{\mathbf{k}=\mathbf{A}_i^{-T}\mathbf{k}} \right) \end{aligned}$$

where \mathbf{k} is a point in the 2D frequency domain, G is a Gaussian function with variance $1/\sigma_m^2$. Simplifying and lumping phase terms into the functions $\varphi_-(\mathbf{k})$, $\varphi_0(\mathbf{k})$, and $\varphi_+(\mathbf{k})$ yields

$$\begin{aligned} & \mathfrak{F} \left\{ \alpha \sum_{i=1}^n \psi(\mathbf{A}_i\mathbf{r} + \mathbf{d}_i) g(\mathbf{A}_i\mathbf{r} + \mathbf{d}_i - \mathbf{p}_i) \right\} \quad (2.2) \\ & = \alpha \sum_{i=1}^n \frac{1}{|\mathbf{A}_i|} \times \left(\left[\begin{array}{l} \varphi_-(\mathbf{k})G(\mathbf{A}_i^{-T}\mathbf{k} - \mathbf{w}) \\ + \varphi_0(\mathbf{k})G(\mathbf{A}_i^{-T}\mathbf{k}) \\ + \varphi_+(\mathbf{k})G(\mathbf{A}_i^{-T}\mathbf{k} + \mathbf{w}) \end{array} \right]_{\mathbf{k}=\mathbf{A}_i^{-T}\mathbf{k}} \right) \end{aligned}$$

Since \mathbf{A}_i is an affine deformation matrix, it is non-singular, and can be expressed using the polar decomposition as

$$\mathbf{A}_i = \mathbf{R}_i \mathbf{U}_i$$

where \mathbf{R}_i is a rotation matrix and \mathbf{U}_i is a symmetric, positive definite stretch tensor.

Equation (2.2) becomes

$$\begin{aligned} & \Im \left\{ \alpha \sum_{i=1}^n \psi(\mathbf{R}_i \mathbf{U}_i \mathbf{r} + \mathbf{d}_i) g(\mathbf{R}_i \mathbf{U}_i \mathbf{r} + \mathbf{d}_i - \mathbf{p}_i) \right\} & (2.3) \\ & = \alpha \sum_{i=1}^n \frac{1}{|\mathbf{A}_i|} \times \left(\left[\begin{array}{c} \varphi_-(\mathbf{k})G(\mathbf{R}_i \mathbf{U}_i^{-1} \mathbf{k} - \mathbf{w}) \\ + \varphi_0(\mathbf{k})G(\mathbf{R}_i \mathbf{U}_i^{-1} \mathbf{k}) \\ + \varphi_+(\mathbf{k})G(\mathbf{R}_i \mathbf{U}_i^{-1} \mathbf{k} + \mathbf{w}) \end{array} \right]_{\mathbf{k}=\mathbf{A}_i^{-T} \mathbf{k}} \right) \end{aligned}$$

Equation (2.3) describes the spectral changes in tagged images due to deformation.

The spectral changes due to deformation are shown in Figure 2.7. The dashed-line circles depict the undeformed spectrum. Regions of the image experiencing local compression and shear will broaden the peaks in the deformed-tagged spectrum, which are depicted by the ellipses in Figure 2.7. Regions of the image experiencing expansion will compress the peaks, but these changes will not affect the filter design. Regions of the image experiencing local rotation will produce spectral changes in regions rotated by the same angle about the center of the frequency domain, as depicted by the rotated ellipses.

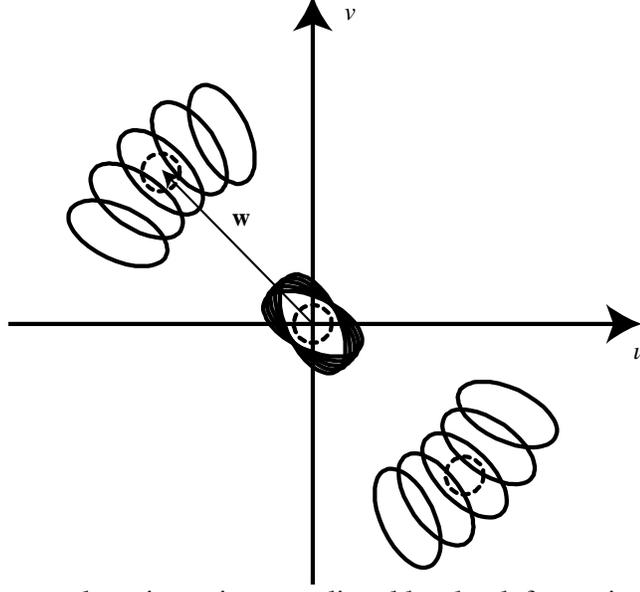


Figure 2.7: Frequency domain regions predicted by the deformation model to contain spectral changes due to local deformation of the myocardium

2.2 C-Gabor Filter Derivation

A tag-tracking pre-filter must capture as much of the expected deformed tag line spectrum as possible while minimizing the coverage of other spectral peaks. Since the exact deformation of the myocardium is unknown, the pre-filter should cover the spectral regions expected to contain the deformed spectrum. The spectral locations depicted in Figure 2.7 suggest a filter structure that covers the two banana-shaped regions defined by the rotated ellipsoids. Hence we propose a new filter which can be obtained with a composite Gabor (C-Gabor) filter that is the sum of several Gabor filters tuned to different rotation angles. If a standard Gabor filter is given by

$$H(\mathbf{k}) = \frac{1}{4\pi\sigma_x\sigma_y} \left[\exp\left\{-\frac{1}{2}(\mathbf{k} - \mathbf{w}_T)^T \mathbf{B}^{-1} (\mathbf{k} - \mathbf{w}_T)^T\right\} + \exp\left\{-\frac{1}{2}(\mathbf{k} + \mathbf{w}_T)^T \mathbf{B}^{-1} (\mathbf{k} + \mathbf{w}_T)^T\right\} \right]$$

where

$$\mathbf{B} = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix},$$

$$\mathbf{w}_T = \begin{bmatrix} (1/T) \cos \theta_T \\ (1/T) \sin \theta_T \end{bmatrix},$$

σ_x^2 and σ_y^2 are frequency-domain variances, T is the tag spacing, and θ_T is the tag angle; then the C-Gabor filter is given by

$$H_B(\mathbf{k}) = \frac{1}{2n+1} \sum_{m=-n}^n H(R(\theta)\mathbf{k}),$$

where

$$\theta = \frac{m\Delta\theta}{n},$$

$$R(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

and $\Delta\theta$ is the expected range of rotation angles. In this thesis, parameters are set to

$$n = \lceil (2\Delta\theta - 1)/2 \rceil$$

and

$$\sigma_x^2 = \sigma_y^2 = \sigma.$$

2.3 Filter Experiments and Results

The performance of the C-Gabor filter is now tested against two other commonly used tag tracking filters, the standard Gabor filter and the HARP filter. The filters are tested by varying key parameters and examining their performance. Parameters varied for the C-Gabor filter are $\Delta\theta$ and σ . The performance of the Gabor filter is tested over the parameters σ_x and σ_y . The HARP filter requires further explanation.

In harmonic phase analysis, called HARP, an elliptically-shaped filter (frequency domain shape) with a Gaussian roll off is used to filter out a single spectral peak for the computation of local phase [30]. Here, we use a two-sided (symmetric) HARP filter for

performance comparisons with the C-Gabor filter described above. Note that a HARP filter covers the regions expected to contain energy in the deformed spectrum by setting the minor and major axes of the ellipse large enough to cover the desired region minimizing the leakage from other spectral peaks. This thesis uses the HARP filter defined in [30] centered at both the positive and negative-frequency peaks of the tag direction to be analyzed. In Section 2.3.1, the performance of this filter is studied for ranges of the semi-minor axis (a) and semi-major axis (b) defined in [30]. For brevity, the reader is referred to [30] for a detailed description of the HARP filter.

2.3.1 Experiments

To evaluate the performance of the three filters, ten human subjects were imaged with a grid-tagged protocol with a tag spacing of 7 pixels. Each study consisted of 11-13 short-axis slices and two long-axis slices. The slices were 8mm thick with no gap between slices. Twenty time frames were imaged through the cardiac cycle. In each study, all slices were analyzed at end-diastole and end-systole.

Each image is pre-filtered with each of the three filters designed above. Tag points are then detected to sub-pixel resolution by detecting local negative peaks, or “valleys”, in the image with the valley detection algorithm of Chapter 3.3. The performance of each filter is quantified by comparing the tag points detected by that filter with a set of expert-edited tag points tracked with a template-matching/active contour algorithm [31] and computing the following two measurements: the mean absolute-value (MAE) of the distance between a detected tag point and the nearest point in the expert-edited tag lines

and the percentage of detected tag points not part of a tag line (PF). For all three filters, the probability of detection was very close to 100%.

The performance of each filter was evaluated in all ten imaging studies for a range of parameter values. The C-Gabor filter was evaluated for values of σ from 0.1 to 0.6 pixels⁻¹ and $\Delta\theta$ from 10 to 50 degrees. The standard Gabor filter was evaluated for values of σ_x and σ_y from 0.1 to 0.6 pixels⁻¹. The HARP filter was evaluated for semi-major and semi-minor axes of $0.1/T$ to $1/T$, where T is the tag frequency.

2.3.2 Results

Figure 2.10 shows images of the mean absolute error (MAE) surfaces and percent false positives (PF) surfaces for the C-Gabor, standard Gabor, and HARP filters. In all three filters, there are relatively large regions where the MAE and PF are relatively small (MAE<0.4mm and PF<10%). This means that all three filters are reasonably robust to changes in their parameters.

Also, in all three filters, the MAE decreases as the size of the pass-band increases (increasing σ and $\Delta\theta$ in the C-Gabor filter; increasing σ_x^2 and σ_y^2 in the standard Gabor filter; and increasing a and b in the HARP filter). As the MAE decreases, however, the percentage of false positives increases, which means the filter does a worse job of smoothing non-tag features in the image. This increase in PF occurs because as the pass-band gets larger, more noise and other high-frequency intensity fluctuations are possible in the filtered image, and the valley detector identifies them as tag points.

Choosing an optimal set of parameters for a given filter is difficult because in each filter the minimum MAE and PF occur at significantly different parameter values. The

minimum MAE was approximately 0.3 mm in all three filters. The C-Gabor filter had the lowest PF (4.6%) followed by the HARP (6.4%) and the standard Gabor (7.1%).

The white lines in Figure 2.10 outline the region where the MAE is less than 0.4mm (the precision to which tag lines in these image can be tracked [32] and PF is less than 10%. This region is largest for the C-Gabor filter and covers the range of expected local rotation in tagged images (30-40 degrees). The region is relatively small in the standard Gabor filter, however, and does not exist in the HARP filter. This means that the C-Gabor filter can operate with an acceptable MAE and PF over a relatively wide range of parameters, while in the other two filters, acceptable MAE or PF comes at the expense of degradation in the other performance measure. While the acceptable values of MAE and PF used to define these regions were chosen arbitrarily to some extent, choosing larger values would still result in a larger acceptable region with the C-Gabor filter than in the other two filters.

Impulse and frequency responses for each filter with parameters in or, for the HARP filter, near the acceptable region are shown in Figure 2.11. For each filter, the parameters determined by the above analysis are set to cover the regions predicted by the models to contain changes due to local deformation of the myocardium. The standard Gabor and HARP filters can only cover these regions by using relatively large variances or ellipse axes in one or both directions, which can potentially pass unwanted high frequencies and allow increased leakage from other spectral peaks. The C-Gabor filter, however, uses a series of relatively narrow-band filters positioned along arcs in the frequency domain to cover these regions more efficiently.

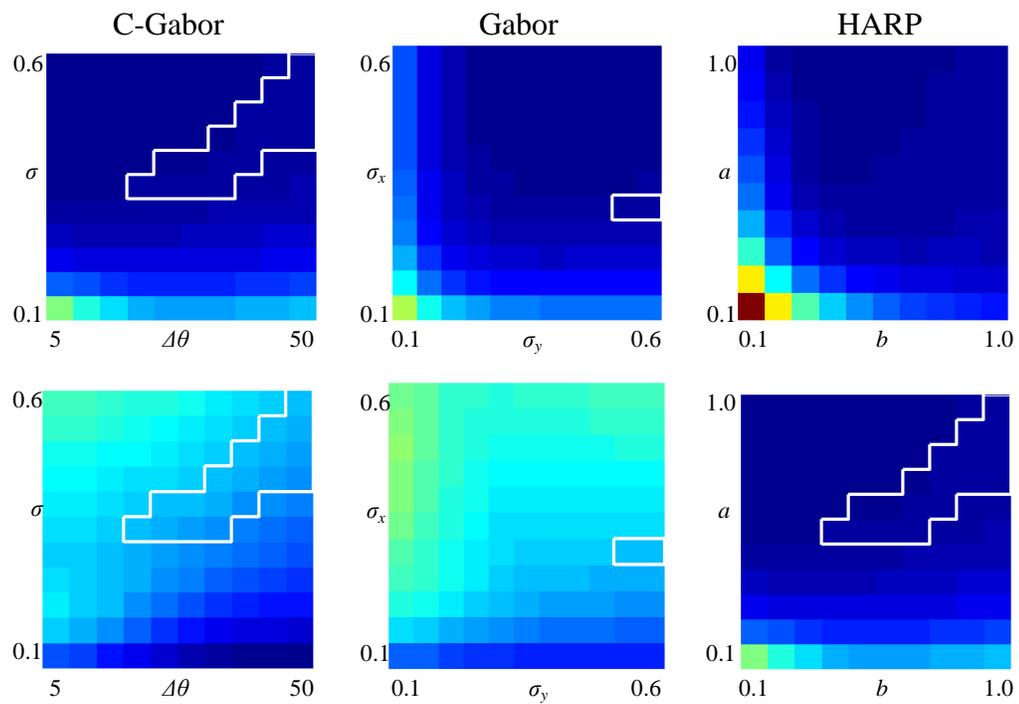


Figure 2.10: Mean absolute error [mm] (top) and percent false positives (bottom) of C-Gabor (left), Gabor (center); and HARP (right) filters

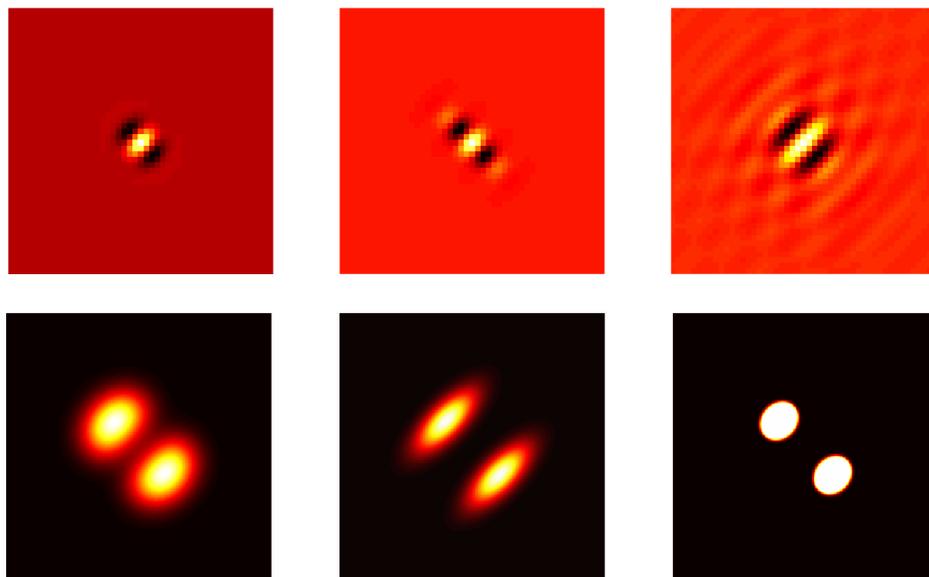


Figure 2.11: Impulse response (top) and magnitude frequency response (bottom) of C-Gabor (left; $\sigma=0.4$, $\Delta\theta=40^\circ$), standard Gabor (middle, $\sigma_y=0.6$, $\sigma_x=0.25$), and HARP (right, $b=0.5$, $a=0.4$) filters.

2.4 Chapter Summary

In this chapter, five models were developed to analytically examine the spectral changes in the 2D Fourier transform of tagged cardiac MRI data due to myocardial deformation. The models predicted that myocardial deformation should cause a spreading outward of the spectral energy from end-diastole to end-systole. The C-Gabor filter was created based on the spectral analysis. The performance of the C-Gabor filter was tested against two other commonly used tag tracking pre-filters. The C-Gabor filter was shown to be superior to standard tag tracking pre-filters as measured by mean absolute error and percent false positives.

CHAPTER 3

TAG LINE IDENTIFICATION FOR TAGGED CARDIAC MRI

3.1 Introduction

Li, *et al* developed tag line tracking based on statistical classification of candidate tag pixels [25] [33]. First, the tag lines of the end-diastolic and end-systolic images are located. Then the end-systolic tag lines are matched to the end-diastolic tag lines by statistical classification. This chapter details two methods for the first part of the statistical classification method, namely, automatic identification of tag lines. Statistical tag line classification is beyond the scope of this thesis. For details on statistical tag line classification, see [25] and [33].

A method for tag line identification based on a multi-channel segmentation of the image, wherein each channel is characterized by a distinct Gabor filter is presented in Chapter 3.2. The method is effective; however, it is only accurate to within ± 1 pixel. A second method of automatic tag line identification is detailed in Chapter 3.3. This method involves pre-filtering an image and subsequently performing a peak/valley detection routine in a direction perpendicular to the tag lines of interest. The peak/valley detection algorithm is shown to be both fast and effective with a detection rate of nearly 100%. Also, it is accurate to sub-pixel measurements. Finally, a chapter summary is presented in Chapter 3.4.

3.2 Tag Line Identification via Multi-Channel Segmentation Algorithm

This section presents an algorithm for tag line identification based on a multi-channel segmentation. Qian *et al* were the first to identify tag lines via a multi-channel segmentation, wherein various texture features are constructed based on Gabor filters [19]. The algorithm presented here is similar to the method first presented by Qian, differing mainly in the use of the nonlinear hyperbolic tangent function and exploitation of the secondary spectral peaks. The method is effective; however, one should note that the accuracy of the method proposed here is limited to ± 1 pixel. The construction of Gabor feature images used in this multi-channel segmentation is overviewed in Chapter 3.2.1. The feature images from Chapter 3.2.1 are combined using k-means clustering [35] in Chapter 3.2.2. The segmentation resulting from combining the feature images is then processed further with some morphological processing to produce the final candidate tag pixels.

3.2.1 Gabor Feature Images

Here, 6 Gabor feature images are constructed for a segmentation of tag lines in tagged cardiac MRI images. Each feature is associated with a spectral peak of the input tagged image, where the term spectral peak refers to a concentration, or peak, of energy in and/or around a particular location in the frequency domain. Both the primary and secondary spectral peaks of the tagged image spectrum are exploited. This is unlike Qian *et al* who only utilize the primary spectral peaks. The term tag spacing is defined as the distance between two adjacent tags in a tagged cardiac MRI image in units of pixels and is given the notation T_t . Also, the term tag frequency is given the notation F_t and is

defined as the mathematical inverse of tag spacing in units of pixels⁻¹. Hence we define the relation $T_t = 1/F_t$. The primary and secondary spectral peaks are found at F_t and $2F_t$ respectively. Exploiting both spectral peaks provides a much more accurate segmentation than simply utilizing the primary spectral peaks. The spectral peaks occur because the tags modulate the image spectrum to multiples of the tag frequency. Tags are quasi-sinusoidal and are, in fact, much thinner and squarer than a true 2D sinusoid. Hence, spectral peaks can be observed at F_t , $2F_t$, $3F_t$, etc. This makes sense if one considers the Fourier series expansion of a $\text{rect}()$ function, wherein the constituent frequencies are located at multiples of the fundamental frequency. Likewise, a more accurate representation of tags can be obtained via a Fourier series expansion. Figure 3.1 shows the spectral energy of a grid-tagged short axis image end-systole. This figure is intended to point out the spectral peaks of a standard tagged image. Notice that the primary peaks are located at $F = 1/T_t$ at an angle of θ_t and secondary peaks are located at $F = 2/T_t$ at an angle of θ_t .

Three Gabor features are constructed at the primary peaks and three are constructed at the secondary peaks at a given tag angle for a total of 6 Gabor features. The three Gabor features of each peak have the same center frequency and standard deviation; however, their orientations are 15 degrees apart. Let

$$h_i(\mathbf{p}) = h(\mathbf{p}) \Big|_{F=F_i, \theta=\theta_i, \sigma_x=\sigma_y=\sigma_i}$$

be the i th symmetric Gabor filter as defined in Chapter 1 with F_i , θ_i , and σ_i defined in Table 3.1 below. Note that the parameters of Table 3.1 were determined by trial and error wherein the decision to accept or reject a parameter value was based solely on a

visual inspection and/or comparison of the results of using one particular parameter value as opposed to using another. Figure 3.2 shows a contour plot of the -3 dB bandwidths of the 6 filters for clarification.

Table 3.1: Gabor filter parameters used to construct feature images

i	F [cycles/sample]	θ [degrees]	σ [cycles/sample]
1	F_t	$\theta_t - 15$	0.2
2	F_t	θ_t	0.2
3	F_t	$\theta_t + 15$	0.2
4	$2F_t$	$\theta_t - 15$	0.4
5	$2F_t$	θ_t	0.4
6	$2F_t$	$\theta_t + 15$	0.4

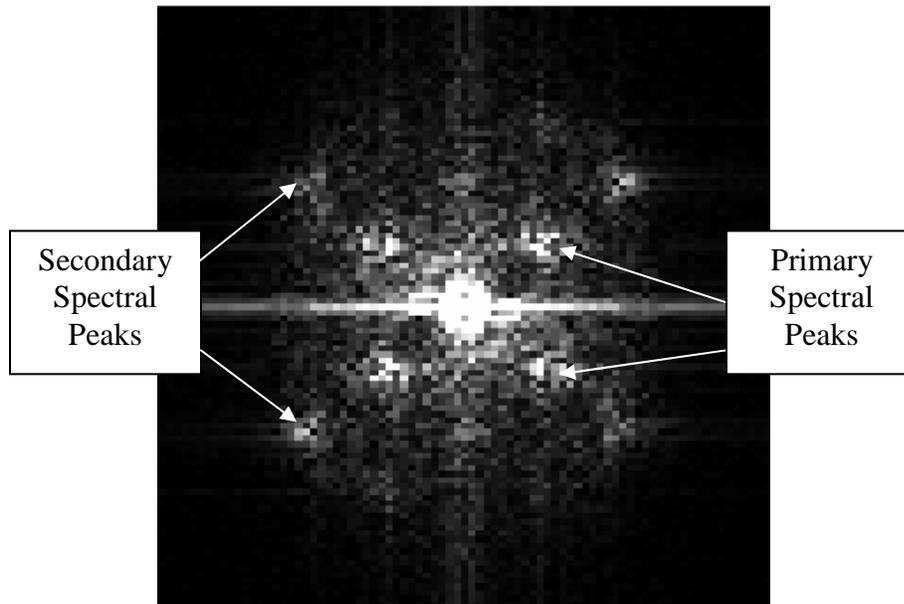


Figure 3.1: Grid-tagged short axis image ind-systole (top) and magnitude spatial frequency spectrum (bottom)

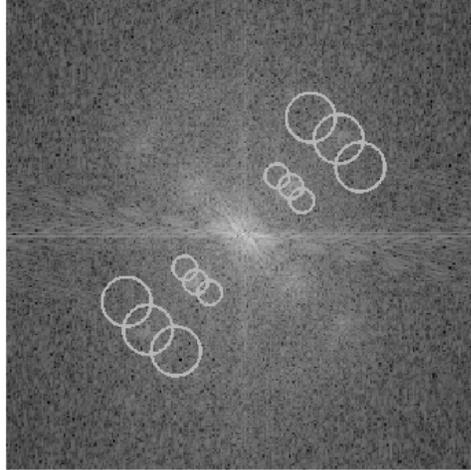


Figure 3.2: Gabor filter bank used to construct 6 Gabor features

The i th response image is given as

$$f_{ri}(\mathbf{p}) = f(\mathbf{p}) ** h_i(\mathbf{p}).$$

After the 6 response images have been constructed, each image is post-processed to produce the final feature images. The post-processing is detailed below:

- 1) Normalize each response image $f_{ri}(\mathbf{p})$ by dividing the image $f_{ri}(\mathbf{p})$ by the maximum of its absolute value

$$f'_{ri}(\mathbf{p}) = \frac{f_{ri}(\mathbf{p})}{\max[f_{ri}(\mathbf{p})]}$$

to yield the normalized response image $f'_{ri}(\mathbf{p})$. This will ensure that all values of $f'_{ri}(\mathbf{p})$ are between ± 1 .

- 2) Apply a $\tanh()$ nonlinearity to each pixel value of $f'_{ri}(\mathbf{p})$. Specifically,

$$f_{fi}(\mathbf{p}) = \tanh(\alpha f'_{ri}(\mathbf{p})),$$

where α is a constant controlling the extent of the nonlinearity.

The constant α is chosen by trial and error to be 10 wherein the decision to accept or reject a value of α was based solely on a visual inspection and/or comparison of the results of using one particular value of α as opposed to using another value of α .

Application of the $\tanh()$ function yields the final set of feature images,

$$F = F(\mathbf{p}, i) = \{f_{f_1}(\mathbf{p}), \dots, f_{f_i}(\mathbf{p}), \dots, f_{f_6}(\mathbf{p})\}.$$

The nonlinearity is intended to produce a quasi-threshold of each response image as pixel values less than 0 will be forced toward -1 and values greater than zero will be forced toward $+1$. This step is applied to give better cluster separation in the feature space of the next step, which in turn, provides easier convergence of the clustering algorithm. Figure 3.3 shows the 6 feature images used in the segmentation. Note that the endocardium and epicardium are outlined in green for reference.

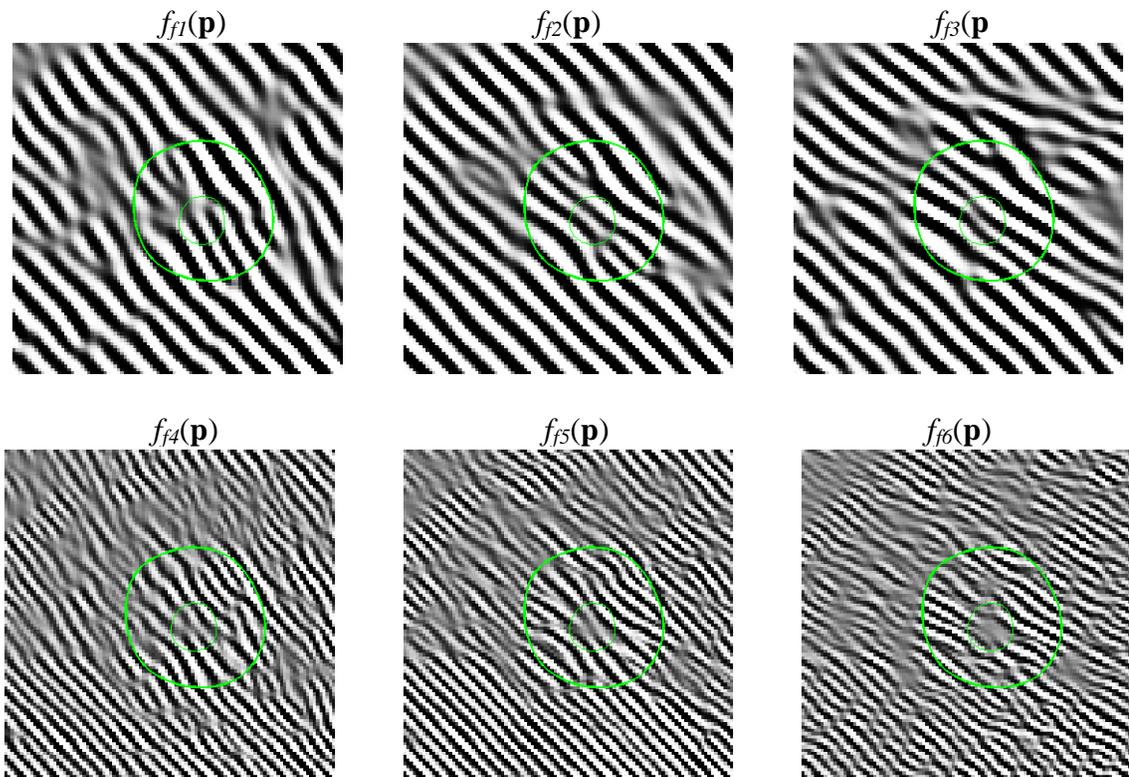


Figure 3.3: Gabor Features

3.2.2 Combining Gabor Features

Here, the six MxN Gabor features are combined via the k-means clustering algorithm [35] where the dimensions M and N refer to the number of rows and columns in an image respectively. The clustering algorithm combines the feature images into a final binary image with one level being tags and the other being myocardium. Before the k-means algorithm may be used the feature images must be reconfigured into feature vectors. Given the 6 Gabor feature images F , a feature vector is constructed which is associated with each pixel. To mathematically express the feature vectors, $F(\mathbf{p},i)$ is rewritten in lexicographical form as $F(l,i)$ by letting

$$l = (y-1)N + x$$

and $i = 1, \dots, I$. Therefore, a feature vector of length 6 is defined at each l in F . From an MxN image, a total of MxN feature vectors of length l are formed. In this case the images are $M = 256$ pixels by $N = 256$ pixels with $I = 6$ Gabor feature images. Therefore $F(l,i)$ will be of dimension $(M \times N) \times (i) = 65536 \times 6$.

Now that the feature images F , have been reconfigured into feature vectors, the k-means clustering algorithm is used to classify the feature vectors into two groups. One group will contain the tag lines and the other will contain the rest of the image. The k-means algorithm can be thought of as a mapping of the feature vectors from I dimensions down to one binary dimension, or,

$$F(l,i) \in \{\square \mid -1 \leq F(l,i) \leq 1\} \Rightarrow f_b(l) \in \{0,1\}.$$

A binary image is formed by reshaping $f_b(l)$ back into an MxN image, $f_b(\mathbf{p})$. Figure 3.4 shows an image segmented by clustering the feature vectors via the k-means clustering

algorithm wherein tag lines are white while other parts of the image are black.

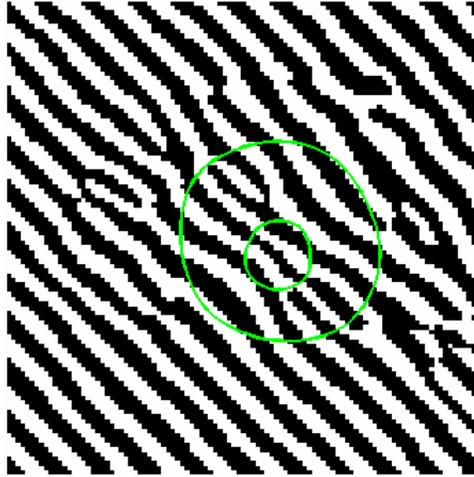


Figure 3.4: Segmentation obtained by clustering $F(l,i)$ into two groups via k-means clustering algorithm

One should note, however, that the tag lines of $f_b(\mathbf{p})$ will be much wider than the tag lines observed in the original image. However, the actual tag lines are located at the approximate center of the areas identified to be tag lines in $f_b(\mathbf{p})$. Therefore, morphological skeletal thinning is applied to the areas determined to be tag lines. Morphological skeletal thinning takes a binary image and thins areas of the image with value +1 until all areas of the image are only 1 pixel in width. This will reveal the actual tag lines to within ± 1 pixel. Figure 3.5 shows the image of Figure 3.4 after morphological skeletal thinning has been applied.

Finally, the tag pixels of interest are those contained within the myocardium. Let the set of identified tag pixels after morphological thinning be Ω_r . Then the set of tags of interest are Ω_T , where

$$\Omega_T = \{ \Omega_r \mid \Omega_r \in \Omega_{myo} \},$$

and Ω_{myo} is defined to be all points inside the contour of the epicardium and outside the contour of the endocardium. Figure 3.6 shows two images. The left image is the original image and the right image is that of the candidate tag pixels, Ω_T , superimposed over the original image $f(\mathbf{p})$. The candidate tag pixels are highlighted in white. Note that in each image the endocardium and epicardium are showcased as green contours. Last, Figure 3.7 shows a block diagram of the algorithm.

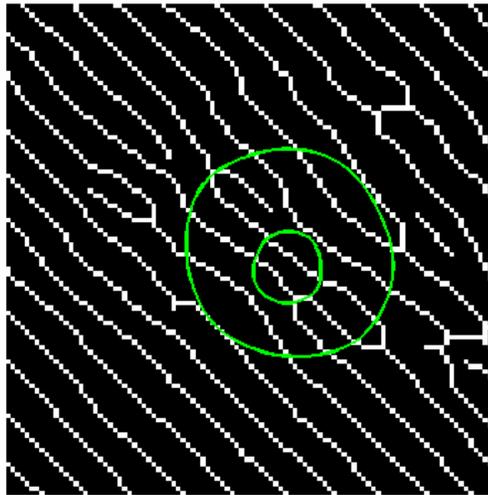


Figure 3.5: Set of identified tag pixels after morphological skeletal thinning

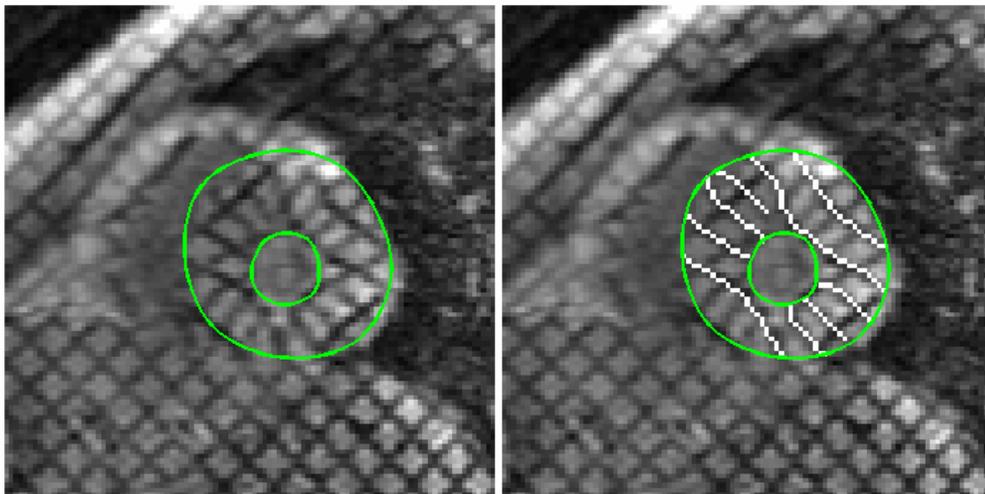


Figure 3.6: Original image(left) and original image with identified candidate tag pixels (right)

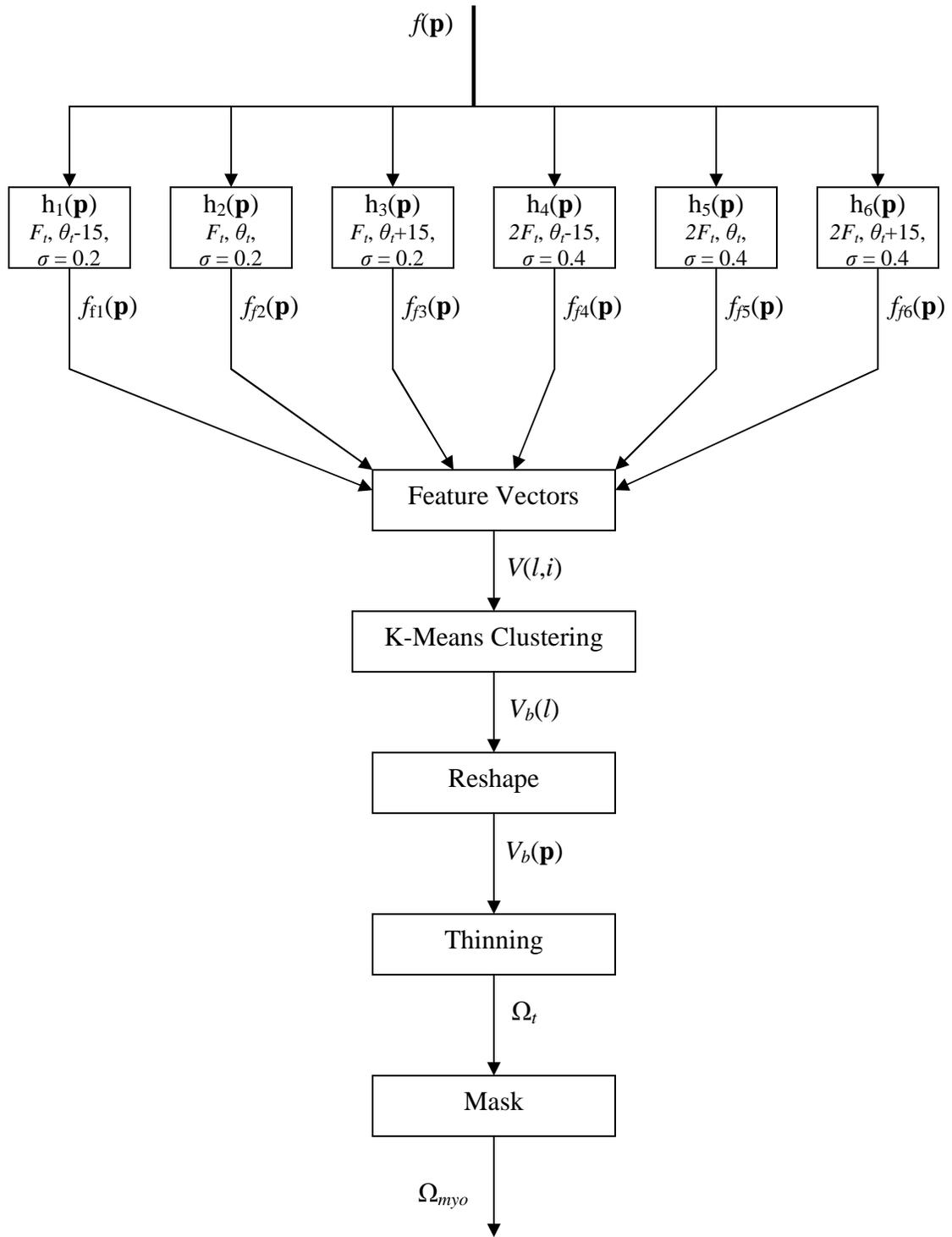


Figure 3.7: Block diagram of multi-channel tag line identification algorithm

3.3 Tag Line Identification via Peak/Valley Detection Algorithm

In this section a new algorithm is presented for tag line identification. The algorithm is based on a peak/valley detection of a pre-filtered image. A brief explanation of the algorithm is given in Chapter 3.3.1. A computationally efficient implementation called QuickTag is given in Chapter 3.3.2. The algorithm has a detection rate of nearly 100% with low rate of false positives. A false positive is defined to be the false classification of a pixel as being located on a tag line. A C-Gabor filter is used for pre-filtering because it was shown in Chapter 2 to be superior to the standard Gabor and HARP filters for this particular application.

3.3.1 Derivation

Candidate tag pixels are located by detecting the local minimum of a pre-filtered image in a direction perpendicular to the tag lines of interest. In short, the algorithm works by finding areas of the filtered image where the first derivative is equal to zero and the second derivative is greater than zero. The derivatives are taken in the direction perpendicular to the tags of interest. Figure 3.8 shows a grid-tagged short axis image near end-systole. From Figure 3.8, the image on the left shows the original image while the image on the right shows the same image with pre-filtering applied. Notice in the image on the right that the tags in the -45 degree direction have been removed and non-tag features have been smoothed.

Now, Figure 3.9 shows the first and second derivatives of the pre-filtered image of Figure 3.8 taken in a direction perpendicular to the tag lines of interest. Below the first derivative image of Figure 3.9 the areas where the first derivative is zero is shown with

red contours. Below the second derivative image, a binary image shows areas where the second derivative image is greater than zero. The tags can be identified as the areas where the first derivative is equal to zero and the second derivative is greater than zero as is demonstrated in Figure 3.10.

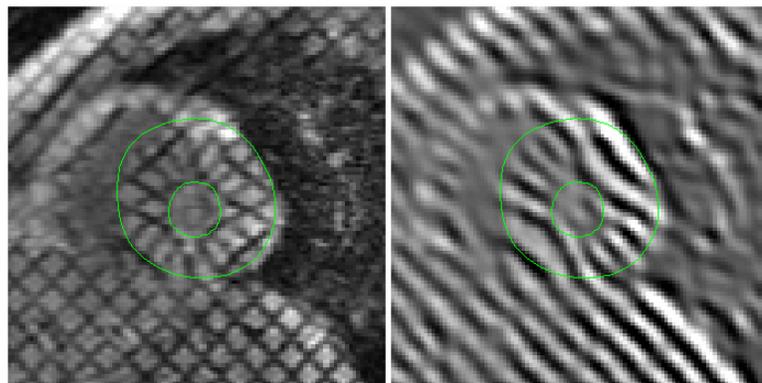


Figure 3.8: Original Image (left) and pre-filtered image (right) for a tag angle of 45 degrees. The endocardium and epicardium are outlined in green

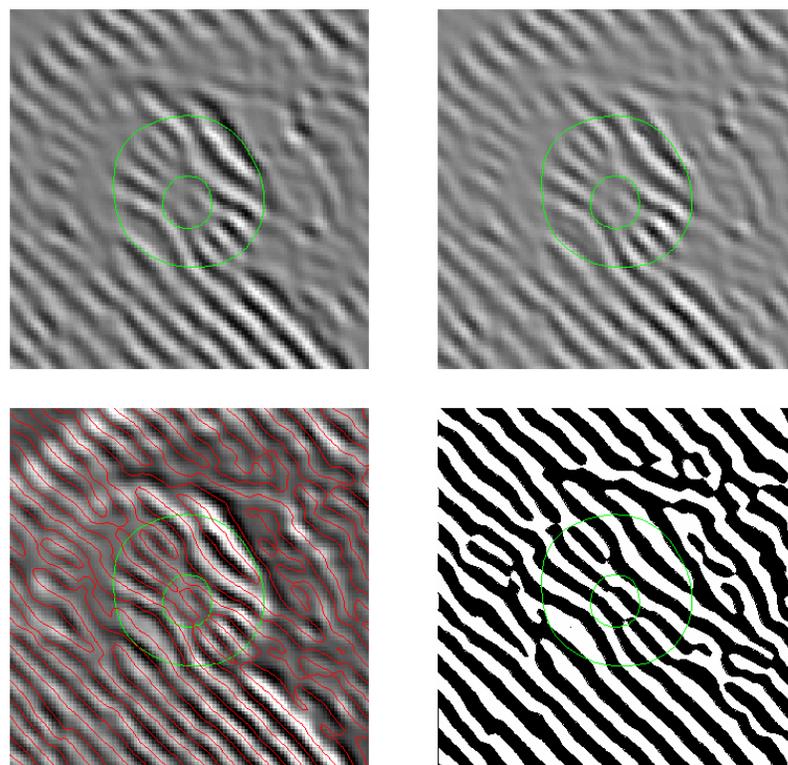


Figure 3.9: First Derivative shown left and second derivative shown right

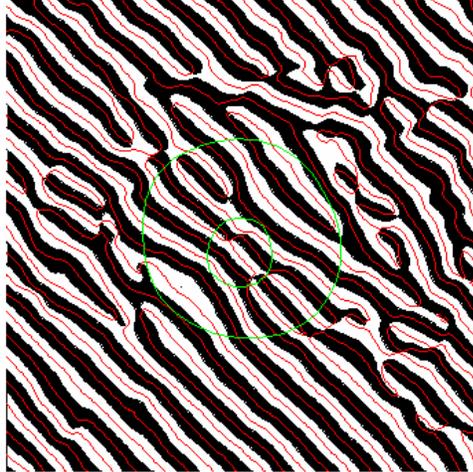


Figure 3.10: Tags are identified as areas in the image above where the red contours are surrounded by white

3.3.2: QuickTag Algorithm

In this section the QuickTag Algorithm is detailed. For illustration purposes, a short axis, end-systolic image will be processed to detect tag lines at a tag angle of 45 degrees throughout this section. One should note, however, that the algorithm works equally as well for other tag angles and other views of the heart. A green contour denotes the endocardium and epicardium of the left ventricle. Also, an enlarged version of the image will be shown revealing only the area of interest. Figure 3.11 shows the example image to be processed (left) along with the area of interest (right).

Let $f(\mathbf{p})$ be the original image and $f_f(\mathbf{p})$ be $f(\mathbf{p})$ filtered with the appropriate filter, $h(\mathbf{p})$, to extract tags at a tag angle of θ_t . Here, the filter of choice is the C-Gabor filter because it is shown, in Chapter 2 of this thesis, to be superior to the standard Gabor and HARP filters for this particular application. For comparison purposes, Figure 3.12 shows $f(\mathbf{p})$ (left) and $f_f(\mathbf{p})$ (right). In Figure 3.12 (right), $f(\mathbf{p})$ has been filtered to reveal tags at an angle of $\theta_t = 45$ degrees. C-Gabor filter parameters used are $\sigma = 0.4 \text{ pixels}^{-1}$ and $\Delta\theta = 40$

degrees because they were determined to be optimal in Chapter 2. Once again, green contours are used to denote the endocontours and epicontours. Notice how filtering with $h(\mathbf{p})$ has virtually eliminated the -45 degree tags from $f(\mathbf{p})$. Also, non-tag image features have been smoothed which helps reduce instances of false positives in subsequent steps.

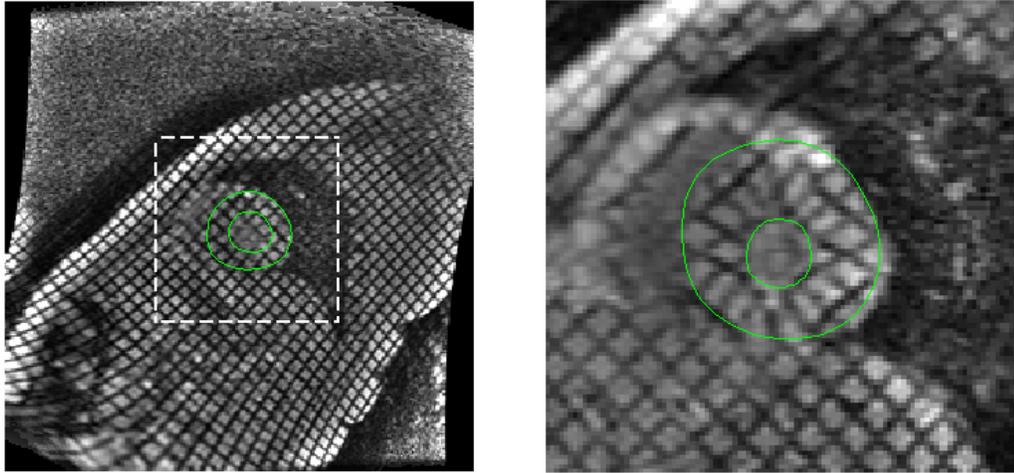


Figure 3.11: Example image to demonstrate proposed algorithm. Green contours outline the left ventricle. Right image shows only the area of interest outlined with the dashed white line in the left image.

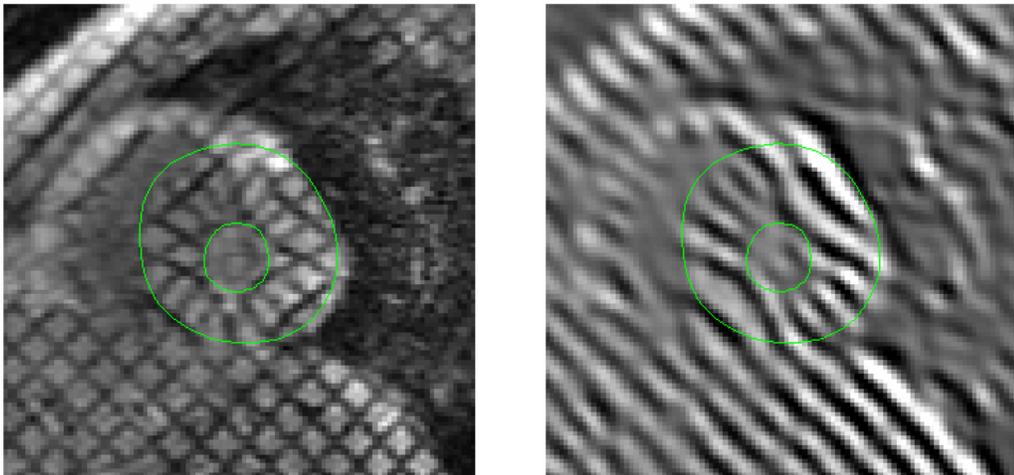


Figure 3.12: Original Image (left) and pre-filtered image (right) for a tag angle of $\theta_t = 45$ degrees

Now two kernels of convolution are defined which will reveal the areas of positive and negative slope of $f_j(\mathbf{p})$ with respect to the direction perpendicular to the tag lines of interest. The two kernels are defined to be dimension 3x3 and will be denoted as c_p for positive kernel and c_n for negative kernel. Table 3.2 summarizes the possible values of c_p and c_n for the standard tag angles of 0, 90, +45, and -45 degrees. Note that the kernels are similar to a Roberts compass gradient operator.

Table 3.2: c_p and c_n for standard tag angles

θ	c_p	c_n
0	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
90	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
-45	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
45	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Filtering with the kernel c_p returns a positive result at pixel locations of $f_j(\mathbf{p})$ where the gradient is positive with respect to the direction perpendicular to the tag lines and negative otherwise. Filtering with c_n return a positive result at pixel locations where the gradient is negative with respect to the direction perpendicular to the tag lines and negative otherwise. It also provides a 1 pixel shift in the tag-normal direction causing the

pixels containing the valleys to align. Two new images are formed after the convolution operations: $f_{fp}(\mathbf{p})$ for the positive-sloping, filtered image and $f_{fn}(\mathbf{p})$ for the negative-sloping, filtered image. One important detail to note is that the negative-sloping image can be calculated by simply taking the negative of the positive-sloping image and then performing a 1-pixel shift in the tag-normal direction. Figure 3.13 shows an example $f_{fp}(\mathbf{p})$ or $f_{fn}(\mathbf{p})$ for the example image.

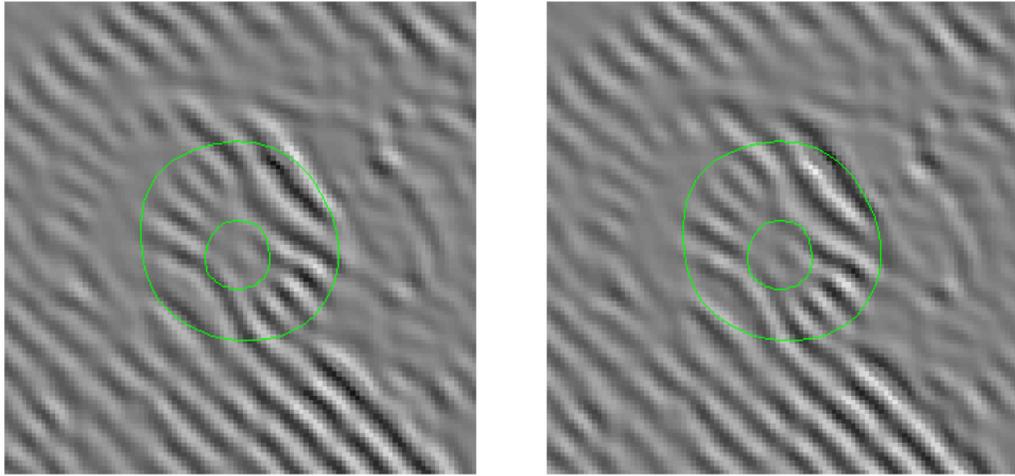


Figure 3.13: $f_{fp}(\mathbf{p})$ and $f_{fn}(\mathbf{p})$ for a tag angle of 45 degrees

Next, two binary images, $f_{bp}(\mathbf{p})$ and $f_{bn}(\mathbf{p})$, are formed by thresholding $f_{fp}(\mathbf{p})$ and $f_{fn}(\mathbf{p})$ at zero. Therefore, any pixel of $f_{bp}(\mathbf{p})$ or $f_{bn}(\mathbf{p})$ with intensity less than zero is set to zero and any pixel of $f_{bp}(\mathbf{p})$ or $f_{bn}(\mathbf{p})$ with intensity greater than zero is set to 1. $f_{bp}(\mathbf{p})$ and $f_{bn}(\mathbf{p})$ are calculated as follows:

$$f_{bp}(\mathbf{p}) = \frac{\text{sign}(f_{fp}(\mathbf{p})) + 1}{2}$$

and

$$f_{bn}(\mathbf{p}) = \frac{\text{sign}(f_{fn}(\mathbf{p})) + 1}{2}$$

These images are called the positive slope and negative slope images because they highlight areas of an image with positive and negative slope with respect to the direction perpendicular to the tag lines. Figure 3.14 shows the two binary slope images for the example image.

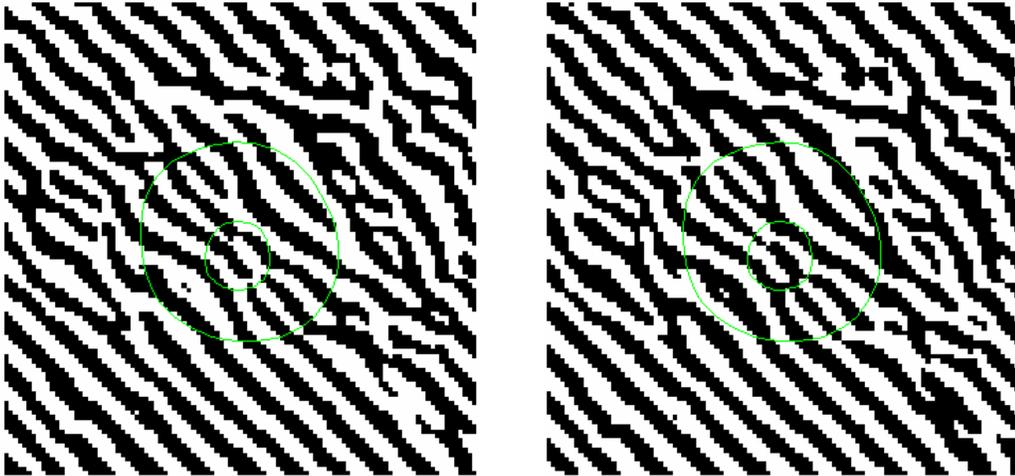


Figure 3.14: $f_{bp}(\mathbf{p})$ and $f_{bn}(\mathbf{p})$ for the example image

An interesting property of $f_{bp}(\mathbf{p})$ and $f_{bn}(\mathbf{p})$ is that pixel values of each image will only be both 1 when the valley of an image has been detected. Hence a tag image $f_t(\mathbf{p})$ can be produced where candidate tag pixels are set to +1 and all other pixels are set to zero by taking the logical AND of $f_{bp}(\mathbf{p})$ and $f_{bn}(\mathbf{p})$; or,

$$f_t(\mathbf{p}) = f_{bp}(\mathbf{p}) \& f_{bn}(\mathbf{p}).$$

Figure 3.15 shows the tag image, $f_t(\mathbf{p})$, for the example image.

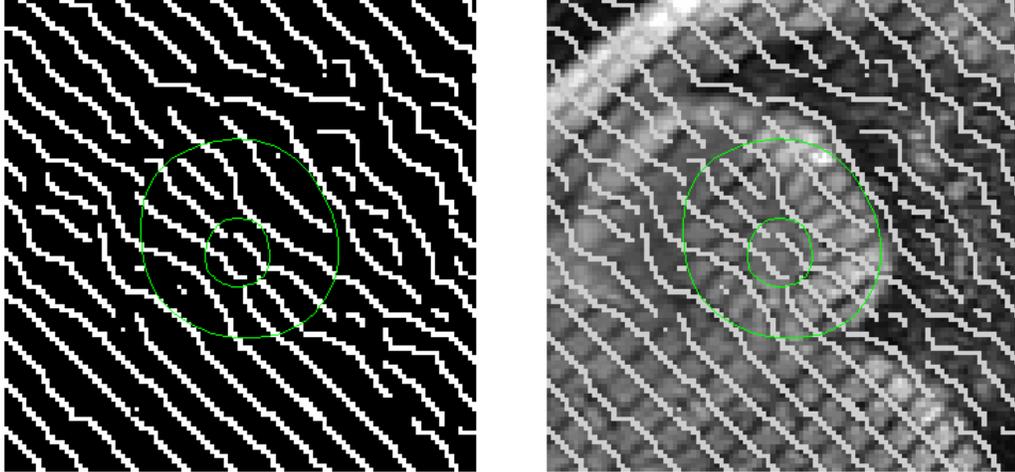


Figure 3.15: $f_t(\mathbf{p})$ shown left and $f_t(\mathbf{p})$ superimposed over original image, $f(\mathbf{p})$ shown right

Together, the kernels have acted to perform the first and second derivative operations. Also, $f_{bp}(\mathbf{p})$ and $f_{bn}(\mathbf{p})$ make ready the information necessary to determine places of $f(\mathbf{p})$ where the first derivative is zero and the second derivative is greater than zero. Finally, the tags pixels of interest are defined to be Ω_t , where

$$\Omega_t = \{f_t | f_t \in \Omega_{myo}\},$$

and Ω_{myo} is defined to be all points inside the contour of the epicardium and outside the contour of the endocardium. Figure 3.16 shows Ω_t superimposed over the original image $f(\mathbf{p})$.

Candidate tag pixels of Ω_t are only accurate to ± 1 pixel. Quadratic polynomial interpolation is used to compute candidate tag point position down to sub-pixel resolution using the binary image, $f_b(\mathbf{p})$, and the filtered image, $f_f(\mathbf{p})$. The sub-pixel refinement uses the candidate tag pixel location and the two pixels in the direction perpendicular to the tag lines of interest from each candidate tag pixel in the pre-filtered image, $f_f(\mathbf{p})$. A quadratic polynomial is fit to these three points using their intensity values and their

spatial coordinates. The candidate tag point location is calculated as the location of the minimum of the quadratic polynomial fit to these three points. The final set of refined candidate tag points is called Ω_T . Figure 3.17 shows Ω_T as the red points superimposed over the original image. Also, Figure 3.18 shows two long axis views end-systole with detected tag points, Ω_T , to demonstrate that the algorithm works equally well on other image views. Last, Figure 3.19 shows a block diagram of the QuickTag algorithm implemented with FFT's.

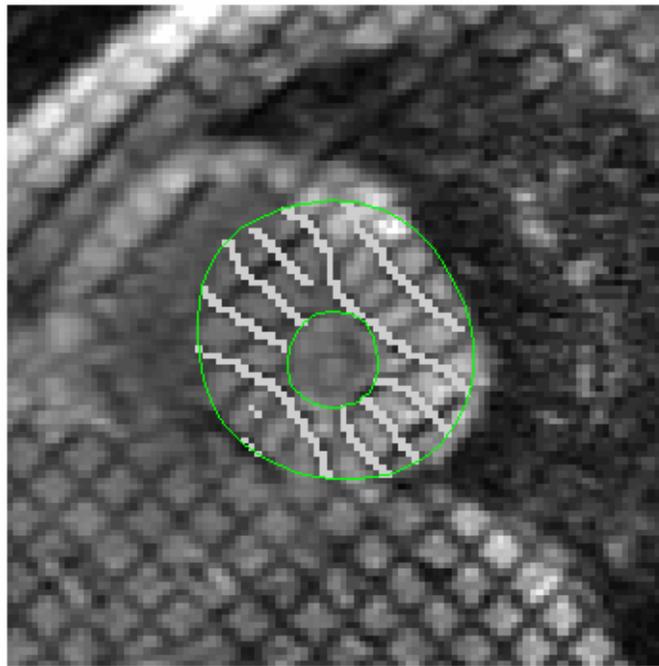


Figure 3.16: Ω_t superimposed over original image, $f(\mathbf{p})$

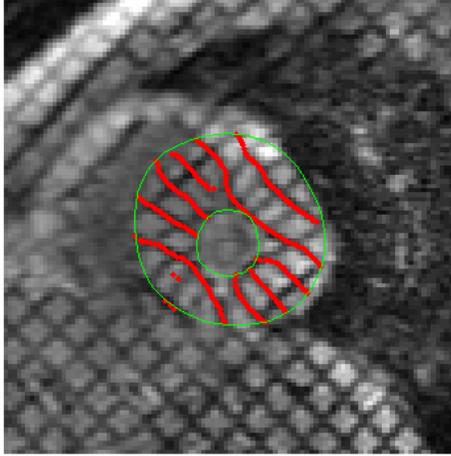


Figure 3.17: Ω_T superimposed over $f(\mathbf{p})$

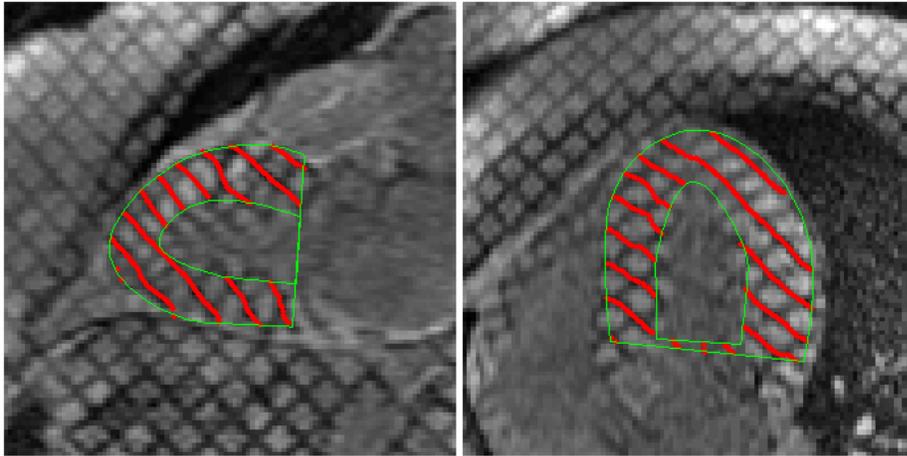


Figure 3.18: 2-Chamber View (left) and 4-Chamber view (right) of tags detected with the QuckTag algorithm

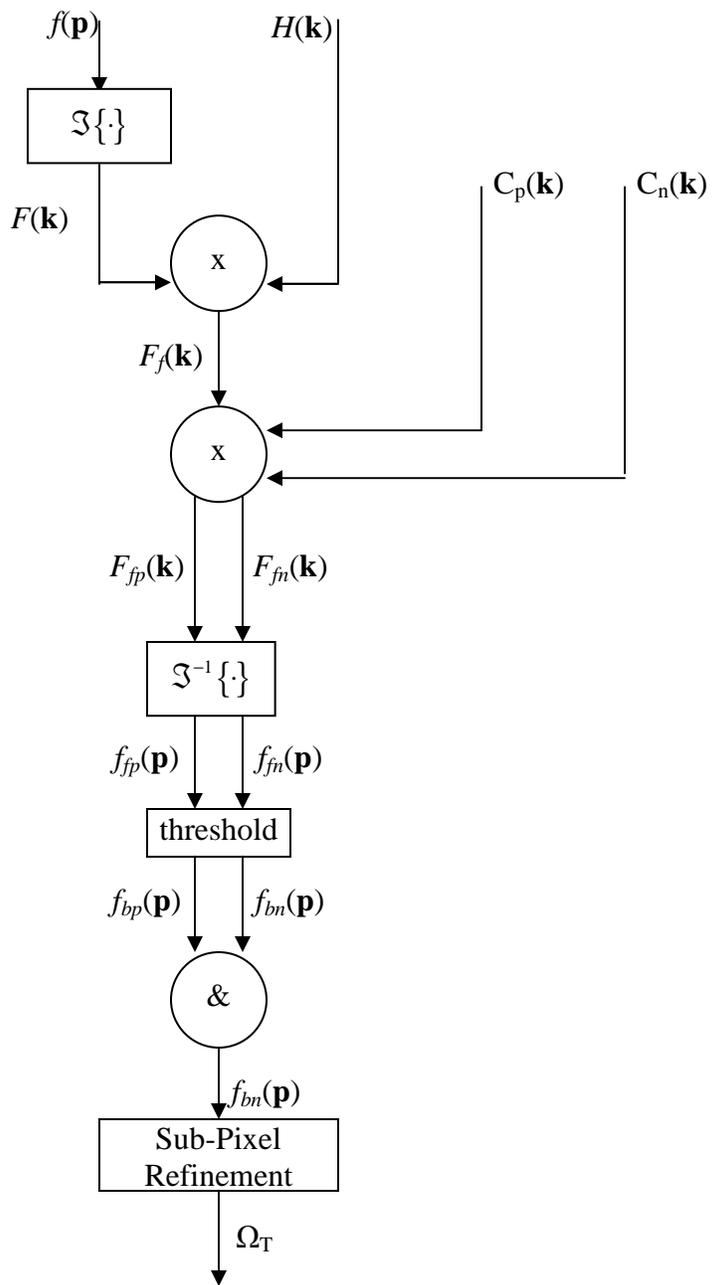


Figure 3.19: QuickTag algorithm

3.4 Chapter Summary

In this chapter two algorithms for tag line identification of tagged cardiac MRI images were detailed. A method for tag line identification based on a multi-channel segmentation was presented in Chapter 3.2. The method is slower than the QuickTag algorithm of Chapter 3.3 due to more filtering and the clustering operation involved. Also, the multi-channel method is only accurate to ± 1 pixel. Future work on this algorithm would include a sub-pixel refinement.

A new algorithm for tag line identification based on a local peak detection of a pre-filtered image is presented in Chapter 3.3. A basis for the algorithm is derived in Chapter 3.3.1 while a detailed an implementation called the QuickTag is presented in Chapter 3.3.2. QuickTag is a computationally efficient implementation of the pre-filtering and peak detection method. The QuickTag algorithm is fast and shown to be accurate down to sub-pixel resolution.

CHAPTER 4
MULTI-CHANNEL SEGMENTATION OF MYOCARDIUM IN CARDIAC MRI
IMAGES

4.1 Introduction

In the last chapter Gabor features to perform an automatic segmentation of tags in tagged cardiac MRI images. In this chapter the use of Gabor features is extended to the segmentation of the left ventricle (LV) in cine cardiac MRI images. Unlike the multi-channel algorithm presented in Chapter 3.1, the algorithm of Chapter 4 is a supervised method in the sense that seed points must be selected for vector classification. To assist with the segmentation a graphical user interface (GUI) was written in order to select seed point samples from the myocardium, blood pool, and air. A detailed description of the segmentation algorithm of this chapter including pre-processing, channel characterization, and response image post processing to produce feature images is given in Chapter 4.2. An experiment to test channel parameters and presents the results of the experiment is given in Chapter 4.3. Finally, a chapter summary is provided in Chapter 4.4.

4.2 Methodology

The method presented in this chapter is a combination of two popular publications on the segmentation of textured images, namely, “Texture Analysis Based on a Human VisualModel” [3], and “Unsupervised Segmentation of Textured Images Using Gabor Filters” [10]. The segmentation algorithm works by constructing several channels wherein each channel is characterized by a total of 4 Gabor filters as in [3]. Image pre-processing involves a local contrast enhancement to remedy the infamous intensity inhomogeneity often found in MRI images. Post processing includes a $\tanh()$ nonlinearity, and local energy computation as is used in [10].

4.2.1 Pre-Processing

MRI images are notorious for exhibiting intensity inhomogeneity [19] across the image (see Figure 4.1). Intensity inhomogeneity simply means that the range of intensity values observed over the image varies with location. A local intensity correction similar to [19] is applied to remedy the problem. If $f(\mathbf{p})$ is the original image and $f_a(\mathbf{p})$ is an image given by the local average of $f(\mathbf{p})$ over a small window, then an intensity corrected image, $f_c(\mathbf{p})$, is founded by

$$f_c(\mathbf{p}) = \frac{f(\mathbf{p})}{f_a(\mathbf{p}) + \delta},$$

where δ is a small value to ensure $f(\mathbf{p})$ is not divided by zero. Figure 4.1 shows an example image before and after local intensity correction.

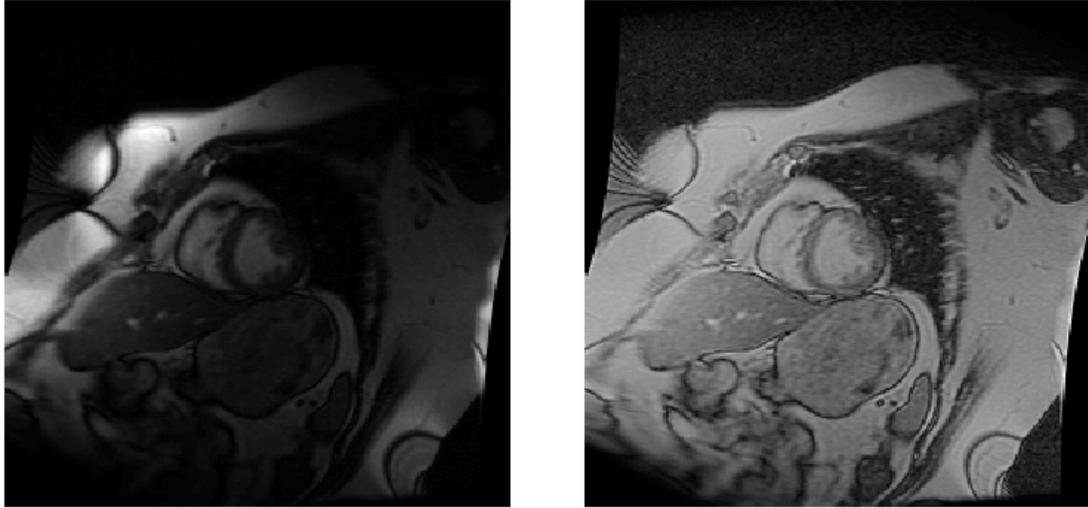


Figure 4.1: Original image (left) and intensity corrected image (right)

4.2.2 Channel Characterization

An HV- based channel characterization is used here to construct feature images.

According to [3], the early stages of the HVS are described by a dense set of channels which decompose the visual scene into a set of response images. Each channel provides limited frequency-domain coverage; however, the multiple channels collectively provide relatively uniform coverage of the 2D spatial-frequency domain.

Simple cortical cells located in the striate cortex are the biological mechanism responsible for channel characterization and are said to perform 2D spatial filtering along with a non-linear half-wave rectification stage at the output [3]. The response of each simple cortical cell is Gabor-like and, therefore, responds to a limited range of frequencies and orientations. Simple cells come in two types, symmetric simple cells and anti-symmetric simple cells, where symmetric and anti-symmetric refers to their receptive field profile (spatial impulse response for us EE's). Each symmetric simple cell is paired with an anti-symmetric simple cell tuned to the same frequency and orientation.

Together, the two filters are a quadrature filter pair. According to [3], the output of each simple cell is half wave rectified. Due to the half-wave rectification, two other simple cells tuned to the same frequency and orientation but with opposite polarity are used to preserve information. Therefore, each channel of the HVS is characterized by four Gabor-like biological filters tuned to the same frequency and orientation, wherein the output of each cell is half-wave rectified and summed together to produce the channel output.

Based on the above model, the following channel characterization is given. The i th channel is characterized by filtering (i.e. 2D convolution) the input image $f(\mathbf{p})$ with the four Gabor filters given by

$$a_{pi}(\mathbf{p}) = \exp\left\{\frac{1}{2}\left[(\mathbf{R}\mathbf{p})^T \mathbf{B}(\mathbf{R}\mathbf{p})\right]\right\} \cos\left(2\pi(\mathbf{w}^T \mathbf{p})\right),$$

$$b_{pi}(\mathbf{p}) = \exp\left\{\frac{1}{2}\left[(\mathbf{R}\mathbf{p})^T \mathbf{B}(\mathbf{R}\mathbf{p})\right]\right\} \sin\left(2\pi(\mathbf{w}^T \mathbf{p})\right),$$

$$a_{ni}(\mathbf{p}) = -\exp\left\{\frac{1}{2}\left[(\mathbf{R}\mathbf{p})^T \mathbf{B}(\mathbf{R}\mathbf{p})\right]\right\} \cos\left(2\pi(\mathbf{w}^T \mathbf{p})\right),$$

and

$$b_{ni}(\mathbf{p}) = -\exp\left\{\frac{1}{2}\left[(\mathbf{R}\mathbf{p})^T \mathbf{B}(\mathbf{R}\mathbf{p})\right]\right\} \sin\left(2\pi(\mathbf{w}^T \mathbf{p})\right).$$

where

$$\mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}$$

is the 2D spatial point,

$$\mathbf{R} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

is the 2D rotation matrix,

$$\mathbf{w} = \begin{bmatrix} 2\pi u_0 \\ 2\pi v_0 \end{bmatrix} = \begin{bmatrix} 2\pi f_0 \cos \theta_0 \\ 2\pi f_0 \sin \theta_0 \end{bmatrix}$$

is a 2D tag frequency, and

$$\mathbf{B} = \begin{bmatrix} 1/\sigma_x^2 & 0 \\ 0 & 1/\sigma_y^2 \end{bmatrix}$$

is a matrix of standard deviation values. The subscripts “ p ” and “ n ” of $a_p(\mathbf{p})$ and $b_p(\mathbf{p})$ refer to positive and negative polarity filters respectively. Also, $a(\mathbf{p})$ and $b(\mathbf{p})$ refer to symmetric and anti-symmetric filters respectively.

After filtering, $a_{pi}(\mathbf{p})$ and $b_{pi}(\mathbf{p})$ are half-wave rectified and subsequently added together to produce $q_{pi}(\mathbf{p})$. Also, $a_{ni}(\mathbf{p})$ and $b_{ni}(\mathbf{p})$ are half wave rectified and subsequently added together to produce $q_{ni}(\mathbf{p})$. Last, $q_{pi}(\mathbf{p})$ and $q_{ni}(\mathbf{p})$ are added together to produce the final channel response image $f_{ri}(\mathbf{p})$. Figure 4.2 shows a block diagram of the i th channel tuned to the frequency f_0 and orientation θ_0 .

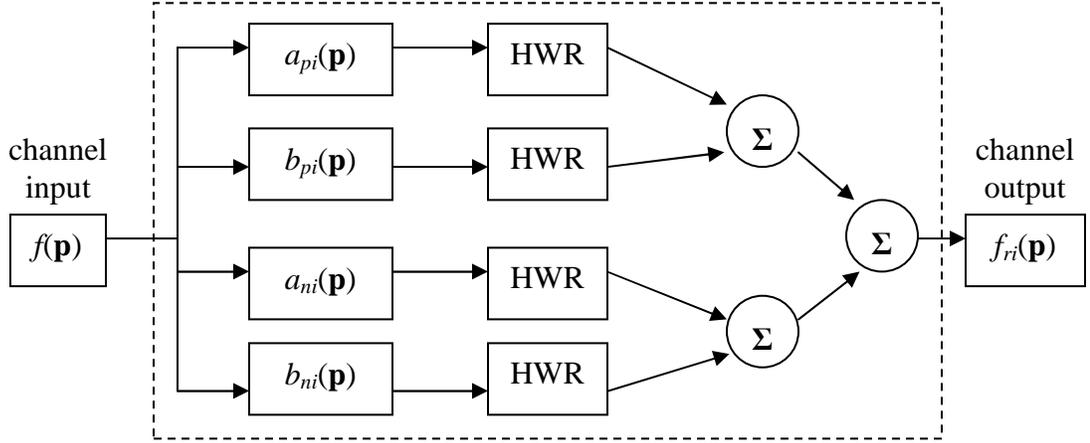


Figure 4.2: Block diagram of i th channel characterization. Note that HWR stands for half-wave rectification.

4.2.3 Channel Post Processing

The i th response image is post-processed by applying the nonlinear hyperbolic tangent function given by

$$f_{ii}(\mathbf{p}) = \tanh\left(\alpha \frac{f_r(\mathbf{p})}{\max(f_r(\mathbf{p}))}\right),$$

where α is a constant which controls the extent of the nonlinearity. This step is used in [10] and has the property of providing better separation in feature space.

Last, a local average energy computation is performed by calculating the average of each image $f_{ii}(\mathbf{p})$ over a small window to produce the final feature images $f_f(\mathbf{p})$ as is done in [10]. Generally, the window should be inversely proportional to the center frequency, f_i , of each channel. Hence, the window size, w , should be $w = \lceil \lambda f_i \rceil$, where $\lceil \cdot \rceil$ is the ceiling operation and λ is a constant to control the window size. It is determined by trial and error that the algorithm is not particularly sensitive to the choice

of λ so we let λ be 2 for this application.

4.2.4 Multi-Channel Bank Construction

Like its HVS counterpart, this algorithm employs multiple channels to provide relatively uniform coverage of the frequency domain. Each channel is characterized by the model of Section 4.2.2 and is tuned to a given center frequency f_i and orientation θ_i . Typically, spatial variations below $f_o = 0.3 \text{ pixels}^{-1}$ can be considered too large to possess texture information; therefore, filters are constructed to have center frequencies above f_o [10]. The multi-channel bank is composed of 16 channels, characterized by a combination of 4 center frequencies and 4 orientations. The center frequencies and orientations of the four filters are given in Table 4.1 below where f_i is given in units of pixels^{-1} and θ_i is given in degrees.

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
f_i	0.3	0.3	0.3	0.3	0.35	0.35	0.35	0.35	0.4	0.4	0.4	0.4	0.45	0.45	0.45	0.45
θ_i	0	0	0	0	45	45	45	45	90	90	90	90	135	135	135	135

Table 4.1 f_i and θ_i used for channel characterization

The Gabor filters used in this algorithm are circular (i.e. $1/\sigma = 1/\sigma_x = 1/\sigma_y$). Also, as the center frequency, f_i , of the channel gets larger the standard deviation of the filter is also increased. Hence the standard deviation of a filter is given by $\sigma = \beta f_i$, where β is a constant greater than zero controlling the channel bandwidth. Increased bandwidth with center frequency is common in the literature and is a property of the HVS [10]. The set of channel response images is referred to as

$$F_r(\mathbf{p}, i) = \{f_{r_1}(\mathbf{p}), f_{r_2}(\mathbf{p}), \dots, f_{r_i}(\mathbf{p}), \dots, r_{r_l}(\mathbf{p})\},$$

where $f_{r_i}(\mathbf{p})$ is the response image of the i th channel and the set is composed of I elements (16 elements in this case). Feature images are constructed by normalizing the intensity values of each response image between -1 and 1 and subsequently applying the nonlinear hyperbolic tangent function. Specifically, the i^{th} feature image, $f_{f_i}(\mathbf{p})$ is calculated as

$$f_{f_i}(\mathbf{p}) = \tanh(\alpha f_{r_i}(\mathbf{p})),$$

where α is a constant to control the extent of the nonlinearity. Optimal values of α and β are calculated in the experimental results section of this chapter. The set $F_f(\mathbf{p}, i)$ is used to construct the set of response vectors to be classified according the method of the next section. The construction of response vectors, $F_f(l, i)$, from a set of response images, $F_f(\mathbf{p}, i)$, is covered thoroughly in Chapter 3 Section 2.2, and, therefore, will not be revisited here.

4.2.5 Vector Classification

A supervised vector classification is used here to assist in locating and segmenting the myocardium. From each image, several sample pixels are selected from the myocardium, blood pool, and air. From the set of sample pixels, a mean response vector is calculated for each type of tissue. Each response vector is then classified as belonging to the myocardium, blood pool, and air based on its vector distance to each of the three mean response vectors. Specifically, a response vector will be assigned to category x if its vector distance to this mean response vector is smaller than the vector distance to the other two response vectors. Figure 4.3 shows the GUI written to assist in the

segmentation. Note that the green “x’s” are myocardium seed pixels, red “x’s” are blood pool seed pixels, and yellow “x’s” are air seed pixels.

4.3 Experiments and Results

The segmentation obtained via the algorithm above is tested for optimal parameters. Specifically, the channel bandwidth constant β and hyperbolic tangent constant α are tested over a wide range of values. The channel bandwidth constant is tested over a range from 0.5 to 3 while the hyperbolic tangent constant is tested over the range 1 to 10.

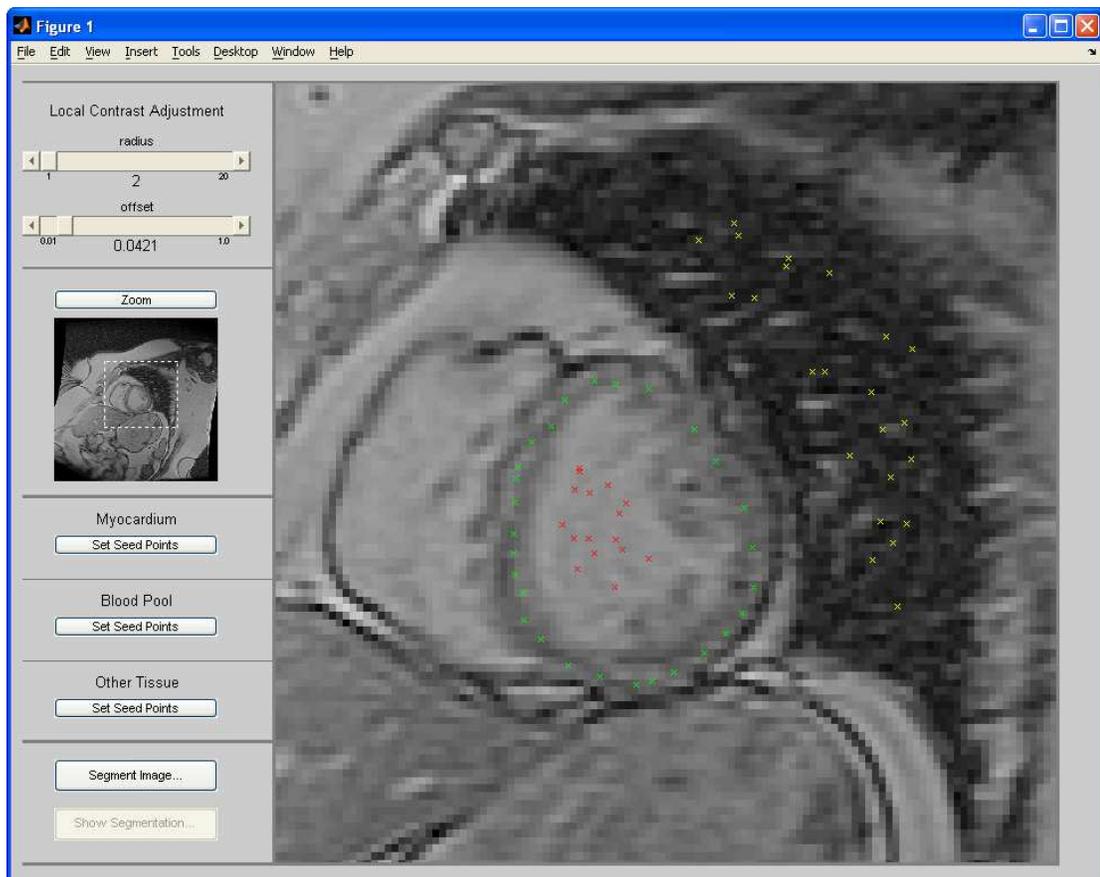


Figure 4.3: GUI written to assist in supervised segmentation

Ten human short axis studies were used to test the parameters α and β . The studies contained 4 hypertensive patients resistant to medication, 2 hypertensive patients responding to medication, 1 mitral valve regurgitation patient, 2 volume overload patients, and 1 diabetic with myocardial infarction. Four mid ventricle slices (2 end-diastole and 2 end-systole) were selected from each study for a total of 40 images. Appendix 4.1 at the end of this chapter lists the patient ID number and visit for each study that was analyzed. Each segmented image obtained for a given value of α and β is compared to a set of hand-segmented contours containing the myocardium by calculating the probability of false positives (P_f) and probability of detection (P_d). Probability of false positive is defined as the percentage of pixels identified as belonging to the myocardium when, in fact, they do not. Probability of detection is defined as the percentage of pixels identified as not belonging to the myocardium when, in fact, they do. P_f and P_d is calculated over the range of parameter values for the 40 test images.

Figure 4.4 shows the results of the experiment as images displaying the average P_f and P_d over the range of parameters tested. Note in Figure 4.4 that the channel bandwidth variable β should be chosen above 1.2, because, below values of 1.2, both the P_f and P_d are high. However, as β is increased above 1.2, both the P_f and P_d images are relatively flat with some local minimums, meaning that the algorithm is not sensitive to choice of β above a value of 1.2.

The algorithm appears to vary little with the choice of hyperbolic tangent constant α with possible local minimums observed in the range of $\alpha = 4$ to $\alpha = 9$. The “best” parameters were determined by finding image locations where both the P_f and P_d are relatively small in both the P_f and P_d images simultaneously. This area is outlined with

the white ellipses in Figure 4.4. Hence the parameters chosen as best for this application are $\beta \sim 1.5$ and $\alpha \sim 7$. Also, Figure 4.5 shows some example segmentation results obtained using optimal parameters.

Last, an experiment is conducted to determine the average P_f and P_d for the optimal parameters of $\beta = 1.5$ and $\alpha = 7$. The same ten human short-axis studies were used; however, the segmentation algorithm was run over different set of images (slices) from each study. This test revealed the average P_f to be 0.446 and the average P_d to be 0.063. The high P_f can be explained by taking into consideration that the hand-contoured studies do not include papillary muscles while the segmentation algorithm does. Also, the epicontour is often set slightly inside the epicardium and the endocontour is often set slightly outside the endocardium in typical hand-contoured study.

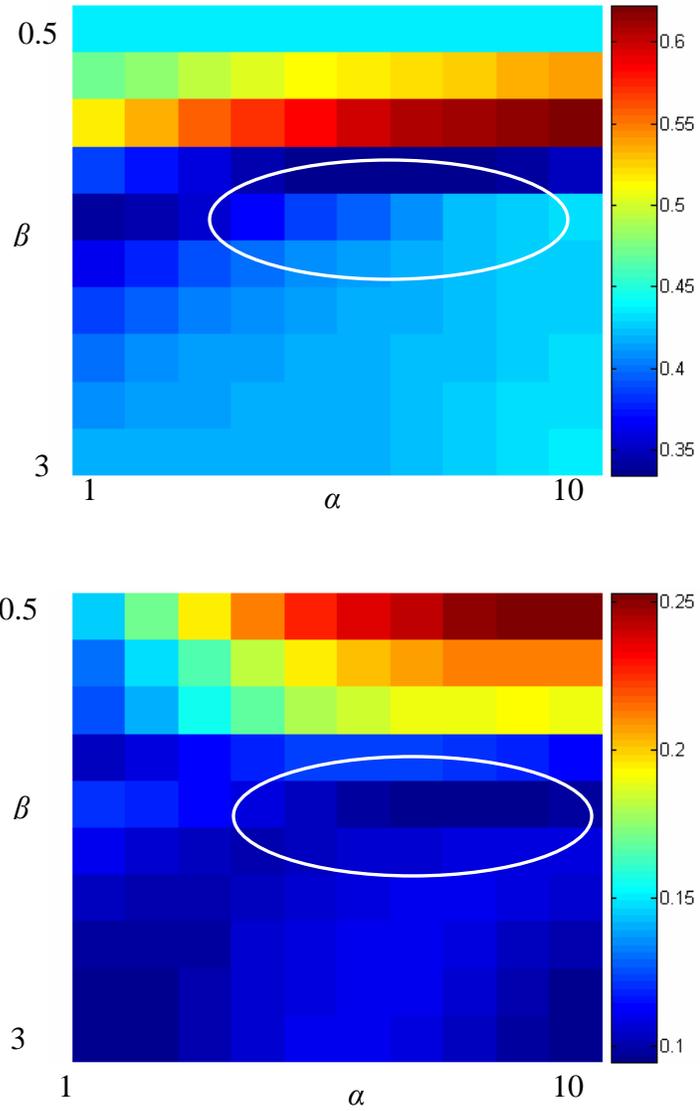


Figure 4.4: Average P_f shown top and average P_d shown bottom. White circles indicate the “optimal” parameters.

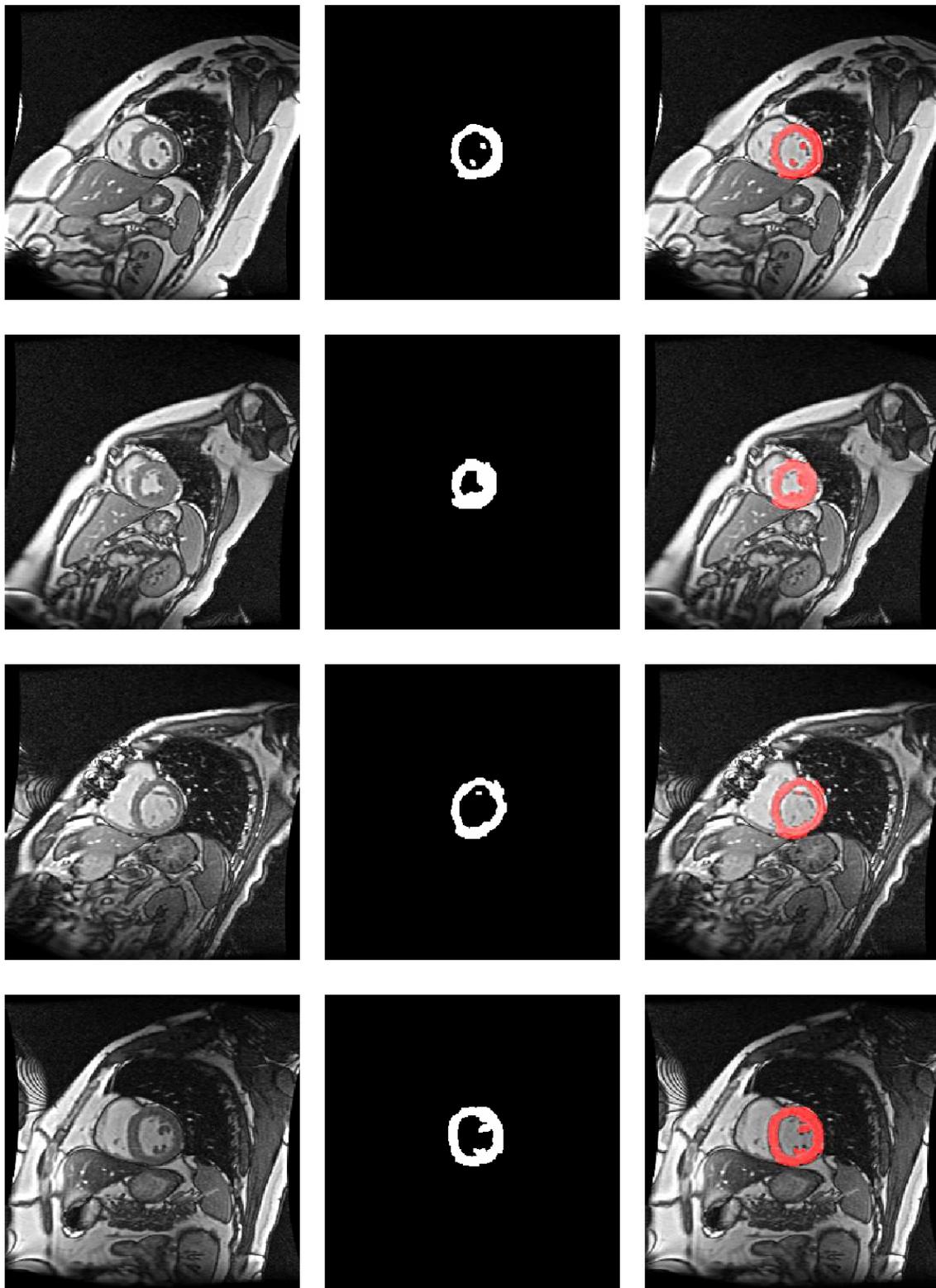


Figure 4.5: Original Images shown left, segmentations shown middle, and segmentations laid over original images and displayed in red shown right.

4.4 Chapter Summary

The use of a multi-channel algorithm for the segmentation of myocardium in cine MRI images is detailed in this chapter. The algorithm is based on two popular texture segmentation algorithms from the literature [3] [10]. This version of the algorithm is supervised in the sense that seed pixels are needed in order to properly segment the image. The algorithm is detailed in Chapter 4.2 including pre-processing, channel characterization, post-processing, and vector classification. Then, an experiment to find the “optimal” parameters for channel characterization (β) and post processing (α) is performed in Chapter 4.3. The optimal parameters were determined to be $\beta = 1.5$ and $\alpha = 1.7$.

Appendix 4.1 Test Studies

Patient ID	Visit	Condition
05DAC006	BAS	Resistant Hypertensive
05DAC033	BAS	Resistant Hypertensive
05DAC040	BAS	Resistant Hypertensive
05DAC048	BAS	Resistant Hypertensive
05DAC055CTL	BAS	Control Hypertensive
05DAC062CTL	BAS	Control Hypertensive
05P1MRR009	BAS	Mitral Valve Regurgitation
05P1VOL001	BAS	Volume Overload
05PVOL507	BAS	Volume Overload
05P3CLN004	BAS	Diabetes w/ Myocardial Infarction

CHAPTER 5

SEGMENTATION OF ASPHALT AGGREGATE FROM MASTIC

5.1 Introduction and Background

In this chapter, the idea of multi-channel segmentation is extended to a non-medical application, namely, the automatic segmentation of aggregate in asphalt images. The structural performance of asphalt pavements is a function of the distribution and type of aggregate, binder, fines, air voids and compaction method [36]. Of particular importance is the need for laboratory compacted asphalt specimens to closely match actual asphalt field samples. This is needed so that the laboratory samples have material properties that are as close to field conditions as possible, yielding valid laboratory experiment results.

Recently, image processing techniques have been used to evaluate the distribution of aggregate, air voids, binder, and fines, as well as to evaluate compaction methods in field extracted and laboratory compacted asphalt samples. The use of such techniques allows for timely and accurate assessment of asphalt mixtures. For instance, Tashman *et al* used X-Ray Computed Tomography (CT) to analyze various asphalt compaction methods [36]. In another paper, Tashman *et al* used CT images to determine air void distribution in asphalt samples [37]. Also, Wang *et al* correlated asphalt mixture performance to the volume fraction of air voids and the gradient of volume fraction of air

voids using CT images of asphalt cores [38].

5.2 Experimental Procedure

Aggregate orientation analysis can be defined as the analysis of the mean orientation characteristics of an asphalt sample, wherein the constituents of the mean characteristics are the individual aggregates in the sample. Orientation analysis is one method of comparing lab asphalt samples to field asphalt samples [39]. The size, shape, and orientation of constituent aggregates in a mix have been shown to affect the local structural properties of asphalt. Moreover, the mean global orientation of constituent aggregates gives a homogeneous measure of the asphalts' material and structural properties and is used here as a measure of the likeness of two asphalt samples. The concern of this analysis is to compare the aggregate orientation of laboratory compacted samples to actual field cores under different compaction thicknesses, or stated explicitly:

Hypothesis: Laboratory samples compacted to about 45 [mm] have an aggregate orientation that more closely matches asphalt cores taken from the field than laboratory specimens compacted to about 115 [mm].

Although the entire experimental procedure is described briefly, the focus of this chapter is on the image processing component of the experiment. The overall experimental procedure for determining aggregate orientation of a given asphalt sample is given in the following paragraphs.

First, cores were obtained from both the field and created in the laboratory. Asphalt samples taken from the field are referred to as field cores, whereas asphalt samples created in a laboratory are referred to as laboratory cores. Both laboratory and

field cores were cylindrical in shape. The field cores ranged in height from 35 mm to 50 mm depending on how the asphalt was originally paved on the roadway. Most field core diameters were 150 mm; however there were a few field cores with a diameter of 100 mm. One laboratory core was compacted to the height of each field core. The remaining the laboratory cores were compacted to a standard height of about 115 mm. Laboratory core diameters were approximately 150 mm.

Next, slices were taken from each core by physically cutting the cores. The slices were taken vertically up through the center of each cylinder at approximately 25 mm spacing. Both sides of each slice were imaged in the next step and labeled side “A” and side “B”.

Next, each slice was imaged with an HP 2200c flatbed scanner. The scanner resolution was set to 200 dpi and the image is quantized at 8 bits grayscale. Additionally, a standard ruler was imaged atop each asphalt core slice to ensure resolution quality control. The photographs were stored digitally in the TIFF format for subsequent image processing. Next, each image was processed for analysis of aggregate orientation. Image processing involved segmentation of aggregate from mastic, labeling individual aggregates, and determining the orientation of individual aggregates. A detailed explanation of the image processing is given in the Section 5.3.

Last, aggregate orientation measures was calculated. Aggregate orientation measures calculated were orientation angle, absolute value of orientation angle, and vector magnitude. Formal definitions of these measures are given in Section 5.4. The mean and standard deviation of orientation measures from all aggregates from each core sample were calculated. These statistics were used to compute a confidence interval for

determining the validity of the original hypothesis.

5.3 Segmentation of Aggregate from Mastic

The focus here is on the image processing involved in segregation of aggregate from mastic in asphalt core slice images. In a typical asphalt core image, the individual aggregates are the rock or granite like material which usually exhibit a higher intensity value than the surrounding mastic. The mastic, or bonding agent, shows up as the darker, in-between, material of the image. Some typical asphalt cores slices are shown in Figure 5.1. The top image of Figure 5.1 is that of a field core taken from a roadway in Alabama, while the bottom image is that of a field core taken from a roadway in Colorado. Notice how the pavement samples are different. The top image exhibits much more aggregate per unit area than does the Colorado sample. Also, the texture of the aggregates varies significantly in each sample. Local granite and/or rock quarries are used when constructing a roadway; therefore, the aggregate texture is a function of the local mineral properties in a given area. These differences make the segmentation task even more challenging.

Images were processed in Matlab, and an automated graphical user interface called ProcessAggregate was developed for this project. The ProcessAggregate program is intended to assist the user through the entire segmentation, labeling, and orientation analysis process. The 8 bit intensity values of the image were represented in MATLAB as double precision values between 0.0 and 1.0.

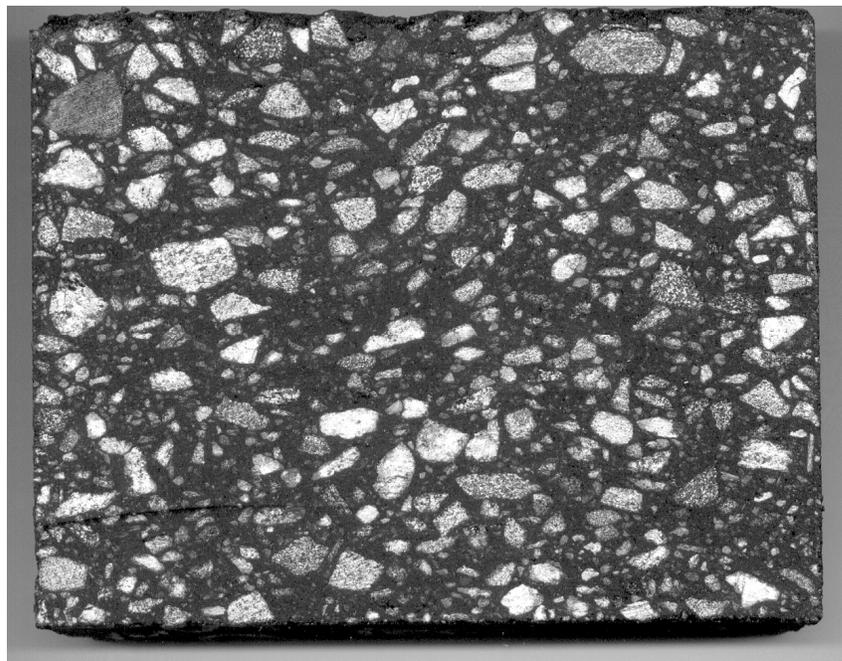
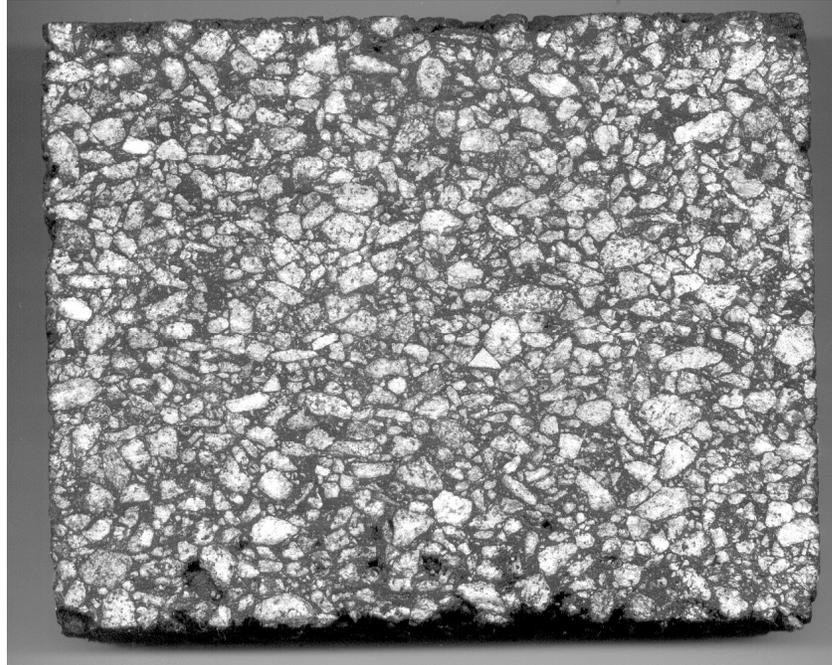


Figure 5.1: Typical asphalt core samples. Field core from Alabama roadway shown top and field from Colorado roadway shown bottom.

5.3.1 Pre-Processing

Prior to an automatic segmentation, two pre-processing tasks were performed. First, the image was cropped to the area of interest. Rather than use an automated routine to perform this task, a simple mouse drag operation was used. Ruler measurements of resolution were determined to be essentially the same as those obtained from the TIFF header file; therefore, subsequent resolution variables were obtained by reading the TIFF headers.

Next, intensity equalization was performed on the image in order to enhance the light and dark regions of the image. The effect of using intensity equalization can be seen by examining the two images of Figure 5.2. Figure 5.2 shows the ProcessAggregate program at the intensity equalization step. The user could set the Eq % value with the slider on the left. The image on the top is the original image while the image on the bottom is the equalized image. Given an Eq value between 0.0 and 1.0, the brightest $\text{Eq} \times 100\%$ of intensity values were mapped to 1.0 and the darkest $\text{Eq} \times 100\%$ of intensity values are mapped to 0.0. The remaining values were linearly scaled between the range 0.0-1.0. In Figure 5.2, one can markedly see the difference in the two images as the aggregate seems to be much brighter and the mastic much darker in the bottom image.

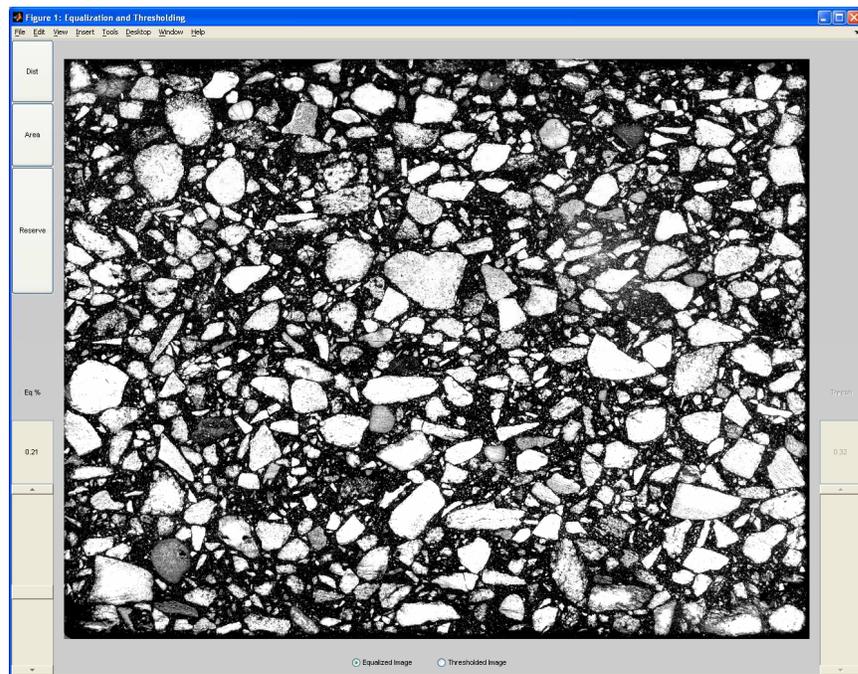
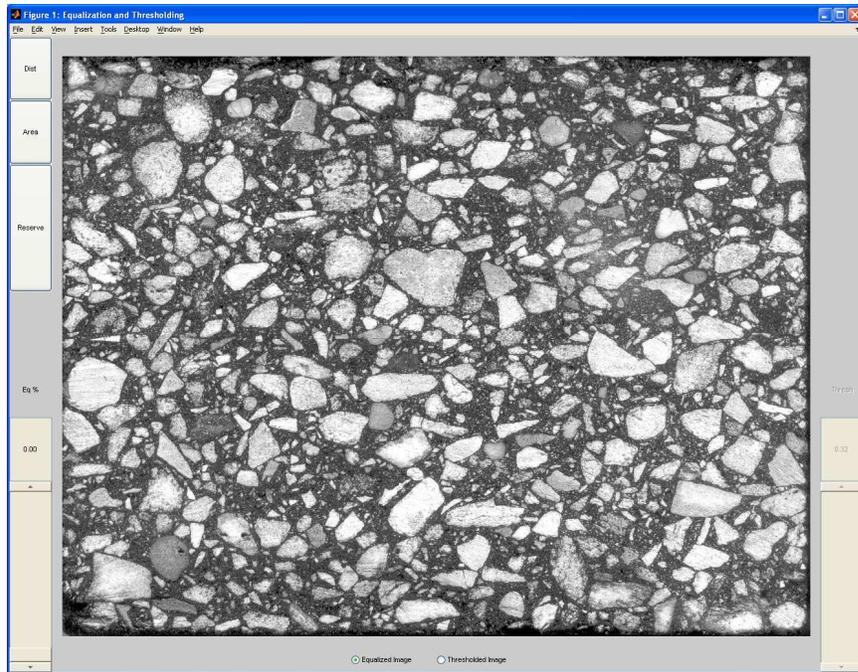


Figure 5.2: Effect of intensity equalization. Original image shown top and intensity equalized image shown bottom

5.3.2 Multi-Channel Segmentation

A multi-channel segmentation was used to segment the mastic from aggregate.

Let $f(\mathbf{p})$ be the intensity-equalized image to be segmented, where

$$\mathbf{p} = \begin{bmatrix} x \\ y \end{bmatrix}$$

is a 2D spatial point. Next, define a bank of symmetric Gabor filters $H(\mathbf{p},i)$ such that

$$H(\mathbf{p},i) = \{h_1(\mathbf{p}), h_2(\mathbf{p}), \dots, h_i(\mathbf{p}), \dots, h_I(\mathbf{p})\},$$

where i is the index of the filter, I is the total number of filters in the bank, and $h(\mathbf{p})$ is the symmetric Gabor filter defined in Chapter 1.4.2. Unlike the segmentation algorithm of Chapter 4, each channel in this algorithm is composed of a single Gabor filter. Due to the finite memory of computers and the relatively high resolution of the images, the size of $H(\mathbf{p},i)$ is limited to about 16 filters. Now, let the i th filter be defined to have the parameters of Table 5.1. Filter parameters of Table 5.1 were determined by trial and error to be most suitable for this particular application. Figure 5.3 shows the -3 [dB] bandwidths of each of the 16 filters from the bank with a white contour. Note that the bank is constructed to give a relatively uniform coverage of the spatial-frequency domain.

A set of response images were obtained by filtering the image with each of the I images from the filter bank $H(\mathbf{p},i)$. The set of response images were referred to as

$$F_r(\mathbf{p},i) = \{f_{r1}(\mathbf{p}), f_{r2}(\mathbf{p}), \dots, f_{ri}(\mathbf{p}), \dots, f_{rI}(\mathbf{p})\},$$

where the i th response image, $f_{ri}(\mathbf{p})$, was calculated as

$$f_{ri}(\mathbf{p}) = f(\mathbf{p}) ** h_i(\mathbf{p}).$$

Figure 5.4 shows an example set of response images.

The next step was to apply a half wave rectification as is done in [3]. Let $F_{abs}(\mathbf{p},i)$ be the set of response images after a half-wave rectification. Then the i th image of $F_{abs}(\mathbf{p},i)$ was denoted as $f_{absi}(\mathbf{p})$ and was calculated as

$$f_{absi}(\mathbf{p}) = \begin{cases} f_{ri}(\mathbf{p}) & , f_{ri}(\mathbf{p}) > 0 \\ 0 & , otherwise \end{cases}.$$

Figure 6.4 shows an example of the 16 images $F_{abs}(\mathbf{p},i)$.

Next, a nonlinearity was applied to each pixel of $F_{abs}(\mathbf{p},i)$ in order to accentuate the response images, as is done in [10]. Prior to applying the nonlinearity, each image of $F_{abs}(\mathbf{p},i)$ was normalized between 0 and 1. Specifically, the i th normalized image was calculated as

$$f'_{absi}(\mathbf{p}) = \frac{f_{absi}(\mathbf{p})}{\max(f_{absi}(\mathbf{p}))}.$$

The set of images obtained by application of the nonlinearity is denoted as $F_n(\mathbf{p},i)$, where the i th image was denoted as $f_{ni}(\mathbf{p})$ and was calculated as

$$f_{ni}(\mathbf{p}) = \tanh(\alpha \cdot f'_{absi}(\mathbf{p})),$$

where α is a constant that controls the extent of the nonlinearity. Here, α was determined by trial and error to be 10. Figure 5.6 shows an example $F_n(\mathbf{p},i)$.

Next, the average response over an image region was calculated for each of the images of $F_n(\mathbf{p},i)$. A Gaussian averaging window was used with the standard deviation of each window calculated as $\sigma_i = 0.4/f_i$, where f_i is the center frequency of the Gabor filter used in the i th channel. The set of averaged images was denoted as $F_f(\mathbf{p},i)$, and was calculated as

$$f_{f_i}(\mathbf{p}) = g_i(\mathbf{p}) ** f_{n_i}(\mathbf{p}),$$

where

$$g_i(\mathbf{p}) = \exp\{-0.5(\mathbf{p}^T \mathbf{S} \mathbf{p})\},$$

and where

$$\mathbf{S} = \begin{bmatrix} 1/\sigma_i & 0 \\ 0 & 1/\sigma_i \end{bmatrix}.$$

Figure 5.7 shows an example $F_f(\mathbf{p}, i)$. Note that $F_f(\mathbf{p}, i)$ is the final set of feature images.

Table 5.1: Filter bank parameters of $H(\mathbf{p}, i)$

i	F_i (pixels⁻¹)	θ_i (degrees)	σ_{xi}	σ_{yi}
1	0.1	0	1.7 x F _i	3 x F _i
2	0.1	45	1.7 x F _i	3 x F _i
3	0.1	90	1.7 x F _i	3 x F _i
4	0.1	135 (-45)	1.7 x F _i	3 x F _i
5	0.1667	0	1.7 x F _i	3 x F _i
6	0.1667	45	1.7 x F _i	3 x F _i
7	0.1667	90	1.7 x F _i	3 x F _i
8	0.1667	135 (-45)	1.7 x F _i	3 x F _i
9	0.2333	0	1.7 x F _i	3 x F _i
10	0.2333	45	1.7 x F _i	3 x F _i
11	0.2333	90	1.7 x F _i	3 x F _i
12	0.2333	135 (-45)	1.7 x F _i	3 x F _i
13	0.3	0	1.7 x F _i	3 x F _i
14	0.3	45	1.7 x F _i	3 x F _i
15	0.3	90	1.7 x F _i	3 x F _i
16	0.3	135 (-45)	1.7 x F _i	3 x F _i

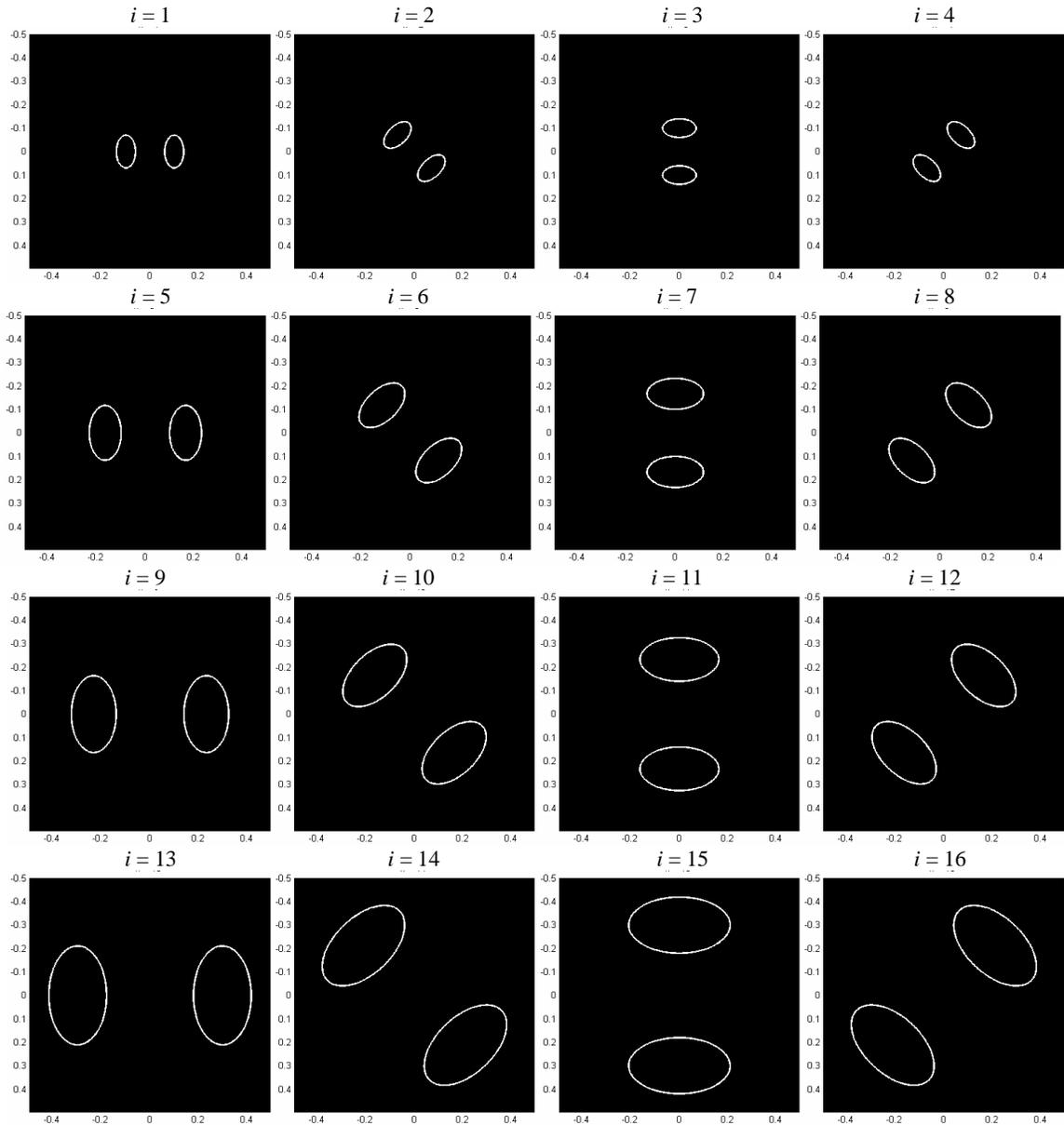


Figure 5.3: Spatial frequency contour plot of Gabor filter bank used in this segmentation. White contours indicate the -3db bandwidth of each filter. The vertical axis of each plot represents the spatial frequency in the vertical dimension in units of cycles/sample and horizontal axis represents the spatial frequency in the horizontal dimension in units of cycles/sample.

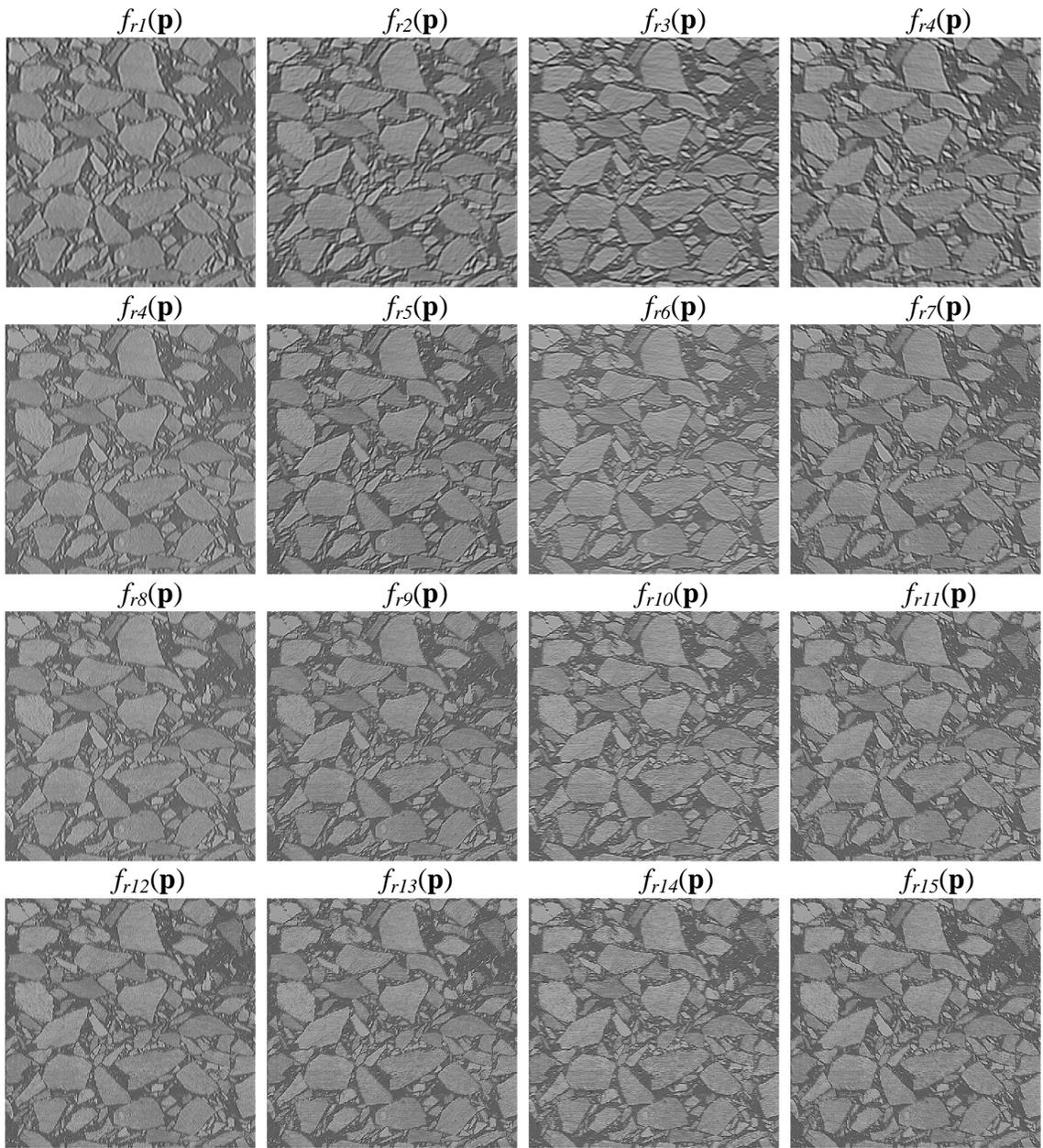


Figure 5.4: Response images, $F_i(\mathbf{p}, i)$

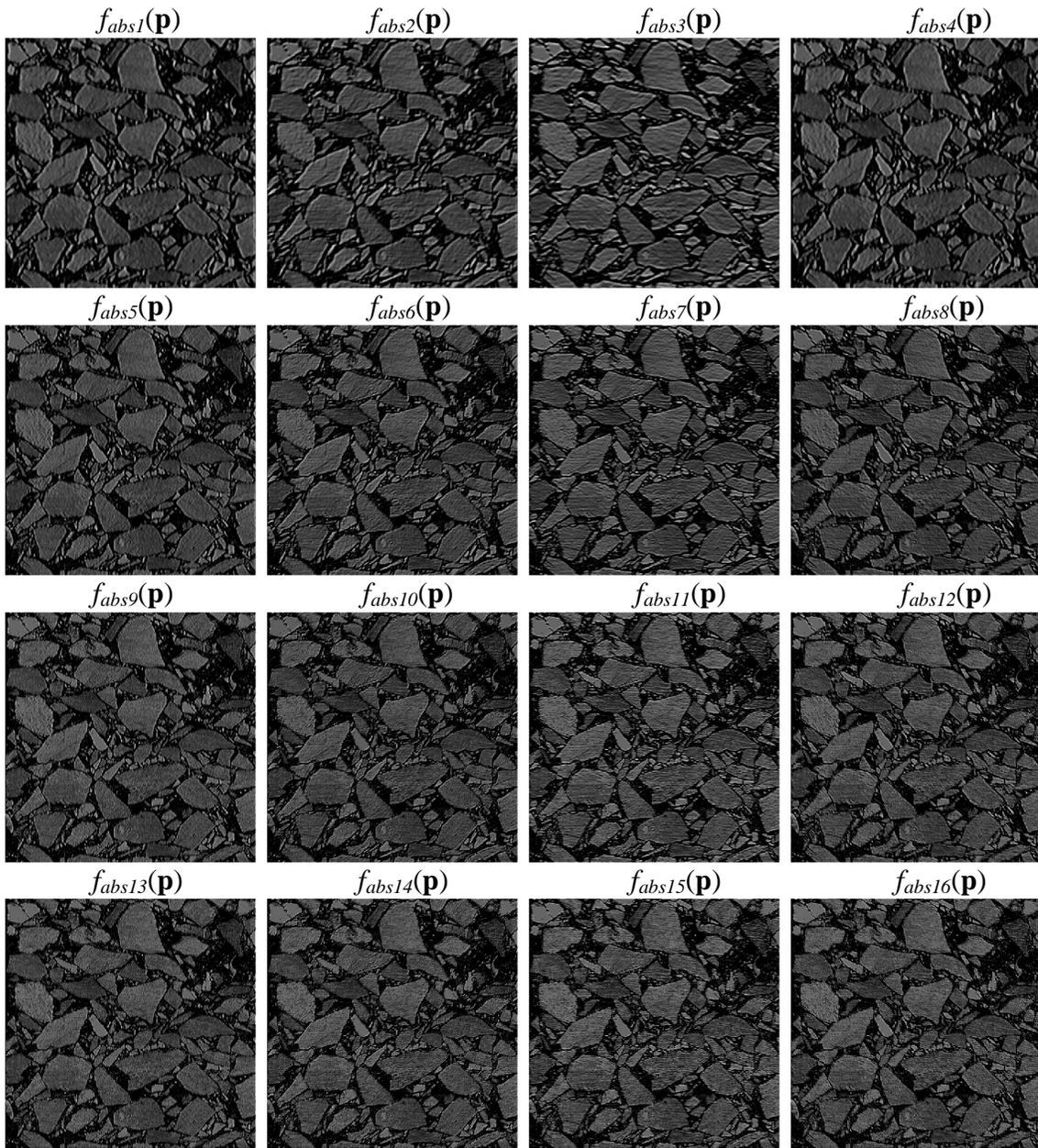


Figure 5.5: Half wave rectified images, $F_{abs}(\mathbf{p}, i)$

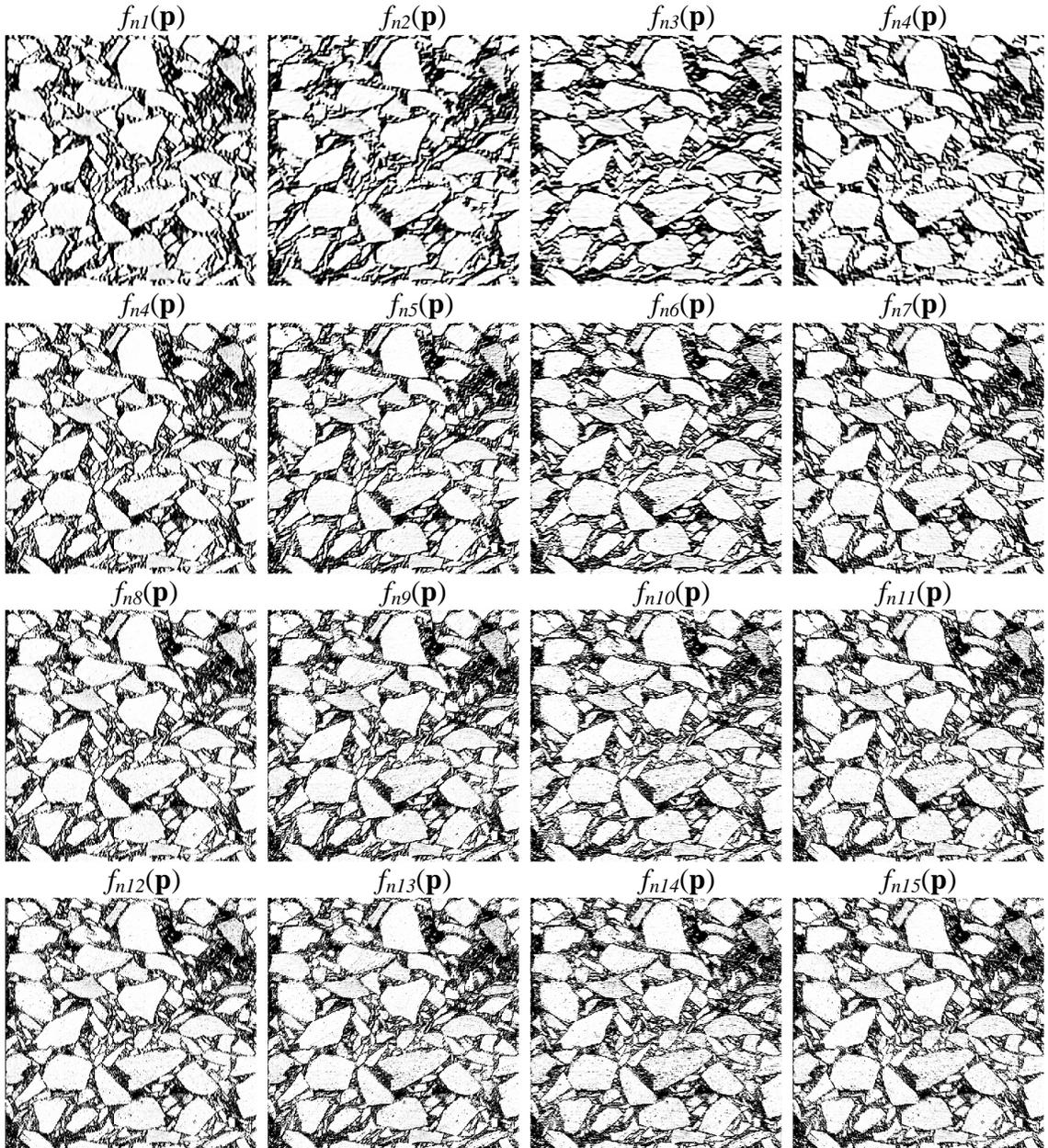


Figure 5.6: Hyperbolic tangent nonlinearity step $F_n(\mathbf{p},i)$

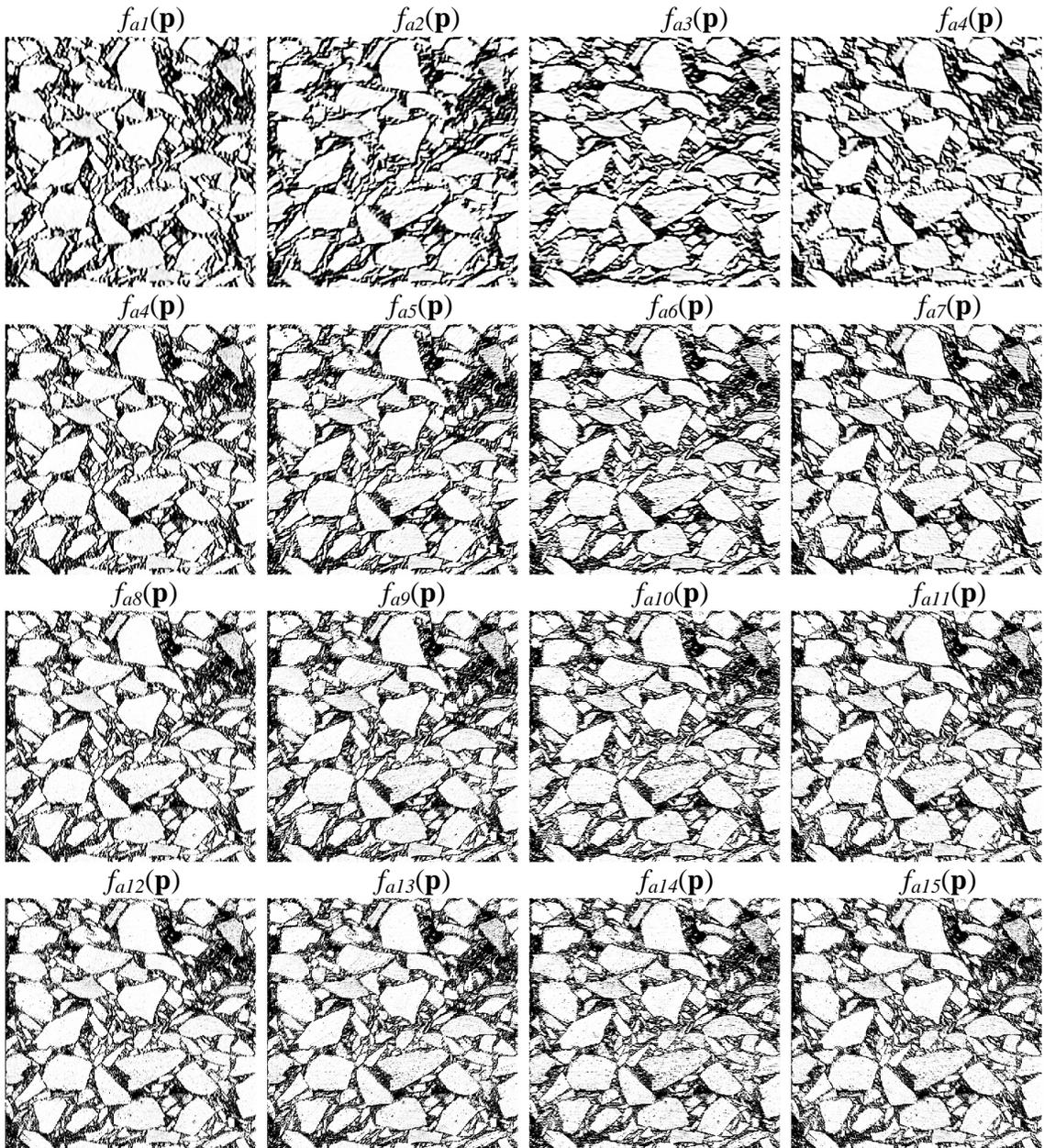


Figure 5.7: Window Averaging Step, $F_a(\mathbf{p}, i)$

Last, the channels were grouped via the k-means clustering algorithm [35]. Prior to clustering, the feature images were reconfigured into feature vectors. Construction of feature vectors from a set of feature images is detailed in Chapter 3.2; therefore it is not revisited here.

Now that the feature images have been reconfigured into feature vectors, the k-means clustering algorithm was applied to the entire set of feature vectors. The vectors were classified as belonging to one of two groups; aggregate or mastic. Figure 5.8 shows a final binary image, $f_b(\mathbf{p})$ where the aggregate is shown as +1 (white) and the mastic is shown as 0 (black). Figure 5.8 also places the original image next to the binary image for comparison purposes.

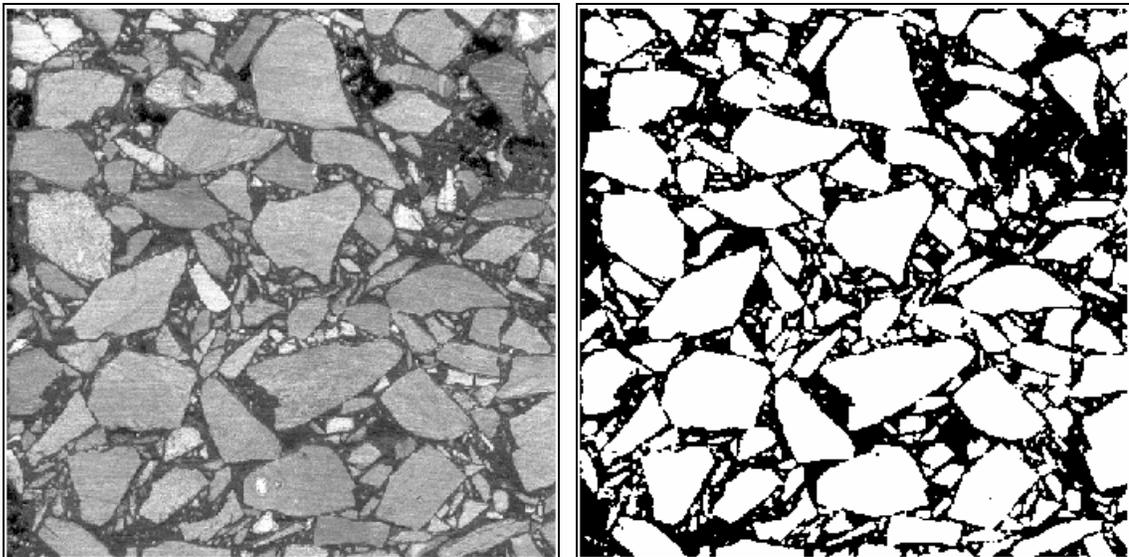


Figure 5.8: Example original image (left) and segmented image (right)

5.3.3 Post-Processing

The next step was to label each connected region of $f_b(\mathbf{p})$ with an integer value, starting at 1 and going up to the number of connected regions of the image. This task was made simple with the built-in Matlab function `bwlabel` (). Figure 5.9 shows an image labeled in such a manner. Each color represents a different connected region. Also, note that the image is displayed with a random colormap for readability.

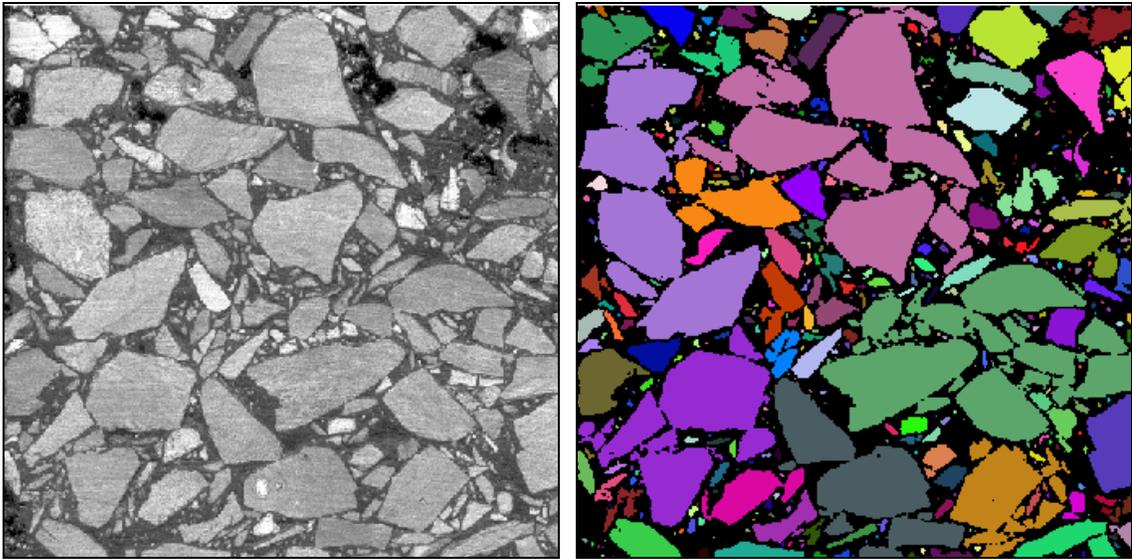


Figure 5.9: Example original image (left) and labeled image (right)

Notice in Figure 5.9 that some regions are connected when in fact it is clear that they should be separated. An automated region splitting routine was used to split most of these regions by morphologically eroding each connected region twice. This disconnected any regions that were only connected by a few pixels by removing any pixel at the perimeter of a connected region. After the two erosions, any new connected regions were assigned a unique label. Each region was then morphologically dilated

twice in order to restore the image to approximately its original size. Morphological dilation is the opposite of morphological erosion as it adds pixels around a connected region's perimeter. Finally, the original shape of each aggregate was retained by taking the logical AND of each labeled region with the binary image $f_b(\mathbf{p})$. The image obtained by regions splitting was denoted as $f_s(\mathbf{p})$. Figure 5.10 shows an example $f_s(\mathbf{p})$.

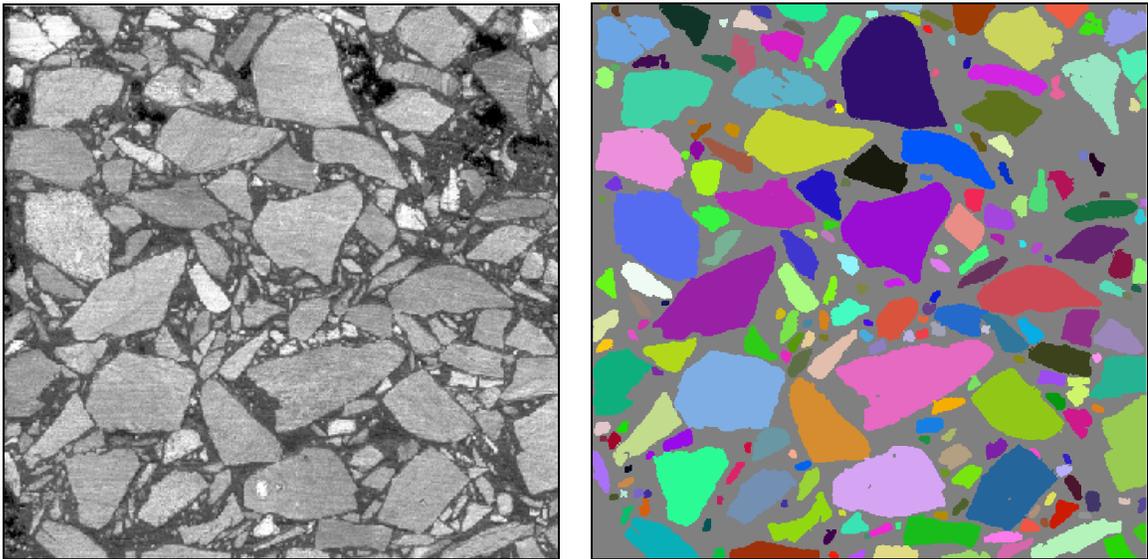


Figure 5.10: Example original image (left) and labeled image after region splitting routine (right)

After the region splitting routine was applied, any other falsely joined or falsely split regions were split or joined with the ProcessAggregate program, which has manual CAD tools to join and split regions of interest. The final step was to clean the image by removing labeled regions with an area less than a pre-determined threshold. A labeled region's area was determined by the number of pixels of the labeled regions times the image resolution. The final labeled image with constituent aggregates labeled was called

$f_A(\mathbf{p})$. Figure 5.11 shows an example $f_A(\mathbf{p})$.

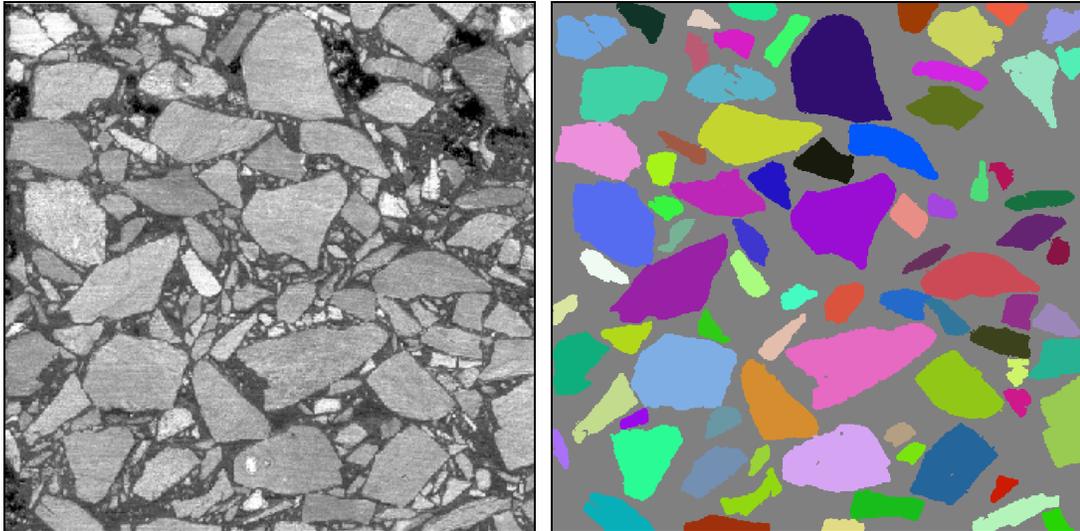


Figure 5.11: Example original image (left) and labeled image after region splitting routine and cleaning operation (right)

5.4 Aggregate Orientation Analysis

Once the individual aggregates were identified, an aggregate orientation analysis was performed. By comparing the mean orientation characteristics of different laboratory samples, one can determine whether a significant statistical difference in aggregate orientation exists between the samples.

5.4.1 Aggregate Orientation Analysis Definitions

The aggregate orientation analysis began by fitting an ellipse to each individual aggregate found by the segmentation. Each ellipse was fit to have the same normalized second central moment given by the equations:

$$c_{xx} = \frac{1}{Area_{ellipse}} \iint (x - c_x)^2 dx dy$$

$$c_{xy} = \frac{1}{Area_{ellipse}} \iint (x - c_x)(x - c_y) dx dy$$

$$c_{yy} = \frac{1}{Area_{ellipse}} \iint (x - c_y)^2 dx dy$$

Properties defined for each labeled region are:

Area = area of aggregate in [mm²] \approx # pixels in this aggregate x pixel size

Equivalent Diameter = diameter of a circle with same area as this aggregate =

$$(4Area/\pi)^{1/2}.$$

Orientation Angle (θ) = Angle b/n horizontal axis of image and major axis of the ellipse.

Ranges from -90 to 90 degrees.

Major Axis Length = Length of major axis of ellipse in [mm].

Minor Axis Length = Length of minor axis of ellipse in [mm].

Eccentricity = Ratio of the distance b/n the foci of the ellipse and its major axis length.

Ranges between 1 and 0 with 0 being a perfect circle and 1 being a straight line of infinite length.

The above measures were used to calculate the mean and standard deviation of the orientation angle, the absolute value of the orientation angle and the vector magnitude was computed. The vector magnitude was defined as:

$$VectorMagnitude = VM = \frac{100}{N} \sqrt{\left(\sum_{k=1}^N \sin 2\theta_k\right)^2 + \left(\sum_{k=1}^N \cos 2\theta_k\right)^2} \%,$$

where N is the number of aggregates segmented in the image, and θ_k is the orientation

angle of the k th aggregate.

5.4.2 Aggregate Orientation Analysis Results

Here, a set of example results are shown and interpreted. For this example, a field core from Georgia (called ga1core1) was selected and compared with a lab core compacted to a thickness of 115 mm (called ga1sgc115) and 45 mm (called ga1sgc45). The ProcessAggregate program was used to perform an automatic segmentation as described in this chapter. The results of the aggregate orientation analysis obtained from the segmentation are given in Figure 5.12. Figure 5.12 shows two plots. The first is a plot of the average absolute angle versus eccentricity for the three samples. The second is a plot of vector magnitude versus eccentricity for three samples. Both plots show that there is no significant difference in field cores and lab cores for eccentricities below 80 percent as measured by average absolute angle and vector magnitude. However, as the eccentricity of the aggregate becomes greater than 80 percent, the 115 mm thick asphalt cores are shown to be more like field cores than the 45 mm cores.

5.5 Chapter Summary

In this chapter, images of asphalt cores were imaged and segmented using a multi-channel segmentation algorithm. The purpose of the algorithm is to segment the asphalt binder from the aggregate in the core image samples to subsequently perform an aggregate orientation analysis. The experimental procedure including core sample acquisition and imaging is explained in Chapter 5.2. The image processing procedure to perform the multi-channel segmentation is detailed in Chapter 5.3. Also included in this

section are the post-processing steps used to obtain labeled aggregate from a segmented image. Last, the process of aggregate orientation analysis with example results from sample asphalt cores using the ProcessAggregate program is explained in Chapter 5.4.

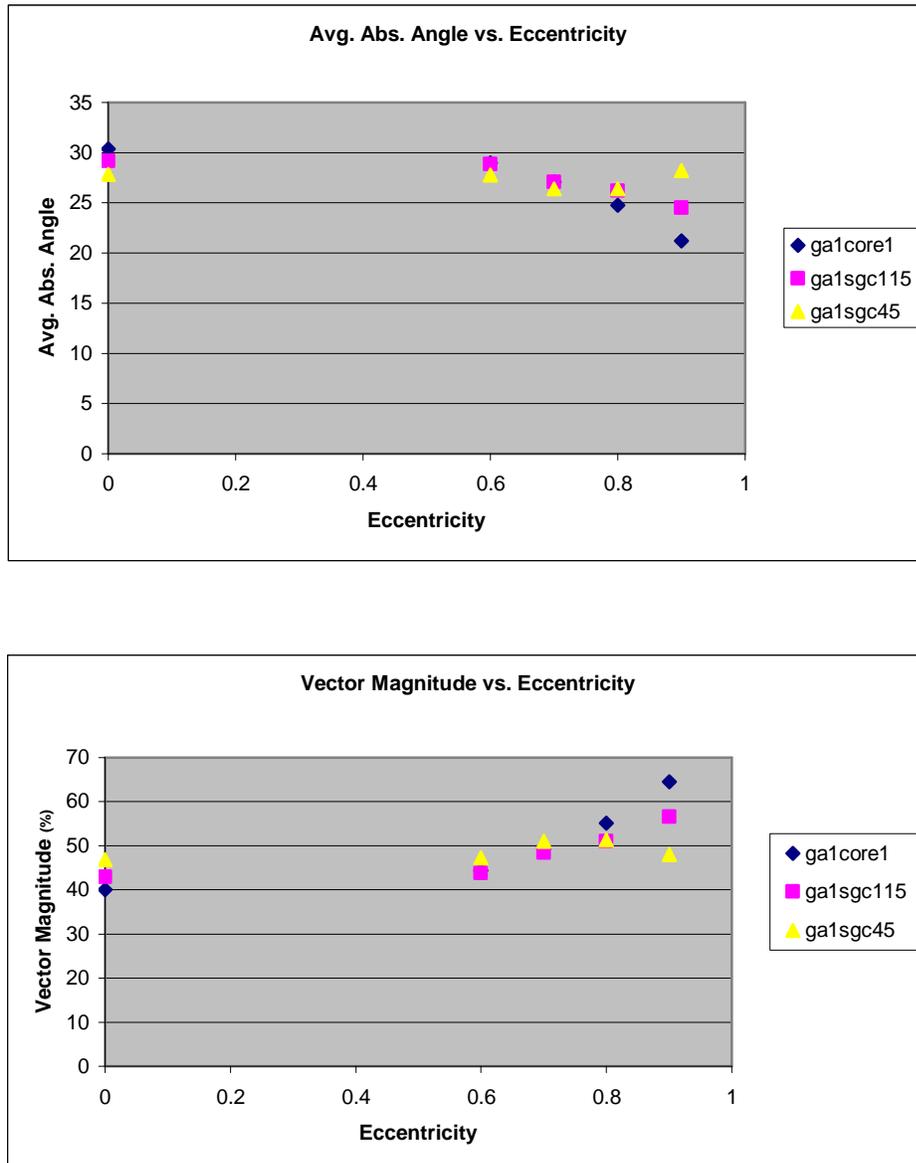


Figure 5.12: Plots of average absolute angle vs. eccentricity and vector magnitude vs. eccentricity

CHAPTER 6

CONCLUSION

6.1 Thesis Summary

In this thesis the use of multi-channel processing has been explored with specific applications to filtering of tagged cardiac MRI images and segmentation of tagged cardiac MRI images, cine cardiac MRI images, and asphalt core images. Several models for spectral changes due to tag line deformation in tagged cardiac MRI images were presented in Chapter 2. The models were used to motivate the design of a C-Gabor filter pre-filter intended to be used in conjunction with a tag-tracking algorithm. The C-Gabor filter was tested against two commonly used filters, namely, the standard Gabor filter and the HARP filter. The C-Gabor filter was shown to be superior to the other two filters tested. The tests also determined optimal parameters for each filter for this particular use.

Two tag line identification algorithms to be used in conjunction with a statistical tag line classification algorithm were presented in Chapter 3. The first method was based on a multi-channel segmentation where each channel is tuned to the tag frequency and the first harmonics of the tag frequency. It was shown to be effective; however, sub-pixel refinement is not demonstrated. The second method presented in Chapter 3 was based on pre-filtering the tagged image with a C-Gabor filter and subsequently performing a peak detection in the direction normal to the tags. The QuickTag algorithm was detailed in

this chapter. QuickTag was shown to be effective for tag line identification and allows sub-pixel refinement.

A supervised multi-channel segmentation algorithm based on two popular literature publications was presented in Chapter 4. Pre-processing, channel characterization, and post-processing were explained. The algorithm was tested for optimal filter bank parameters by comparing the segmentation to hand segmented images.

A final application of multi-channel filter banks was provided in Chapter 5. A detailed an unsupervised segmentation algorithm for segmentation of asphalt pavement images was detailed in this chapter. The purpose of developing the algorithm was to help perform an aggregate orientation analysis intended to determine the difference between laboratory asphalt cores and field asphalt cores. Preliminary results were given at the end of the chapter showing a difference in field and laboratory cores at eccentricities above eighty percent.

6.2 Directions for Future Work

This thesis has explored the use of multiple channels tuned to different spatial frequencies to process and extract information from images. In particular, channel characterization was a primary focus and the bulk of the work. Future research into this area should focus on channel combination as opposed to channel characterization. For instance, the segmentation algorithms of Chapter 3 and 5 used simple k-means clustering to combine image features into a final segmented image. Other methods such as neural networks and fuzzy clustering should be explored. Another direction for further research would include the combination of multi-channel filter methods with other segmentation

methods to achieve better segmentations. For instance, the myocardial segmentation of Chapter 4 could very well be combined with an active contour method to achieve a better segmentation.

REFERENCES

- [1] Bovik, A. C., and Clark, Marianna, "Multichannel Texture Analysis Using Localized Spatial Filters," *IEEE Trans. On Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, Jan. 1990, pp. 55-73
- [2] Cerqueira, M.D., Weissman, N.J., Vasken, D., Jacobs, A.K., Kaul, S., Laskey, W.K., Pennell, D.J., Rumberger, J.A., Ryan, T., and Verani, M.S., "Standardized Myocardial Segmentation and Nomenclature for Tomographic Imaging of the Heart," *AHA Scientific Statement*, 2002
- [3] Tan, T. N., and Constantinides, A. G., "Texture Analysis Based on A Human Visual Model"; *Acoustics, Speech, and Signal Processing, 1990. ICASSP-90, 1990 International Conference on* 3-6 April 1990 pp: 2137-2140 vol. 4
- [4] Gabor, Dennis, "Theory of Communication," *J. IEE London* 93, 429-457 (1946)
- [5] Daughman, J. G., "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters," *J. Opt. Soc. Am. A*, vol. 2, No. 7, pp. 1160-1169, (1985)
- [6] D.G. Hubel and T.N Wiesel, "Receptive fields, binocular interaction, and functional architecture in the Cat's Visual Cortex," *J. Physiology London* 160, 106-154 (1962)
- [7] Marcelja, S., "Mathematical Description of the Response of Simple Cortical Cells," *J. Opt. Soc. Am.*, vol. 70 , no. 11, November 1980
- [8] Daughman, J. G., "Spatial Visual Channels in the Fourier Plane," *Vision Res.*, vol. 24, no. 9, pp 891-910, 1984
- [9] Daughman, John G., "Two-Dimensional Spectral Analysis of Cortical Receptive Field Profiles," *Vision Research* Vol. 20, pp. 847-856
- [10] Jain, A. K., and Farrokhnia, F., "Unsupervised Segmentation of Textured Images Using Gabor Filters," *1990 Conference Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*; 4-7 Nov. 1990, pp. 14-19
- [11] Perry, A.; Lowe, D.G., "Segmentation of Textured Images," *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR '89.*, IEEE Computer Society Conference on, 4-8 June 1989 Page(s):319 - 325
- [12] Kameyama, K.; Taga, K., "Texture Classification by Support Vector Machines with Kernels for Higher-Order Gabor Filtering," *Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on, Volume 4, 25-29 July 2004* Page(s):3009 - 3014
- [13] Kachouie, N.N.; Alirezaie, J., "Texture Segmentation using Gabor Filter Banks and Multi-Layer Perceptron," *Systems, Man and Cybernetics, 2003. IEEE International Conference on, Volume 3, 5-8 Oct. 2003* Page(s):2897 - 2902
- [14] Mital, D.P., "Texture Segmentation Using Gabor Filters," *Knowledge-Based*

- Intelligent Engineering Systems and Allied Technologies, 2000. Proceedings. Fourth International Conference on, Volume 1, 30 Aug.-1 Sept. 2000 Page(s):109 - 112
- [15] Kumar, A., Pang, G., “Defect Detection in Textured Materials Using Gabor Filters,” IEEE Transactions on Industry Applications, vol. 38, issue 2, Mar.- Apr. 2002, pp. 425-440
- [16] Garcia-Consuegra, J., Cinerors, G., Ballesteros, J., and Monila, R., “Remote Sensing Segmentation Through A Filter Bank Based on Gabor Functions,” Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 2, May 12-15, 1998, pp. 1169-1171
- [17] Mangko, T. R., and Pramudito, J. T., “Implementation of Gabor Filter To Texture Analysis of Radiographs in the Assessment of Osteoporosis,” 2002 Asia-Pacific Conference on Circuits and Systems, vol.2, Oct. 28-31, 2002, pp. 251-254
- [18] Montillo, A., Metaxas, D., and Axel, L., “Extracting Tissue Deformation Using Gabor Filter Banks,” Proceedings of SPIE: Medical Imaging: Physiology, Function, and Structure from Medical Images, Vol. 5369, p. 1-9; Amir A. Amini, Armando Manduca; Eds., San Diego, CA, Vol. 5369 pp 1-9 (2004)
- [19] Qian, Z; Montillo, A.; Metaxas, D.N.; Axel, L., “Segmenting Cardiac MRI Tagging Lines Using Gabor Filter Banks,” Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE, Volume 1, 17-21 Sept. 2003 Page(s):630 - 633
- [20] Hamlick, R.M., K. Shanmugam, and I. Dinstein, “ Texture Features for Image Classification,” IEEE Transactions on Systems, Man and Cybernetics, SM-3, pp. 610-621, 1973
- [21] Chen, C. H., Pau, L. F., and Wang, P. S., “The Handbook of Pattern Recognition and Computer Vision,” 2nd Edition,, pp. 207-248, World Scientific Publishing Co., 1998.
- [22] Soh, L. K., and Tsatsoulis, C., “Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices,” IEEE Transactions on Geoscience and Remot Sensing, Volume 37, Issue 2, Part 1, March 1999, pp. 780-795
- [23] Soo Chin Liew; Hock Lim; Leong Keong Kwoh; Geok Kee Tay; Geoscience and Remote Sensing Symposium, 1995. IGARSS '95. 'Quantitative Remote Sensing for Science and Applications', International Volume 2, 10-14 July 1995 Page(s):1412 - 1414 vol.2
- [24] Lopez, L.; Moctezuma, M.; Parmiggiani, F.; Geoscience and Remote Sensing Symposium, 2005. IGARSS '05. Proceedings. 2005 IEEE International Volume 3, 25-29 July 2005 Page(s):1781 – 1784
- [25] Lee, Jin, Denney, Thomas S., Davis, Craig A., “Tag Point Classification in Tagged Cardiac MRI Images,” submission to IEEE International Symposium on Biomedical Imaging, Arlington VA, April 6-9
- [26] Waks, E., Prince, JL., and Douglas, AS., “Cardiac Motions Simulator for Tagged MRI,” in *Proc. Workshop Meathematical Methods in Medical Image Analysis*, 1996, pp. 182-191.
- [27] Bracewell, Ronald N., *Two-Dimensional Imaging* Prentice Hall, Englewood Cliffs, New Jersey 07632, 1995
- [28] Jones, R. T., “Annual Reviews of Fluid Kinematics,” Annual Reviews, Palo Alto,

- CA. 1969.
- [29] Denney, Thomas S., Prince, Jerry L., “Optimal Brightness Functions for Optical Flow Estimation of Deformable Motion,” *IEEE Transactions on Image Processing*, Vol 3, No. 2, March 1994
 - [30] Osman, Nael F., McVeigh, Elliot R. , and Prince, Jerry L., “Imaging Heart Motion Using Harmonic Phase MRI,” *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, March 2000
 - [31] Denney, TS, Gerber, BL, and Yan, L, “Unsupervised Reconstruction of a Three-Dimensional Left Ventricular Strain From Parallel Tagged Cardiac Images,” *Magn Reson Med* 49:743-754, 2003
 - [32] Atalar, E. and McVeigh, E.R., “Optimization of tag thickness for measuring position with magnetic resonance imaging,” *IEEE Trans Img Proc*, Vol 13, No. 1, March 1994
 - [33] Li, Jin, “Tag Line Tracking and Cardiac Motion Modeling from Tagged MRI,” Ph.D. Dissertation Proposal, Auburn University, 2006
 - [34] Qian, Z, Montillo, A., Metaxas, D., Axel, L., “Segmenting Cardiac MRI Tagging Lines Using Gabor Filter Banks,” *Engineering in Medicine and Biology Society, 1003 Proceedings of the 25th Annual International Conference of the IEEE*, vol. 1, 17-21 Sep., 2003, pp 630-633
 - [35] MacQueen, J. B., “Some Methods for Classification and Analysis of Multivariate Observations,” *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*”, Berkely, University of California Press, pp281-297
 - [36] Tashman, L., Masad, E., Peterson, B., and Saleh, H., “Internal Structural Analysis of Asphalt Mixes to Improve the Simulation of Superpave Gyratory Compaction to Field Conditions,” *Journal of Asphalt and Paving Technologists*, vol. 70 , 2001, pp. 605-645
 - [37] Tashman, L., Masad, E., Angelo, J. D., Bukowski, J. and Harman, T., “X-ray Tomography to Characterize Air Void Distribution in Superpave Gyratory Compacted Specimens,” *International Journal of Pavement Engineering*, vol. 3 (1), pp. 19-28, 2002
 - [38] Wang, Y., Mohammed, L., Harman, T. “Voids Distribution and Performance of Asphalt Concrete,” *IJP* 2002 vol. 1 num. 3, Sep. 2002, pp 22-33
 - [39] Nasad, E., Muhunthan, B., Shashidhar, N., and Harman, T., “Aggregate Orientation and Segregation in Asphalt Concrete,” *Geotechnical Special Publication*, n 85, *Application of Geotechnical Principles in Pavement Engineering*, 1998, pp 69-80