



Target Structure Enhancement of Cardiac Magnetic Resonance Images (MRI).

A study on the Rician Noise Model of Magnitude MRI

**Erich Leung**

**John Tsotsos**

Technical Report CSE-2009-05

Jun 28 2009

Department of Computer Science and Engineering  
4700 Keele Street Toronto, Ontario M3J 1P3 Canada

## Abstract

This report presents a computational framework for the enhancement of magnitude magnetic resonance (MR) images of the heart. It addresses two sources of degradation of image quality from the perspective of an algorithmic procedure that seeks to recover the anatomical structures of the organ of interest: (1) the effect of image noise and other confounding features that obscure the relevant information of these structures and (2) inadequate image contrast. An analysis of the signal properties of these confounding features leads to the wavelet-based approach presented in this report. An enhancement framework, that is adaptive to both image structures and signal intensity, is developed to attenuate image noise, to remove the confounding features present in the low-intensity regions and to enhance image contrast. The framework is implemented and evaluated on both synthetic and medical test cases with strong supporting evidence for the hypothesis of this report. The experimental findings also demonstrate that the scope of application of the framework can be extended to the class of non-cardiac images.

# Contents

<b>Abstract</b>	<b>iv</b>
<b>List of Acronyms</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 The Problem of Adaptive Enhancement of Cardiac Magnetic Resonance (MR) Images . . . . .	2
1.2 Background and Related Research . . . . .	6
1.3 Contribution . . . . .	13
1.4 Summary . . . . .	17
<b>2 Computational Framework</b>	<b>18</b>
2.1 Signal Representation in the Wavelet Domain . . . . .	21
2.2 Signal Properties of Confounding Features in Magnitude MR Images	22
2.2.1 Noise Behavior in Magnitude MR Images . . . . .	23
2.2.1.1 Rician Noise Model of Image Signals . . . . .	24
2.2.1.2 Convergence Range of Rician Signals . . . . .	26
2.2.2 Signal Properties of Squared Magnitude Representations	33
2.3 Smoothing-Enhancement (SE) Approach to Adaptive Enhancement of MMR images . . . . .	41
2.3.1 Smoothing Step: Attenuation of Fine Scale Signal Oscillations due to Gaussian-like Noise . . . . .	42
2.3.1.1 SU Decision Rule: Wavelet Shrinkage with the Universal Threshold . . . . .	43
2.3.1.2 MSE Decision Rule: Wavelet Shrinkage with Minimum Mean Squared Error (MSE) . . . . .	46
2.3.1.3 SSR Decision Rule: Semisoft Shrinkage with Rician based Thresholds . . . . .	48
2.3.2 Enhancement Step: Adaptive Coarse Scale Signal Enhancement . . . . .	50
2.4 Summary . . . . .	57

<b>3</b>	<b>Experimental Results</b>	<b>59</b>
3.1	Methodology of Performance Evaluation . . . . .	59
3.2	Performance Evaluation . . . . .	66
3.2.1	Synthetic Images . . . . .	67
3.2.2	Clinical Magnitude MR images . . . . .	72
3.3	Summary . . . . .	84
<b>4</b>	<b>Conclusion</b>	<b>85</b>
4.1	Discussion . . . . .	85
4.2	Future Work . . . . .	88

## Appendices 91

<b>I</b>	<b>Technical Notes</b>	<b>92</b>
<b>A</b>	<b>Mechanics of Wavelet Analysis</b>	<b>93</b>
<b>B</b>	<b>Statistical Properties of Rician Noise</b>	<b>103</b>
B.1	Rician Noise in Signal Free Regions . . . . .	104
B.2	Rician Noise in Infinity Signal Regions . . . . .	104
<b>C</b>	<b>Noise Parameter <math>\sigma</math>: Estimation and Relevant Range</b>	<b>107</b>
C.1	Estimation . . . . .	107
C.2	Relevant Range . . . . .	109
<b>D</b>	<b>Moments, Cumulants and Convergence Analysis of Rician Noise</b>	<b>110</b>
D.1	Statistical Moments and Cumulants . . . . .	110
D.2	Moments and Cumulants of the Gaussian Family . . . . .	112
D.3	Moments and Cumulants of the Rician Family . . . . .	116
D.4	Convergence Analysis of Rician Noise . . . . .	119
<b>E</b>	<b>Wavelet Shrinkage and Noise Attenuation</b>	<b>127</b>
E.1	Introduction . . . . .	127
E.2	SU Decision Rule: Wavelet Shrinkage with the Universal Threshold . . . . .	130
E.3	MSE Decision Rule: Wavelet Shrinkage with Minimum Mean Squared Error (MSE) . . . . .	133
E.4	SSR Decision Rule: Semisoft Shrinkage with Rician based Thresholds . . . . .	136
<b>F</b>	<b><math>\chi^2</math> Noise in Squared Magnitude Images</b>	<b>139</b>

<b>II</b>	<b>Output Data Sets of Performance Evaluation</b>	<b>141</b>
<b>G</b>	<b>Synthetic Test Cases</b>	<b>143</b>
<b>H</b>	<b>Medical Test Cases</b>	<b>146</b>

# List of Tables

3.1	Enhancement Procedures Implemented for Performance Evaluation . . . . .	60
3.2	Selected Test Cases for Performance Evaluation . . . . .	72
3.3	Performance Evaluation: Improvement in Image Contrast. . . . .	73

# List of Figures

1.1	A Cardiac Magnitude MR Image. . . . .	4
2.1	Major Components of the SE Framework . . . . .	20
2.2	MR Image Data Transformation. . . . .	23
2.3	Points of Convergence over the Relevant Range of $\sigma$ . . . . .	29
2.4	Signal Intensity Distributing in MR images . . . . .	32
2.5	Signal Fluctuations Due to Noise . . . . .	36
2.6	Wavelet Decomposition of the Squared Magnitude Representation of a Binary (Ring) Image. . . . .	38
2.7	Wavelet Decomposition of the Squared Magnitude Representation of an Image with Fading Intensity. . . . .	39
2.8	Shrinkage Functions. . . . .	44
2.9	Contrast Improvement by Means of Gradient Amplification . . . . .	52
2.10	Behavior of Operator $\Phi$ on Signal $I$ . . . . .	53
3.1	Image Representations for Performance Evaluation. . . . .	65
3.2	Noise-free Image with Region Indices. . . . .	67
3.3	Synthetic Image: A Specimen of Noisy Input. . . . .	68
3.4	Comparative Performance on Synthetic Image: Contrast Improvement. . . . .	69
3.5	Comparative Performance on Synthetic Image: Error of Approximation. . . . .	71
3.6	Cardiac Image: SRS0000-10. . . . .	74
3.7	Cardiac Image: SRS0000-70. . . . .	75
3.8	Cardiac Image: Eman-83. . . . .	75
3.9	Cardiac Image: SRS0006-24. . . . .	76
3.10	Brain Image: BG-3. . . . .	77
3.11	Cardiovascular Image: Carotid Arteries 2. . . . .	78
3.12	Comparative Performance on Cardiac Image: Eman-70 . . . . .	79
3.13	Carotid Arteries . . . . .	80
3.14	Comparative Performance on Cardiovascular Image . . . . .	81
3.15	Comparative Performance on Brain Image. . . . .	82
3.16	Comparative Performance on Cardiac Image: SRS0006-24. . . . .	83

A.1	Multiresolution Decomposition of Image. . . . .	100
A.2	Histogram of Wavelet Coefficients using Haar Transform. . .	102
C.1	Estimation of Reference Range of Noise Parameter $\sigma$ . . . . .	109
D.1	Point of Convergence of Rician Data. . . . .	122
D.2	First and Second Cumulants of Rician Distribution. . . . .	123
D.3	Standardized Third and Fourth Cumulants of Rician Distribution. . . . .	124
D.4	Convergence of Cumulants of Rician Distribution for $\sigma = 4$ and $\sigma = 6$ . . . . .	125
D.5	Convergence of Cumulants of Rician Distribution for $\sigma = 8$ and $\sigma = 10$ . . . . .	126
G.1	Synthetic Images: Contrast Improvement. . . . .	144
G.2	Synthetic Images: Error of Approximation. . . . .	145
H.1	Cardiac Image: SRS0000-10. Diffusion Group. . . . .	147
H.2	Cardiac Image: SRS0000-10. S Group. . . . .	148
H.3	Cardiac Image: SRS0000-10. SE Group . . . . .	149
H.4	Cardiac Image: SRS0000-70. Diffusion Group. . . . .	150
H.5	Cardiac Image: SRS0000-70. S Group. . . . .	151
H.6	Cardiac Image: SRS0000-70. SE Group. . . . .	152
H.7	Cardiac Image: Eman-70. Diffusion Group. . . . .	153
H.8	Cardiac Image: Eman-70. S Group. . . . .	154
H.9	Cardiac Image: Eman-70. SE Group. . . . .	155
H.10	Cardiac Image: Eman-83. Diffusion Group. . . . .	156
H.11	Cardiac Image: Eman-83. S Group. . . . .	157
H.12	Cardiac Image: Eman-83. SE Group. . . . .	158
H.13	Cardiac Image: SRS0006-24. Diffusion Group. . . . .	159
H.14	Cardiac Image: SRS0006-24. S Group. . . . .	160
H.15	Cardiac Image: SRS0006-24. SE Group . . . . .	161
H.16	Brain Image: BG-3. Diffusion Group. . . . .	162
H.17	Brain Image: BG-3. S Group. . . . .	163
H.18	Brain Image: BG-3. SE Group. . . . .	164
H.19	Cardiovascular Image: Carotid Arteries 2. Diffusion Group. . . . .	165
H.20	Cardiovascular Image: Carotid Arteries 2. S Group. . . . .	166
H.21	Cardiovascular Image: Carotid Arteries 2. SE Group. . . . .	167

## Acronyms

Acronym	Meaning
CIR	Contrast improvement ratio
CMR	Cardiac magnetic resonance
CMRI	Cardiac magnetic resonance imaging
CT	Computer-aided tomography
CVD	Cardiovascular disease
MCMC	Markov chain Monte Carlo
MR	Magnetic resonance
MMR	Magnitude magnetic resonance
MAD	Median absolute deviation
MRA	Magnetic resonance angiography
MRI	Magnetic resonance imaging
MSE	Mean squared error
SI	Signal intensity
SNR	Signal-to-noise ratio
T-NT	Tissue-to-non-tissue
T-T	Tissue-to-tissue

Acronym	Wavelet Base
DB $k$	The $k^{th}$ order Daubechies wavelet
Haar	Haar wavelet (DB 1)
Sym $k$	The $k^{th}$ order Symmlet or least asymmetric wavelet

## Notation

Symbol	Meaning
$\sim$	Asymptotically approximate
$\langle \cdot \rangle$	Inner product
$\ \cdot\ $	$\mathcal{L}_2$ -norm
$\tilde{\mathcal{X}}$	Median of $\mathcal{X}$
$1_{\mathcal{A}}$	Indicator function of a set $\mathcal{A}$
$f^{(r)}(x) = \frac{d^{(r)}f(x)}{dx^{(r)}}$	Derivative taken $r$ times
<b>Card</b> $\mathcal{X}$	Cardinality of the set $\mathcal{X}$
$D_i$	Distance in the $i^{th}$ cumulant between a Rician distribution and a Gaussian distribution
<b>E</b> ( $\cdot$ )	Mean or expected value of a random variable
${}_1F_1(\cdot; \cdot; \cdot)$	Confluent hypergeometric function of the first kind
$\Gamma(\cdot)$	Gamma function
<b>H</b> ( $\cdot$ )	Heaviside step function
$\inf \mathcal{A}$	Infimum (or the greatest lower bound) of the set $\mathcal{A}$
$I_n$	Bessel function of the first kind of the $n^{th}$ order
$\mathcal{K}$	Cumulant-generating function of a distribution
$\kappa_i$	The $i^{th}$ cumulant of a distribution
$\kappa_i^s$	The standardized $i^{th}$ cumulant of a distribution
$\mathcal{R}_{\kappa_i}$	The $i^{th}$ cumulant of a Rician distribution
$\mathcal{R}_{\kappa_i^s}$	The standardized $i^{th}$ cumulant of a Rician distribution
$\mathcal{L}_2$	The space of square-integrable (finite-energy) functions (signals)
$\mathcal{L}(f, \hat{f})$	The loss function of estimation of $f$ by $\hat{f}$
$M(t)$	Moment generating function of a probability distribution
$m_i$	The $i^{th}$ moment of a distribution
$\mathbb{R}^n$	The $n$ -dimensional space of real numbers
<b>R</b> ( $f, \hat{f}$ )	Risk of estimation of $f$ by $\hat{f}$
$\mathcal{S}^c$	The point of convergence
$\widetilde{\mathcal{S}^c}$	The noisy realization of point of convergence in empirical data
$\phi_{j,k}$	A scaling function at scale $j$ and location $k$
$\mathcal{S}_{j,k}$	an approximation coefficient associated with $\phi_{j,k}$
<b>sgn</b>	Signum function
$\sup \mathcal{A}$	Supremum (or the least upper bound) of the set $\mathcal{A}$
<b>X</b> <sup>T</sup>	Transpose of a matrix or vector <b>X</b>
$\mathcal{V}_J$	An approximation space spanned by scaling functions at scale $J$
$\psi_{j,k}$	A wavelet function at scale $j$ and location $k$
$\psi_\eta$	A wavelet function at scale, location and orientation indexed by $\eta$
$\omega_{j,k}$	A wavelet coefficient associated with $\psi_{j,k}$
$\omega_\eta$	A wavelet coefficient associated with $\psi_\eta$
$\xi_{\omega_\eta}$	Noise component of a wavelet coefficient $\omega_\eta$

*continued on next page*

Table 3: *continued*

Symbol	Meaning
$\tilde{\gamma}_{\omega_{\eta}}$	Noise-free component of a wavelet coefficient $\omega_{\eta}$
$\mathbb{Z}$	The space of integers $\{\dots - 3, -2, -1, 0, 1, 2, 3, \dots\}$

# Chapter 1

## Introduction

Cardiovascular disease<sup>1.1</sup> (CVD) remains the leading cause of mortality and hospitalization in North America. The prevalence of CVD is not restricted to any particular age groups. Indeed it ranks among the top three causes of death for children under the age of 15 in the USA. Early diagnosis and treatment are effective means in reducing the prevalence of CVD. It accounts for 25% of the decline in mortality from ischemic heart disease in the USA. The advances in cardiac imaging technology play an important role in visualizing the pathology and physiology of the cardiovascular system and therefore allow for early detection of CVD, improvement of the accuracy of diagnoses, measuring the effects of treatments, and determination of CVD-related risks. In this regard, cardiac magnetic resonance imaging (CMRI) is one of the most flexible and sophisticated imaging techniques with which to examine the heart. It provides a set of images for visualizing the cardiovascular system as well as for reliably measuring the anatomical structures

---

<sup>1.1</sup>For the social consequences of the disease, see American Heart Association [5] and Heart and Stroke Foundaton of Canada [55].

of the organ and the patterns of its deformation over the cardiac cycle<sup>1,2</sup>.

## 1.1 The Problem of Adaptive Enhancement of Cardiac Magnetic Resonance (MR) Images

The purpose of this report is to present a computational framework for signal enhancement of magnitude Magnetic Resonance Imaging (MRI) for medical visualization and analysis. Although developed in the context of Cardiac Magnetic Resonance (CMR) image analysis, the framework finds a much broader scope of application in other (non-cardiac) classes of clinical images. Nevertheless, CMR image analysis remains the natural context that motivates and defines the problems for this thesis.

The task of CMR image analysis is to transform a set of cardiac MRI data<sup>1,3</sup> into diagnostically-relevant information about the imaged heart. Diagnosis of cardiac disease relies upon an accurate assessment of the morphology and function of the patients' hearts. Recovery of the morphology of the heart and its anatomical structures is critical to the reliable modeling of the heart and the measurement of the diagnostic parameters of the heart.

---

<sup>1,2</sup>A complete period of contraction and relaxation of the heart with the electrical and mechanical events that control the opening and closing of the valves and the flow of blood in and out of the organ constitutes a cardiac cycle.

<sup>1,3</sup>A set of CMRI data consists of a number of temporal sequences, each representing a dynamic view of a cross-section of the thorax which contains the heart over a cardiac cycle<sup>†</sup>. These sequences are ordered by spatial location of the cross sections along an axis of imaging. In a typical CMR image, the heart appears as a configuration of structures located within the bony structures of the thoracic wall, on the diaphragm, above the liver and stomach, and against the background of the thoracic cavity; See Fig. 1.1. This image sequence provides a discrete, volumetric representation of the heart and its internal anatomical structures over time.

<sup>†</sup> see footnote 1.2.

The detection of structure in medical images usually relies on the difference in the signal intensity of different tissues. These signal intensity differences determine image contrast; see Weishaupt et al. [121]. Compared with the image contrast along the interfaces of the anatomical parts of the organ, the appearance and signal intensity of tissue in a clinical MR image is, in general, relatively insignificant because signal intensity depends on imaging techniques and scanning parameters employed. Under different parametric set-ups, the same scanner may produce varying intensity profiles of the same scene. Depending upon the information acquired for clinical diagnosis and analysis, signal intensity of different tissues as well as their visibility is a function of intrinsic (tissue-related) and extrinsic (instrument-related) parameters of data acquisition. The range of choices over these parameters partially attributes to the versatility of MRI for clinical analysis. It allows a sufficient contrast for discriminating many tissues which are indistinguishable on computer-aided tomography (CT). These choices are however made under competition among many imaging requirements, such as imaging time, spatial and temporal resolution, and reduction of imaging artifacts, under resource constraints, such as magnetic field strength, choice of transmitters and receiver coils, and the others. This research explores a solution to enhance image contrast along those interfaces of anatomical structures that carry diagnostic information in magnitude MR images. These structures are hereafter referred to as target structures.

In general, the quality of clinical MRI does not allow an easy interpretation of the data for extracting diagnostic information. Recovery of the boundary contours/surfaces of target organs remains a non-trivial endeavor.

---

This burden would be greatly simplified if these interfaces were embedded in less ambiguous configurations of structures, that is, if those features attributable to the target structures were not significantly obscured by similar, but irrelevant, structures in their proximity. In the context of CMR images, for instance, two factors contribute to obscure target structures and reduce the image contrast: image noise and dynamic structures cluttering the thoracic cavity; see Fig. 1.1. They are hereafter referred to as confounding factors.

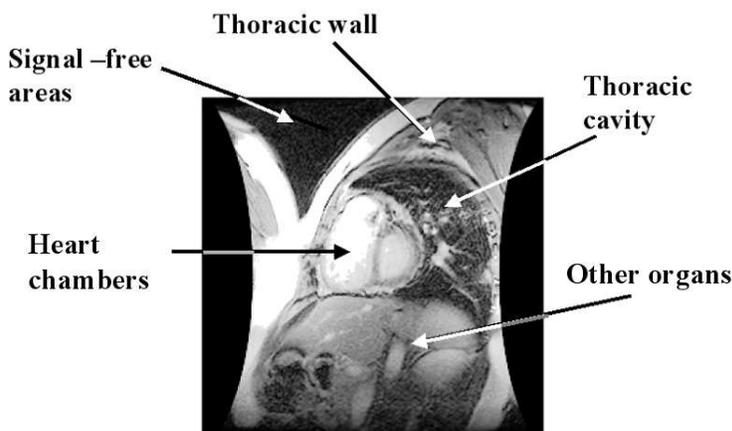


Figure 1.1: A Cardiac Magnitude MR Image.

The noise of a magnitude MR image, generally modeled by a Rician distribution, is signal dependent with a distribution that tends to a Rayleigh distribution<sup>1.4</sup> over the low intensity regions and to a Gaussian distribution

---

<sup>1.4</sup>Both the Rician and the Rayleigh distribution arise in the context of signal detection with narrowband<sup>†</sup> noise processes. Given in Eq. (2.2.2) is the probability density function of the Rayleigh distribution, that is defined only over the positive values of the random variable with a positive (non-zero) mean. The magnitude of a complex addition of Gaussian real and imaginary noise components resulting from a collection of small, independent wave disturbances due to scattering of transmitted signals follows a Rayleigh distribution. The superimposition of the signal and these disturbances gives rise to the Rician distribution, the probability density function of which is given in Eq. (2.2.1). For details, see Goodman [49]; Minkoff [78, 79] and Whitaker and Benson [123]. This model of noise disturbance provides an appropriate description of the noise processes in the for-

over the high intensity regions; see Gudbjartsson and Patz [52]; Macovski [73] and Kisner et al. [68]. In a magnitude CMR image, the low-intensity signal is usually present along the interface of the anatomical parts and over the background of the heart. The bias introduced by the Rician noise significantly reduces the image contrast essential for extracting the information about the boundaries of the heart and its anatomical parts. Furthermore, this noise effectively reduces the lossless compression ratio, which is important to image transmission; see Wu et al. [127].

Noise is not the only confounding factor that obscures the target structures. The thoracic cavity appears as a low intensity region where the lungs, blood vessels and other tissues create patterns of dynamic structures in the immediate periphery of the heart. It is usually not easy to discriminate these networks of structures from the target organ. These fibrous structures lack a definite shape. Like the heart, many of them are dynamic structures in periodic motion. Both characteristics represent a challenge to modeling them for identification and discrimination. These low-intensity structures in the proximity of the target organ are hereafter referred to as background structures. The interactions between these background structures and the Rician noise further confound the tasks of boundary detection and anatomical reconstruction both spatially and dynamically.

---

mation of magnitude MR images. The details are covered in Section 2.2 in the following chapter and Appendix B.

‡ A process with no signal energy distributed outside a frequency interval  $(\omega_1, \omega_2)$  is called a bandpass process. If the bandwidth,  $\omega_2 - \omega_1$ , is small compared with the center frequency, the process is called a narrowband or quasi-monochromatic process. See Papoulis [86] and Qian [95].

In this light, this thesis seeks an efficient computational framework to offset the confounding effects of these factors, and to provide a data set with enhanced image contrast, especially that attributable to the target structures. To be precise, its objectives are as follows : (1) to recover from the image noise the signals of the anatomical structures of the heart; (2) to enhance the image contrast attributable to these signals and (3) to suppress the background structures. The problem so defined seeks a solution which is inherently adaptive to both signal gradient as well as intensity. It will be referred to hereafter as the problem of adaptive enhancement.

## 1.2 Background and Related Research

Enhancement of medical images has been an active research area; specifically, reduction of image noise has been a major focus of attention. To the best knowledge of the author, however, no existing research directly addresses the problem of adaptive enhancement in its full scope. To put this report in perspective, this section provides a brief review of the major research approaches to the problem. Signal enhancement by way of noise reduction has been developed for MRI applications with the assumption of Gaussian noise as the working model. Diffusion based filtering in the spatial domain and selective reconstruction by series expansion in the wavelet domain have been widely adapted for biomedical image application. Both are useful approaches for enhancing the target organs and their anatomical structures.

Non-linear diffusion filtering<sup>1.5</sup> is developed to overcome the blurring effect of classical Gaussian smoothing techniques. Its diffusion mechanism preserves image structures, especially edges. Signal oscillations are reduced with a diffusion kernel, which adapts its smoothing operators to local oscillating structures, steering the direction of diffusion in such a way so that it preserves geometric structures of the data, such as edges and contours. These techniques provide good results for image enhancement in a variety of applications. In the present context, however, diffusion-based techniques pose some problems. First, diffusion preserves local averages of signal intensity and therefore provides no natural mechanism to eliminate the bias introduced by Rician noise. This is especially true for low intensity regions. Second, the goal of adaptive enhancement is not to enhance every geometric structure. Enhancement of target structures is desired whereas confounding factors, such as background structures, should be suppressed. This idea of enhancement is not intrinsic to the diffusion model. Furthermore, diffusion is an iterative process which brings with it a cost in terms of computational complexity.

The wavelet domain denoising technique introduced in Weaver et al. [117] with a user-defined threshold<sup>1.6</sup> is among the earliest application of wavelet analysis to image denoising. Briefly, wavelet analysis<sup>1.7</sup> maps image signals

---

<sup>1.5</sup>This approach is introduced in the work of Perona and Malik [87] and extended subsequently; for a review see Weickert [118, 120].

<sup>1.6</sup>It is worth pointing out the conceptual similarity between the coring and the wavelet approaches to noise reduction. Coring is essentially a nonlinear approach to noise reduction by means of selecting signal components according to thresholding rules in terms of their absolute intensity level. This similar idea underlies the technique proposed in Weaver et al. [117] and the subsequent approaches of wavelet shrinkage. For more details, see Carlson et al. [16]; Simoncelli and Adelson [103] and Portilla et al. [93].

<sup>1.7</sup>For details on the mechanics of wavelet analysis, see Appendix A.

to a set of coefficients by signal decomposition in terms of a set of analyzing functions (or basis functions). This set of coefficients constitutes a compact representation that provides a localized measurement of signal oscillations in both scale and space. The operation is invertible, that is, a corresponding inverse mapping, which achieves signal synthesis, exists to generate a lossless reconstruction of finite-energy signals with the coefficients. If analysis is simple-mindedly followed by synthesis without changing the coefficients, the operation is of course a trivial endeavor producing no better output than the input. Image processing seeks a new representation of input data. By the multiscale properties of a wavelet system, there is a set of coefficients that yields a perfect synthesis of this target representation. The whole problem of image transformation is thus turned into a search for the mapping from the decomposition of the input representation (i.e., the input coefficient set) to a set of coefficients (i.e., the target coefficient set) that supports the generation of the target representation. Wavelet domain image transformation can therefore be summarized as an analysis→transformation→synthesis operator sequence:

**Analysis :** Compute the coefficients from a spatial domain data set.

**Transformation :** Apply transformation operators to the coefficients.

**Synthesis :** Compute the spatial-domain representation with the coefficients generated in the transformation step.

Multiresolution decomposition of a signal into its constituent waveforms is fundamental to wavelet analysis. A technical note on the mechanics of wavelet analysis that are relevant to this research is provided in Appendix A. It is sufficient at the moment to say that a set of image data can be expressed

as a superimposition of a coarse scale approximation and the details capturing spatial oscillations at successive finer scales. The coarse scale approximation generated by wavelet decomposition represents smoothly varying signals. The rapidly oscillating components and the transient elements due to meaningful structures are captured in a small number of large coefficients, the spikes towering over a sea of small coefficients due to the fluctuations caused by random (Gaussian white) noise. This stark contrast in signatures between coherent oscillations and random fluctuations provides useful information for discriminating these two important classes of image features for noise suppression and structure preservation. This observation provides the point of departure for the classical approach of wavelet shrinkage. Signals due to coherent structures are recovered from a subset of the largest coefficients selected according to some decision rules. Hard thresholding and soft thresholding are exemplars. Given a threshold, hard thresholding suppresses all coefficients smaller than the threshold. Soft-thresholding shrinks coefficients toward zero so that it does not only suppress small coefficients, but also attenuates the components due to random fluctuations of those surviving coefficients.

While Gaussian noise remains a tacit, working assumption used in many applications, recent findings suggest that a Rician noise model provides a more accurate account for the stochastic behavior of the noise commonly found in the magnitude MRI used in the clinical setting; see Andersen [6] and Macovski [73]. In the general context of clinical MRI, there are two major approaches to the problem of Rician noise. The distinction between them can be traced back to the very process which brings about the Rician noise in MRI analysis. Rician noise does not occur in signal acquisition, but arises

from the process of constructing magnitude MR images. A complex-valued spatial-domain representation is generated from the frequency-domain data read out from an MRI machine through an inverse Fourier transform. The magnitude image is the square root of the sum of square of the complex pair. The real and imaginary data after the inverse transform remain Gaussian distributed with means of zero and equal variance  $\sigma^2$ . The non-linear operator in the last step of image construction, however, transforms the distribution from Gaussian to Rician - a distribution which is signal dependent.

Rician noise is present in the magnitude image, but not in the complex images from which the magnitude image is generated. One approach to noise enhancement is to eliminate the Gaussian noise from the complex images before the construction of magnitude images. This approach gives rise to a number of methods: Wood and Johnson [125, 126] use wavelet packets, Alexander et al. [4] derive a wavelet-domain Wiener-type filter and Bao and Zhang [11] apply multi-scale product thresholding to denoising the complex image pair.

The problem of Rician noise in a magnitude MR image is addressed almost exclusively within the framework of wavelet analysis. The implications of the non-Gaussian nature of image noise in the context of machine understanding of visual data are raised in Gregg and Nowak [51] and Nowak [83] which start with a demarcation of two MRI regimes according to the signal-to-noise ratio (SNR) of the acquired images: (1) low resolution and high SNR and (2) high resolution and low SNR. As a result, a magnitude MR image is deemed as either a high SNR image or a low SNR image. A wavelet-domain filter derived for the high SNR images is a set of weights

which shrink the wavelet coefficients toward zero in such a way that the shrinkage factors approximately minimize the mean squared errors (MSE) of the estimates. This filter relies on the assumptions that (1) noise in the high SNR image is signal independent, and (2) the wavelet coefficients are unbiased. Obviously these conditions fail to hold for the low-SNR images. For the low SNR (below 10dB) MRI, Gregg and Nowak [51] and Nowak [83] apply the framework of noise filtering to squared magnitude images with two additional steps: (1) to provide an unbiased estimator of noise variance using the discrete wavelet squared transform of the squared image and (2) to remove the bias from the approximation coefficients of the image. This second algorithm with the additional steps, however, fails to outperform the first one for any MRI with the SNR  $> 15$  dB, a condition that clinical applications of MRI can usually meet; see Nowak [83]. The technique of removing bias from the squared image was adapted in some later algorithms, such as the algorithm developed by Xue et al. [128].

The division of images into two categories ignores the fact that even a so-called high-SNR image is composed of both high and low intensity regions. Over these low intensity regions, however, this filtering framework fails to remove the Rician noise satisfactorily. Wu et al. [127] are concerned with the noise in the signal-free regions of the image. An additional step is proposed to identify the background noise for removal with a map which indicates the locations of signal-free regions according to a threshold scheme.

The issue of general relevancy of the effects of Rician image noise is raised in Wood and Johnson [126]. While an image may attain an overall high SNR, important image features may fall within an intensity region where a skewed

Rayleigh distribution may dominate. As a result, Rician noise degrades edge and contrast resolution. A wavelet packet denoising technique<sup>1.8</sup> is applied to both complex image pair and magnitude images for the purpose of noise reduction. A compact basis is selected for decomposing the image data so that Gaussian noise is distributed evenly throughout the coefficients in the hierarchical representation in different subspaces. Shrinkage procedures are then applied to the coefficient of the resulting decomposition. Compared to magnitude denoised images, complex denoised images are found to have better noise suppression, resulting in a sharper definition of the anatomical structures of the target organs. These results indicate for our purpose that Rician noise in the high signal intensity regions can be smoothed out by Gaussian approximation techniques, if given an accurate estimate of variance of the Rician noise, but the problem of noise suppression in the Rician situations is at best partially solved.

A spatial domain technique is proposed in Sijbers et al. [100] to address the problem of Rician noise. The noise reduction is achieved by a diffusion process using a Gaussian kernel which is pointwise adaptive to local structures. The shape of the kernel is controlled by the local gradient strength along the main direction of the patterns in a local neighborhood. These local gradients are estimated through maximization of the likelihood function based on the assumption of a signal which is locally piecewise linear over a neighborhood with additive Rician noise. The Rician-based filter is found to perform better than the Gaussian-based filter in terms of contrast restoration; the difference is most apparent in regions of low SNR. The

---

<sup>1.8</sup>Wavelet packet decomposition can be seen as a generalized wavelet decomposition. For details, see Ogden [84] and Walnut [112].

signal dependent nature of Rician noise requires a diffusion kernel which is adaptive to both local structures and signal intensity. The performance gain comes mainly from the improved estimation of local structures.

Other techniques are also proposed. Woo and Yang [124] propose using additive, student t-statistics to model the Rician noise under a Bayesian framework using a Markov Chain Monte Carlo (MCMC) technique to denoise low SNR MR images. This alternative noise model which maintains the assumption of signal independence and a symmetric shape of the distribution does not go much farther than the Gaussian model in addressing the Rician nature of noise behavior. The iterative nature of the MCMC used for parameter estimation also incurs a very high cost in terms of computational resources. A generic framework of noise reduction in the wavelet-based domain is proposed in Pizurica [88] and Pizurica et al. [92] based on the method of joint signal detection and estimation. Each wavelet coefficient is modeled as a representation of a signal of interest with a probability defined in terms of the global coefficient histogram and the local measure of spatial activity. In the absence of the prior knowledge, the probability densities are estimated from the empirical histogram. An MR image is taken as one of its potential applications, but no detailed analysis is provided for magnitude MR images. Indeed, it is not even clear how the noise parameters are estimated.

### **1.3 Contribution**

The approaches presented in the previous section provide important clues about the nature of the noise effect and its possible solution. Rician noise properties degrade the quality of magnitude MR images. Regardless of the

general intensity level or SNR, the effects of noise behavior have a practical impact on the low intensity regions which are not only restricted to signal-free regions, but also include those low contrast structures with diagnostic significance.

One common theme running through all these approaches is noise reduction. The objectives of the techniques are defined in terms of two considerations: how to (1) remove the noise and (2) at the same time preserve the integrity of the signal. This is a classic problem with medical image enhancement. The definition of the problem, however, ignores other confounding factors in medical images, that may obscure the diagnostically significant information, such as background structures mentioned in the foregoing discussion. These factors may interact with each other. No previous work addresses these problems, even tangentially, in the context of medical image enhancement.

The objective of this thesis is to explore the problem of adaptive enhancement where Rician noise occurs. This approach to the problem of data quality is largely motivated by the problem of image analysis in the context of cardiac magnitude images, a class of MR images, which, unlike brain images, receives relatively little attention in research on medical image enhancement. There are two major components of inquiry. First, Rician noise remains a major concern in an adaptive enhancement. This report provides a detailed, statistical analysis of the effect of the statistical behavior of Rician noise in magnitude MRI. It clarifies the theoretical underpinnings of some assumptions used in the existing literature. In particular, the analysis leads to a meaningful demarcation between high intensity and low intensity

regions of a magnitude MR image in terms of the qualitative difference in noise effect. The latter are the active zones where non-Gaussian behavior is active and therefore impacts negatively upon the image quality. In these regions Gaussian-based techniques are inadequate to noise attenuation and are therefore inadequate for the task of image enhancement.

This analysis leads to an approximation model of Rician noise in terms of its behavior and its impact upon an image. Accordingly a computational framework is developed for the task of adaptive enhancement. Apart from the signal-dependent behavior of Rician noise, the task of achieving these objectives simultaneously is a challenge for a number of reasons. Contrast enhancement and attenuation of noise are two major components of adaptive enhancement. It is a long-standing problem to remove noise and enhance contrast gradients in a jointly optimal way. This report seeks a better solution to the problem of contrast and structure enhancement within the constraints defined by the objectives of adaptive enhancement that it suppresses not only random noise, but also the background structures while preserving the target structures of the organ in their proximity.

This research shows that an optimal decomposition of the signal in the wavelet domain and the signal-varying behavior of confounding factors allows an adaptive scheme to achieve the objectives of adaptive enhancement:

1. Gaussian random noise manifests itself as high frequency components of a signal. Noise attenuation therefore means attenuation of these components either by smoothing, shrinkage or other means. This same idea plays the guiding principle underlying the existing approaches to the problem of Rician noise. In contrast, this research works on the

hypothesis that the non-Gaussian noise active in low intensity regions of an image and background structures may be captured in the coarse scale approximation.

2. Image contrast is always a concern for image enhancement. In the context of image denoising, contrast improvement is usually achieved as a by-product of noise reduction and signal preservation. This report postulates that intensity gradients can be much better enhanced by adaptively attenuating the average intensity of the smooth parts of the image while preserving the rapidly oscillating components of image signals. The wavelet decomposition of the signal provides the mechanism for the task.

It is a hypothesis of this thesis that the adaptive enhancement framework can jointly achieve the objectives of adaptive enhancement. The performance of the adaptive enhancement framework is systematically evaluated in comparison with both wavelet-based and diffusion-based techniques on a variety of images, both cardiac and non-cardiac. Not only does the evaluation support the hypothesis of the report and find its performance significantly better than the other techniques, but also it shows that its scope of application can be extended outside the scope of cardiac images. For the performance evaluation, this report provides an analysis of the measurement metrics for measuring and comparing system performance in terms of image contrast improvement. It shows the source of bias in the metric used in previous research and proposes a solution to correct the problem. Performance measured in terms of both metrics – the original and the improved version – are reported for comparison purposes. In general, the framework presented in this report provides an efficient solution to the problem of adaptive enhancement.

## 1.4 Summary

The purpose of this report is to explore the problem of adaptive enhancement in a Rician situation, a problem that is first motivated in the context of cardiac MR image analysis. The problem of adaptive enhancement is defined in terms of its threefold objectives as removal of confounding factors, preservation of target structures and enhancement of image contrast. A brief review of previous research reveals a growing awareness of the effects of the signal-dependent noise on the image quality of magnitude MR images and on the performance of algorithmic approaches to image analysis. The image noise of magnitude MR images, especially over their low-intensity regions, remains a challenge to the currently available techniques of noise removal. This report suggests seeking a better solution to the problem of image enhancement in the broader context of adaptive enhancement. Chapter 2 starts with an analysis of the signal properties of the confounding factors, that leads to a wavelet-based framework devised to jointly achieve the objectives of adaptive enhancement. Experimental results are presented in Chapter 3, which shows strong support for the adaptive enhancement framework in terms of its objectives and in comparison with other techniques. Chapter 4 concludes the report with some general comments on the approach and suggestions for further research. A number of appendices are included for reference on technical details and a systematic presentation of experimental results.

## Chapter 2

# Computational Framework

This chapter lays out a computational framework, referred to hereafter as the smoothing - enhancement (SE) framework, for adaptive enhancement of magnitude MR images. Conceptually, the framework has three components: (1) multiresolution analysis of the image signal; (2) analysis of signal properties of confounding features; and (3) an adaptive scheme for image enhancement. I will discuss each component in turn.

The first component provides a powerful tool for signal analysis, transformation, and synthesis. This chapter begins by outlining the general ideas underlying multiresolution (or multiscale) analysis of an orthogonal wavelet system and its role in noise enhancement. The goal of doing so is to introduce the notations in use throughout the thesis. A more detailed discussion of the mechanics of wavelet analysis as it relates to the framework can be found in Appendix A.

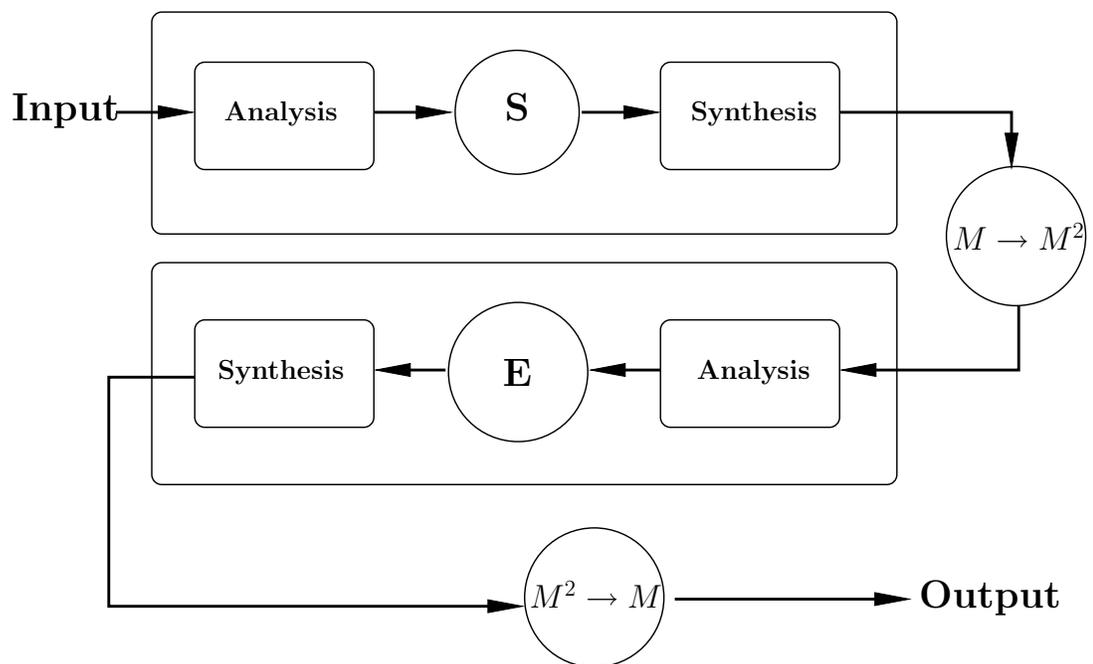
The focus of Section 2.2 on the second component of the SE framework is an analysis of the signal properties of confounding features. The first

part of the analysis provides an analytic description of noise in a magnitude MR image, and is followed by a discussion on the transformed behavior of the confounding factors in the squared magnitude images. These analyses lead to the key observations of this report. First, Rician and Gaussian noise behave differently in regions of low signal intensity where features along the interface between the anatomical parts and background structures reside. Second, the wavelet representation of squared data captures the effects of the confounding factors in the coarse scale sub-image. Third, wavelet decomposition provides a natural means of contrast enhancement by manipulating the average intensity over the regions of smoothly varying signals.

The last part of this chapter discusses the implications of these observations, showing that an effective enhancement scheme is possible to achieve the dual goal of contrast enhancement and confounding feature suppression. This is done in two steps: (1) the smoothing (S) step preprocesses the image data to remove the fine scale signal fluctuations due to random noise according to an approximation to its behavior in high intensity regions; and (2) the enhancement (E) step applies an adaptive enhancement scheme to the squared magnitude data of the pre-processed image generated from the previous step to attenuate the effects of confounding features and enhance image contrast( see Fig. 2.1).

In order not to distract attention from the main line of reasoning, the details of each of these sections are covered in a set of technical notes in Part I of the Appendix.

Figure 2.1: Major Components of the SE Framework



**Legends**

- S** Fine Scale Smoothing.
- E** Coarse Scale Enhancement.
- $M \rightarrow M^2$  Construction of a squared magnitude image.
- $M^2 \rightarrow M$  Construction of a magnitude image.

## 2.1 Signal Representation in the Wavelet Domain

In the wavelet domain, the observed data  $\mathcal{I}$  is projected onto a multi-resolution analytic space; for a review of the mechanics of wavelet analysis, see Appendix A. Image structures can then be analyzed in terms of a hierarchy of sub-images, each capturing signal oscillations at different scales and orientations. An orthogonal wavelet system is defined in terms of an orthogonal basis which consists of a set of scaling functions  $\{\phi_{j_0, \mu}\}$ , each localized around the location ( $\mu$ ) at a selected scale ( $j_0$ ) and their corresponding wavelet functions  $\{\psi_{\mathbf{j}, \mu}^\kappa\}$ , each capturing the signal oscillations in the neighborhood around locations  $\mu$ , at successively finer scales ( $\mathbf{j}$ ) and along orientations<sup>2.1</sup> ( $\kappa$ ). An image can be expanded in terms of this basis by the following analytical form:

$$\mathcal{I} = \sum_{\mu} \mathcal{S}_{j_0, \mu} \phi_{j_0, \mu}^{2D} + \sum_j \sum_{\mu} \sum_{\kappa \in \Psi} \omega_{j, \mu}^\kappa \psi_{\mathbf{j}, \mu}^\kappa. \quad (2.1.1)$$

Wavelet analysis is a mapping from image data in the spatial domain into a set of coefficients in the wavelet domain, given a basis of analysis. The set of scaling (or approximation) coefficients  $\{\mathcal{S}_{j_0, \mu}\}$  in the first term in Eq. (2.1.1) fully characterizes the coarse scale approximation up to a selected scale  $j_0$  whereas the set of wavelet coefficients  $\{\omega_{j, \mu}^\kappa\}$  in the second term represents a superimposition of sub-images each capturing the detail of image structures visible at successively finer scales. In the subsequent discussion, it is sometimes convenient to use a more compact form, i.e.,

$$\mathcal{I} = \sum_{\mu} \mathcal{S}_{j_0, \mu} \phi_{j_0, \mu}^{2D} + \sum_{\eta} \omega_{\eta} \psi_{\eta}, \quad (2.1.2)$$

---

<sup>2.1</sup>  $\kappa \in \Psi = \{\mathbf{H}, \mathbf{V}, \mathbf{D}\}$  is the set of orientations – horizontal, vertical, and diagonal respectively – of the basis functions of a detail space.

where  $\eta = \{j\}_{j>j_0} \times \{\mu\} \times \{\kappa\}$  is an index set over  $j$  (the scale),  $\mu$  (the location), and  $\kappa$  (the orientation). The representation generated in this analytical framework provides a number of useful properties for adaptive enhancement including the property of locality, multiresolution, statistical independence (or decorrelation), energy compaction, exponential decay and others. For a review of these properties, see Appendix A.

## 2.2 Signal Properties of Confounding Features in Magnitude MR Images

This section seeks an analytical description of the signal properties of confounding features such as noise and background structures. A number of conceptual questions are addressed in connection with adaptive enhancement. Specifically, adaptive enhancement is concerned with the non-Gaussian behavior of image noise, the role of a Gaussian noise model as an approximation to this specific class of image noise, and the signal properties of confounding features such as the background structures in magnitude and squared magnitude image data. This analysis not only leads to the enhancement scheme presented in the last part of this chapter, but also helps to clarify and assess the validity of many assumptions made in the literature. In many cases different assumptions are made based on practical experience or intuition<sup>2,2</sup> without adequate analytical groundwork<sup>2,3</sup>.

---

<sup>2,2</sup>In addition, Gaussian noise models are usually adopted implicitly simply for convenience or assumed by default.

<sup>2,3</sup>At least an adequate analysis is not reported.

### 2.2.1 Noise Behavior in Magnitude MR Images

A magnitude MR image, a representation used in clinical settings, is constructed from the measurement of an MRI scanner in the frequency-domain. The measurement is subject to a variety of random noise sources related to imaging hardware and the physiology of imaged subjects. The resulting noise can be modeled as a set of random variables that follow a multivariate normal distribution; for details, see Wang and Lei [115] and Wang [114]. The non-linear operation of (magnitude) image construction does not preserve the Gaussian nature of image noise. Instead, it creates Rician noise. A squared magnitude image results from pixel-wise squaring of a magnitude

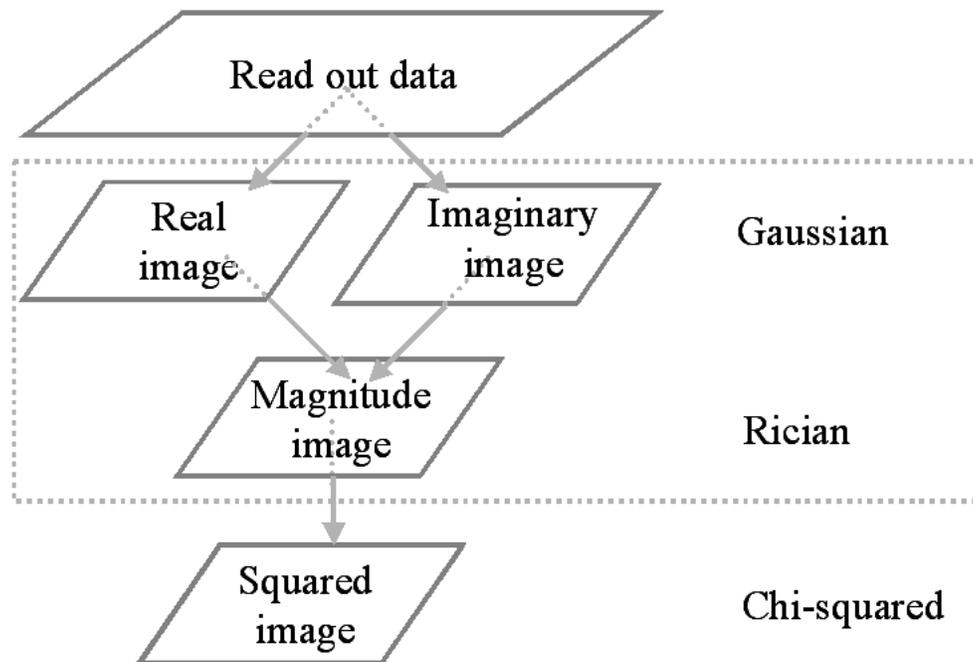


Figure 2.2: MR Image Data Transformation. The major components in the construction of a magnitude image is inside the dotted box.

image. This process further transforms the Rician noise into non-central  $\chi^2$

noise<sup>2.4</sup>. These image and noise transformation processes are summarized in the schematic diagram in Fig. 2.2.

### 2.2.1.1 Rician Noise Model of Image Signals

A signal read out from an MRI scanner contains a real and an imaginary component. Each component can be modeled with additive Gaussian noise with a zero mean and the same variance; see Wang and Lei [115] and Wang [114]. A pair of real and imaginary images, referred to as complex images or a complex pair, in the spatial domain are generated by the inverse Fourier transform. Because the transform is a linear operator that preserves Gaussian properties, the stochastic characteristics of image noise are not qualitatively altered.

For the convenience of display and to remove phase artifacts, it is common in clinical medicine to use magnitude images. A magnitude image is generated after a magnitude (root-sum-squared) operation is performed on the complex images. The resulting data,  $I$ , are characterized by the conditional probability density function given the (noise-free) signal  $S$  and the variance  $\sigma^2$  of the noise in the complex images as the following

$$f(I|S, \sigma) = \frac{I}{\sigma^2} \exp \left\{ -\frac{I^2 + S^2}{2\sigma^2} \right\} I_0 \left( \frac{IS}{\sigma^2} \right) 1_{\{x \geq 0\}}(I) \quad (2.2.1)$$

where  $I_0$  is the modified Bessel function of the first kind of zeroth order<sup>2.5</sup>,

---

<sup>2.4</sup>For details, see Appendix F.

<sup>2.5</sup>For the details of the modified Bessel function of the zeroth order  $I_0$ , see Abramowitz and Stegun [2] and Arfken and Weber [9].

and  $1_{\mathcal{A}}$  the indicator function<sup>2.6</sup> over the set  $\mathcal{A}$ ; see Andrews and Phillips [7]; Minkoff [79], and Sijbers [99]. Eq. (2.2.1) represents the Rician distribution. Since the term  $\frac{S^2}{\sigma^2}$  measures the signal-to-noise ratio (SNR), the distribution depends on the SNR or on the signal intensity if the noise parameter  $\sigma$  is known. This analytical result is verified in both clinical and experimental settings; for an example, see Kisner et al. [68].

In the absence of signal, the density given in Eq. (2.2.1) can be rewritten as

$$\mathcal{F} = \frac{I}{\sigma^2} \exp \left\{ -\frac{I^2}{2\sigma^2} \right\} 1_{I>0}(I), \quad (2.2.2)$$

which is a Rayleigh distribution<sup>2.7</sup>; for details, see Appendix B.1. As SNR increases, the probability density function takes on a more and more symmetric shape. As the SNR approaches infinity, Rician noise,

$$\lim_{\frac{S}{\sigma} \rightarrow \infty} \mathcal{F} \sim \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -\frac{1}{2\sigma^2}(I - S)^2 \right\}, \quad (2.2.3)$$

is approximately a Gaussian distribution  $\mathcal{N}(S, \sigma^2)$ ; for details, see Appendix B.2. Let us call this distribution the limiting distribution of infinite signal intensity or the limiting distribution for short.

At infinite signal intensity, a Rician distribution converges to a Gaussian distribution with mean  $S$  and variance  $\sigma^2$ . This result leads to the conjecture that the stochastic behavior of a Rician variable may be reasonably approximated by the limiting distribution defined in Eq. (2.2.3) over the range of

---

<sup>2.6</sup>An indicator function is defined as  $1_{\mathcal{A}}(w) = \begin{cases} 1, & \text{if } w \in \mathcal{A}, \\ 0, & \text{if } w \notin \mathcal{A}. \end{cases}$

<sup>2.7</sup>For the details of the Rayleigh distribution, see Goodman [49]; Evans et al. [41] and Kay [66].

high SNR. We refer to this subrange of signal intensity as the convergence range (or the range of convergence). The point of convergence  $S^C$  of a Rician distribution is the point of lowest signal intensity of the convergence range. The limiting distribution given in Eq. (2.2.3) may be a reasonable approximation model for the noise behavior over those regions where the signal falls within the convergence range. On the other hand, the behavior of a Rician variable does not follow closely any Gaussian distribution over the range of signal intensity lower than this point; this range of low signal intensity is referred to as the “lower than convergence” range or for short the low-intensity range. Over this range, considerable error of approximation is expected if a Gaussian noise assumption is imposed upon the data. Noise active in this range is referred to as the non-Gaussian part of Rician noise or non-Gaussian-like noise. It is crucial to determine the point of convergence for the relevant members of the Rician family, since it defines the intensity range where an approximation model, which takes the limiting distribution defined in Eq. (2.2.3) as its approximation distribution, remains reasonable.

### **2.2.1.2 Convergence Range of Rician Signals**

The probability distribution of a random variable is described by its statistical moments and the corresponding cumulants, if they exist. The first four members of these sequences measure the location and the shape of a distribution in terms of the mean, variance, skewness and kurtosis of the distribution in that order. Distributions with a finite number of lower moments or cumulants in common are expected to bear resemblance to each other. In practice, distributions with the first four moments equated result in a remarkably good approximation of one another; see Kendall et al. [67].

The convergence analysis detailed in Appendix D takes this approach to find the convergence range of Rician noise. The major results of the analysis are summarized as follows. Let  $\kappa_i$  be the  $i^{th}$  cumulant of random variable  $X$ . For a Gaussian variate where  $X \sim \mathcal{N}(\mu, \sigma^2)$ , the following is true<sup>2.8</sup>:  $\kappa_1 = \mu$ ;  $\kappa_2 = \sigma^2$ ;  $\kappa_3 = 0$ ; and  $\kappa_4 = 0$ . Thus, a Gaussian distribution is completely characterized by its first two cumulants with its third and fourth cumulants set to zero. For any random variable, the distance of its higher cumulants from the origin provides a measure of how close it is to normality. Standardized cumulants are proposed for this purpose to attain scale invariance of the measurement. The standardized skewness given by

$$\kappa_3^s = \kappa_3 \kappa_2^{-\frac{3}{2}} \quad (2.2.4)$$

and the standardized kurtosis

$$\kappa_4^s = \kappa_4 \kappa_2^{-2} \quad (2.2.5)$$

are thus important for our purpose. Given that skewness,  $\kappa_3$ , and kurtosis,  $\kappa_4$ , of a Gaussian distribution are zero, so are their standardized counterparts. Therefore the distances of the standardized skewness and kurtosis of a distribution from zero measure how much the distribution departs from normality; see Davison [33].

The corresponding cumulants,  $\kappa_1, \kappa_2, \kappa_3^s$ , and  $\kappa_4^s$  of Rician variables are required to locate the point of convergence. Appendix D shows that the

---

<sup>2.8</sup>For details, see Appendix D.

$r^{th}$  moment of a Rician variable can be expressed in terms of the Gamma function  $\Gamma(\cdot)$  and the confluent hypergeometric function<sup>2.9</sup>  ${}_1F_1(\cdot, \cdot; \cdot)$  as the following

$$m_{(r)} = (2\sigma^2)^{\frac{r}{2}} \Gamma\left(\frac{r}{2} + 1\right) {}_1F_1\left(-\frac{r}{2}, 1; -\frac{S^2}{2\sigma^2}\right). \quad (2.2.6)$$

The cumulants can be evaluated numerically from the moments; the details are discussed in Section D.1 of the appendix. It is obvious from Eq. (2.2.6) that the cumulants are functions of both  $S$  and  $\sigma$ . To be practically useful, the analysis must be applied to the relevant range of  $\sigma$  for a typical MR image. This range has been estimated<sup>2.10</sup> from empirical data of clinical MR images to fall between three and ten. It will be taken as the reference range for convergence analysis.

The cumulants of a Rician distribution associated with the signal intensity  $S$  over the convergence range should approach their Gaussian distribution counterparts. The distances between the corresponding cumulants of these distributions shrink toward zero<sup>2.11</sup> as the signal intensity approaches the point of convergence from below and stays at zero over the range of convergence. The major findings<sup>2.12</sup> are summarized as follows. As shown in Fig. 2.3, over the reference range<sup>2.13</sup> of  $\sigma$ , the function of the point of

---

<sup>2.9</sup>For details see Appendix D.

<sup>2.10</sup>For details of the estimation procedure, see Appendix C.

<sup>2.11</sup>The situation is more complicated than what this statement appears to suggest. For details, see Appendix D.

<sup>2.12</sup>The details of numerical evaluation, the method used to measure the convergence of distribution and the results are covered in Appendix D.

<sup>2.13</sup>For a discussion of the reference range of noise parameter  $\sigma$  relevant to MR images,

convergence  $S^C$  is linear<sup>2.14</sup> in  $\sigma$ . The convergence analysis based on the cumulant comparison places the upper bound of the point of convergence at roughly 50 above the point at which the approximation distribution can be used to model the noise behavior.

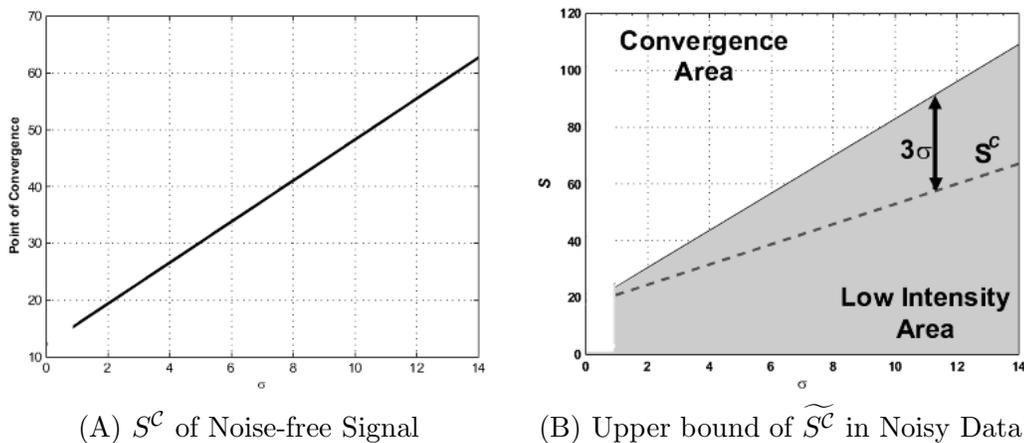


Figure 2.3: Points of Convergence over the Relevant Range of  $\sigma$ .

That the two distributions converge over a substantial range of signal intensity has important implications for our purpose: that is, over significant stretches of image sites, the stochastic behavior of image noise can be modeled by the limiting distribution. Because this distribution is a member of the Gaussian family, enhancement techniques built upon Gaussian noise models can be recruited for noise attenuation over these regions. The impact of these techniques depends on the signal intensity of the image. In practice,

see Appendix C.

<sup>2.14</sup>The linearity does not hold for  $\sigma$  close to zero. For our purpose, this low range is irrelevant.

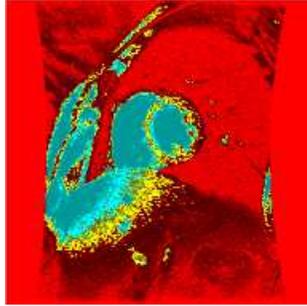
this varies from application to application. At one (hypothetical) end of the continuum, an image in its entirety falls within the convergence range, and thus a Gaussian approximation would be practically adequate for signal recovery. On the other hand, there is no role for Gaussian approximation if the target signals lie outside this range. A typical medical MR image falls in between these hypothetical cases.

It is helpful to explore empirically the role of Gaussian approximation in medical image enhancement. By the very nature of our problem, the noise-free signal  $S$  is not known a priori. A decision rule is needed to determine the point of convergence in a set of empirical data. As the signal approaches the point of convergence, the distribution of image noise becomes more and more Gaussian-like. Thus, it is reasonable to assume that the observed data lie within the range of  $3\sigma$  about the noise-free signal. As a consequence, we can put an upper bound on the point of convergence in a set of noisy data at  $\widetilde{S}^c \leq S^c + 3\sigma$ , where  $\widetilde{S}^c$  denotes the noisy realization of the point of convergence.

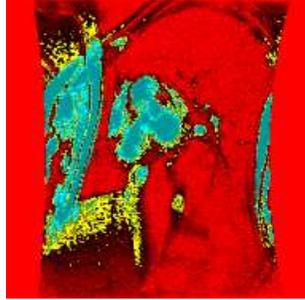
In general, the target organ is positioned in the region of signal within the convergence range against a background of features within the low intensity range. An appropriate model of image noise assuming the limiting distribution can be useful for recovery of the target structures internal to the organ. Accuracy of localization of the external boundaries of the organ may be compromised due to the low intensity signals along its interface with the background structures. Yet, images with relatively low average image intensity or high  $\sigma$  are not exceptional. These less-convenient, but still typical circumstances are illustrated in the examples shown in Fig. 2.4. The

color labels code the range of signal intensity at each image site. The red and yellow regions represent signals falling outside the convergence range – with red labels coding for data below  $S^C$  and yellow labels coding for data within  $3\sigma$  above the point of convergence. For both cardiac and non-cardiac images, the tissues of internal structures are represented by image features outside the convergence range (labeled as red and yellow regions in the images). Images (A) - (D) are examples of cardiac MRI. In the last image, the yellow region close to the center of the image represents the apex of the heart in its entirety. Image (E) is a brain image while image (F) is an MRA of carotid arteries. The small blue patches fringed with a red and yellow border are the arteries to be recovered. In all these cases, the target features are partially embedded in the low-intensity regions where non-Gaussian-like noise is active.

In general, many of the target features related to the internal structures of the organ fall within the convergence range. As a consequence, these components of the target signal can easily be enhanced using the limiting distribution of image noise. However, this approximation remains a partial solution to adaptive enhancement for three important reasons. First, even under the most favorable circumstance, the interface of the target structures with its background is usually embedded in low-intensity regions outside the convergence range. Second, in the immediate periphery of the target organ there are complex patterns of low intensity structures over a background where the non-Gaussian-like behavior of Rician noise is most active. As a consequence, the signatures of the external boundaries of the organ are obscured and their observability is reduced. Third, under less favorable circumstances, signal intensity of some internal structures is too low for an



(A)

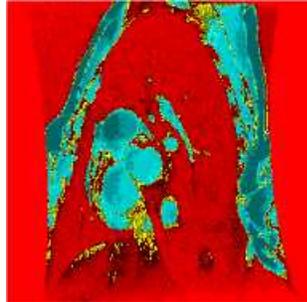


(B)

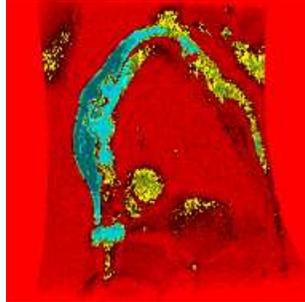
---

Cardiac MRI:  $\sigma = 8.22$

---



(C)

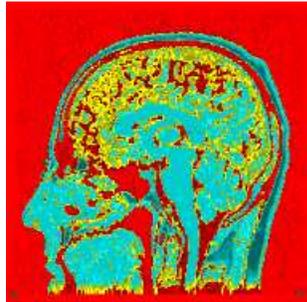


(D)

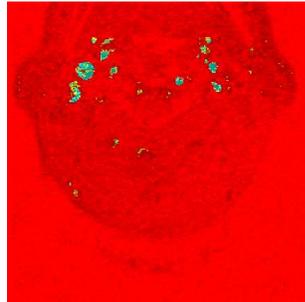
---

Cardiac MRI:  $\sigma = 9.06$

---



(E)



(F)

Brain MRI :  $\sigma = 9.09$

Cartoid Arteries MRA:  $\sigma = 7.16$

---

Figure 2.4: Signal Intensity Distribution in MR images. Red regions:  $S \leq S^C$ . Yellow regions:  $S^C < S \leq S^C + 3\sigma$ . Blue regions:  $S > S^C + 3\sigma$ .

approximation model that assumes the limiting distribution to be a valid model of the image noise active in those regions. Finally, most of the image background is characterized by a signal of very low intensity. This is also a region where the non-Gaussian behavior is most active. Given these general properties of medical MR images, it seems reasonable to use the approximation distribution to smooth out the noise components of the image signal. This smoothing step is however not adequate for our purpose. An additional enhancement operation must follow to clean up the remaining noise, remove the irrelevant structures of the background and enhance the image contrast.

### 2.2.2 Signal Properties of Squared Magnitude Representations

This section seeks a transformation of the image data that provides a better representation for our purpose. Consider a squared magnitude representation of the image data generated from the output of an MR scanner. A squared magnitude image can be generated under the pixel-wise squaring transformation on a magnitude image:  $I \longrightarrow I^S$ . Since a magnitude image is generated by a root-sum-square operation on a complex pair of images, the construction of squared magnitude images is mathematically equivalent to performing a sum-of-squares operation on a pair of images each with additive Gaussian noise with mean zero and variance  $\sigma^2$ . It is a special case of a sum-of-squares transformation of Gaussian variates, where the conditional distribution of  $I^S$  can be modeled as a noncentral  $\chi'^2$  distribution<sup>2.15</sup>:  $f(I^S|S, \sigma) \sim \sigma^2 \chi'^2(\lambda)$  with two degrees of freedom ( $\nu = 2$ ) and the

---

<sup>2.15</sup>The sum of squares of random samples from a Gaussian distribution, i.e.,  $\sum_{i=1}^{\nu} (u_i + \lambda_i)$ , where  $u_1 \dots u_{\nu} \sim \mathcal{N}(0, 1)$  are independent standard normal random variables, follows

non-centrality parameter  $\lambda = \frac{S^2}{\sigma^2}$ ; for details, see Gudbjartsson and Patz [52]; Gregg and Nowak [51]; Nowak [83] and Appendix F. In the spatial domain, the expected value of the signal intensity is given by<sup>2.16</sup>

$$\mathbf{E}(I^S) = S^2 + 2\sigma^2 \quad (2.2.7)$$

which is biased, departing from the (noise-free) signal  $S^2$ . As the bias term is no longer dependent on noise-free signals, a simple correction scheme has been suggested<sup>2.17</sup> for the spatial domain representation:

$$\widetilde{I^S} = I^S - 2\sigma^2, \quad (2.2.8)$$

where  $\widetilde{I^S}$  is the bias-corrected estimate of squared signals,  $S^2$ . A similar scheme<sup>2.18</sup> can be applied in the wavelet domain.

Given an orthogonal wavelet basis, the squared magnitude image is decomposed into a coarse scale sub-image and a set of detail representations at successive scales of increasing resolution:

$$\begin{aligned} I^S &= \underbrace{\sum_{\mu} \mathcal{S}_{j_0, \mu}^{I^S} \phi_{j_0, \mu}^{2D}}_{\mathcal{P}_v I^S} + \underbrace{\sum_{\eta} \omega_{\eta}^{I^S} \psi_{\eta}}_{\mathcal{R}_w I^S}, & (2.2.9) \\ &= \mathcal{P}_v I^S + \mathcal{R}_w I^S & (2.2.10) \end{aligned}$$

---

a central  $\chi_{\nu}^2(\lambda)$  distribution, with  $\nu$  degrees of freedom and non-centrality parameter  $\lambda = \sum_{i=1}^{\nu} \lambda_i$ ; for details, see Evans et al. [41]; Johnson et al. [62]; Stuart et al. [106] and Appendix F.

<sup>2.16</sup>See Appendix F.

<sup>2.17</sup>See Gudbjartsson and Patz [52].

<sup>2.18</sup>See Gregg and Nowak [51] and Nowak [83].

where  $\mathcal{S}^{I^S}$  and  $\omega^{I^S}$  are the approximation coefficients of the squared magnitude image and the wavelet coefficients respectively. The coarse-scale projection,  $\mathcal{P}_v I^S$ , of the squared magnitude image onto the approximation space is the polynomial approximation of the image whereas the detail representation is the projection,  $\mathcal{R}_w I^S$ , of the residues of the approximation onto the detail spaces. Given an orthogonal wavelet basis, the expected value of the wavelet coefficients in the detail sub-images equal the noise-free coefficient and therefore is unbiased; whereas, the expected value of the scaling coefficients is biased<sup>2.19</sup>. Thus, the bias term in its entirety is projected onto the coarse scale approximation.

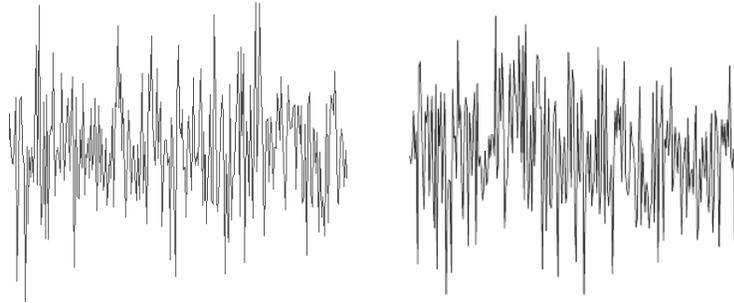
A crucial observation for our purpose is made regarding the variance of surviving noise in the squared image. The variance of  $\chi'^2$  noise in a squared image is given by<sup>2.20</sup>

$$\text{var}_{IS|S^2, \sigma} = 4\sigma^2 S^2 + 4\sigma^4 \quad (2.2.11)$$

As the variance of noise is quadratic in  $S$ , signal fluctuations due to image noise are small over the low-intensity regions but their magnitude will rapidly increase with signal intensity. This property is illustrated in Fig. 2.5. A pair of identical synthetic images corrupted by Gaussian noise with zero mean and a standard deviation ( $\sigma = 10$ ) are used to generate a squared image. The noise-free images are shown in the second row. Shown in the left column is a binary image with a white circle against a black background, while on the right is an image with fading intensity, constructed from 256

<sup>2.19</sup>See Gregg and Nowak [51] and Nowak [83].

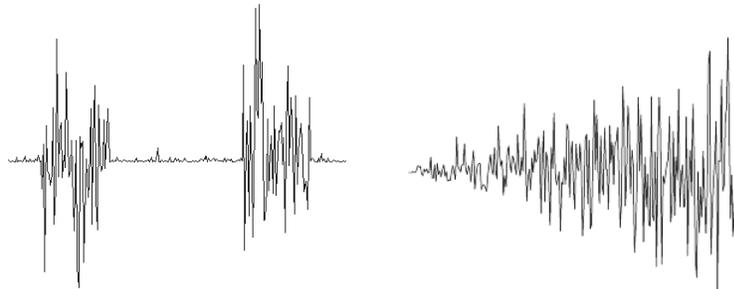
<sup>2.20</sup>See Appendix F.



Signal Fluctuations Due to Gaussian noise



Noiseless Image



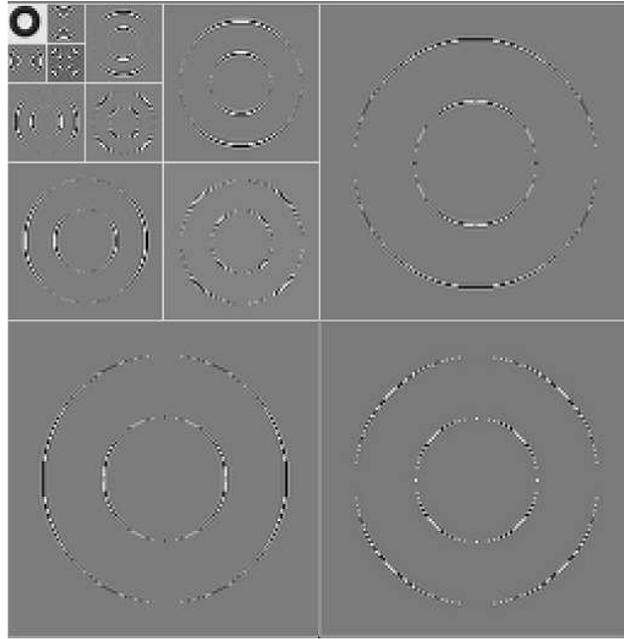
Signal Fluctuations Due to  $\chi^2$  Noise

Figure 2.5: Signal Fluctuations Due to Noise: A pair of identical synthetic images corrupted by additive Gaussian noise are used to generate a squared magnitude image. The results of two synthetic images are shown: a binary image on the left and an image with fading intensity on the right. Shown on the first row is the signal fluctuation caused by the Gaussian noise, on the second row the noise-free image with the line of reference and on the third row the fluctuation caused by the noise in the squared magnitude image.

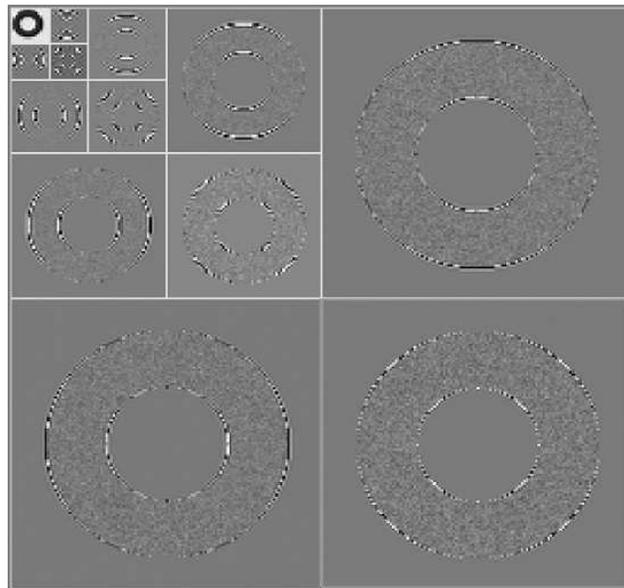
equi-luminous columns with signal intensity increasing from 0 (i.e., black) on the far left to 255 (i.e., white) on the far right. As expected, signal fluctuations due to Gaussian noise (shown in the first row) are random but self-similar over the entire line of reference. In contrast, the  $\chi'^2$  noise in a squared magnitude image causes signal fluctuations that depend on the signal intensity, as shown in the bottom row. Over the low-intensity interval of the reference line, image noise contributes only negligible perturbations around the noise-free signal. Oscillations of increasingly greater magnitude may be created outside this interval. The larger the signal intensity is, the larger the fluctuations that are caused by the image noise.

This result gives rise to the conjecture that the Rician noise outside the convergence regions is transformed in the construction of squared magnitude images into some low-frequency signal oscillation. The construction of a squared magnitude image does not remove noise, but rather it transforms noise behavior in a non-linear way. Wavelet representation allows perfect reconstruction. Signal oscillations are captured somewhere in the hierarchy of the wavelet decomposition. In the wavelet domain, low-frequency signal oscillations are captured in the coarse scale representation. According to the conjecture, image noise of squared magnitude images over the low-intensity regions does not contribute to the wavelet coefficients in any significant way.

Shown in Fig 2.6 and Fig. 2.7 are the wavelet decompositions of the squared magnitude representation of the noisy image with a ring and that with fading intensity. Wavelet coefficients capture signal discontinuities. In the case of the binary image with a ring, in the absence of any noise, as shown in Fig. 2.6(A), the significant coefficients represent the silhouette of the ring.

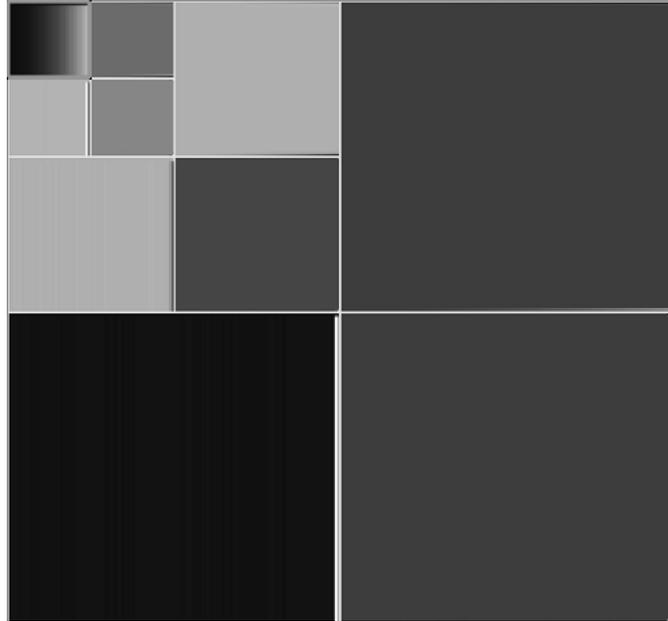


(A) Decomposition of Noise-free Image

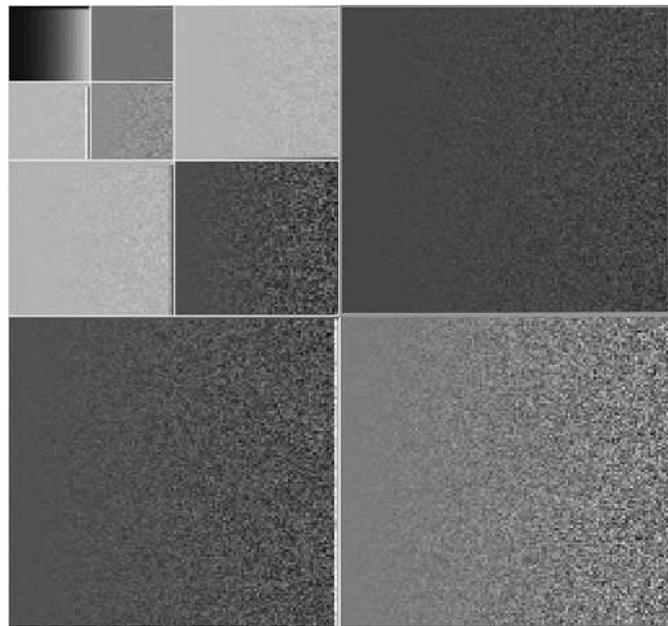


(B) Decomposition of Noisy Image

Figure 2.6: Wavelet Decomposition of the Squared Magnitude Representation of a Binary (Ring) Image.



(A) Decomposition of Noise-free Image



(B) Decomposition of Noisy Image

Figure 2.7: Wavelet Decomposition of the Squared Magnitude Representation of an Image with Fading Intensity.

Under the effect of image noise, as shown in Fig. 2.6(A), however, significant coefficients cover the entire ring but remain absent or almost absent in the low-intensity background. The image of fading intensity is created without abrupt signal transition. As a result, in the absence of noise, there are no significant wavelet coefficients anywhere in the wavelet domain representation. The coarse scale approximation gives back the original image. On the other hand, the effect of noise in the squared magnitude image is expressed as a gradual transition of increasingly greater amounts of significant coefficients over the higher intensity regions (on the left). As in the case of the binary ring, in all the detail sub-images in Fig 2.7(B), the low intensity regions are characterized by an absence of any significant coefficients.

The foregoing discussion gives rise to an assumption of this report. That is, apart from bias, the surviving noise causes small perturbations around the noise-free signal over the low intensity regions. Apart from this noise effect, the structures in the background may be assumed to be some slow oscillations over regions prevailed by low intensity signals. From this perspective, these coarse-scale oscillations, due to both the surviving noise and background structures, are similar to the effect of the long-term noise used in chemometrics to model the background interference (or drifting signals) arising in the experimental data of analytical chemistry. Noise of this class, in contrast to the well-known classes used in conventional settings, is characterized by lower frequencies than those of the analyzed signal<sup>2,21</sup>. This noise component or drifting signal can be removed by attenuating the approximation coefficients with experimentally determined factors. The estimated

---

<sup>2,21</sup>See Chau et al. [22].

coefficients  $\tilde{\mathcal{S}}$  are obtained according the the following scheme:

$$\tilde{\mathcal{S}} = \mathcal{S} - \gamma\zeta \quad (2.2.12)$$

where  $\zeta$  is an experimentally determined component attributable to the drifting signals and  $\gamma$  is the dampening factor which controls the force of noise suppression. This report will follow a similar path in setting up the enhancement scheme to solve the problem defined in the foregoing discussion.

### **2.3 Smoothing-Enhancement (SE) Approach to Adaptive Enhancement of MMR images**

The SE framework addresses the problem of adaptive enhancement in two steps – the smoothing (S) step and the enhancement (E) step. The first step (smoothing) preprocesses the image data in its magnitude representation to remove the fine scale signal oscillations according to the approximation model of the noise behavior active in high intensity regions. The second step implements an adaptive enhancement scheme on the squared magnitude representation of the preprocessed data to achieve the objectives of adaptive enhancement. The computational structure of the framework is illustrated in Fig. 2.1. This section sets out the computational details of the transformation rules of these two steps.

### 2.3.1 Smoothing Step: Attenuation of Fine Scale Signal Oscillations due to Gaussian-like Noise

The smoothing step preprocesses image data to attenuate noise using an approximation model which approximates the noise distribution by its limiting distribution. Because the limiting distribution is a member of the Gaussian family, the model allows for many existing enhancement techniques for this purpose. The choice depends on the requirement of the given application. A survey of these techniques<sup>2.22</sup> is beyond the scope of this section. The remainder of this section focuses on the basic ideas of the wavelet shrinkage; these ideas play a significant role in many of these techniques, as well as the implementation of the smoothing step in the performance evaluation of the framework, the subject matter of Chapter 3.

Many of the wavelet-based techniques are variants based on the idea of selective reconstruction<sup>2.23</sup>. These approaches start with the properties of wavelet decomposition – namely the properties of localization and energy compactness<sup>2.24</sup> of multiresolution analysis. The wavelet representation is characterized by a small number of large coefficients exclusively at the locations of rapid signal (function) oscillations and discontinuities that define coherent image structures. The presence of noise changes the empirical co-

---

<sup>2.22</sup>A brief review is provided in Appendix E.

<sup>2.23</sup>The idea can be traced back to the first wavelet-based approaches proposed by Weaver et al. [117] to reducing random noise in MR images. The wavelet shrinkage framework of signal approximation from a rigorous statistical point of view is developed independently in Donoho and Johnstone [37] and DeVore et al. [34].

<sup>2.24</sup>The property of locality refers to the space-scale localization of the structure of the signal in its wavelet expansion, whereas the property of energy compaction refers to the energy concentration of wavelet representation of the signal in a sparse number of scale-location coordinates where signal singularities occur within the support of the wavelet basis; see Appendix A.

efficients: a large number of non-zero, but small coefficients that contribute noise and a few large coefficients that contribute signal<sup>2.25</sup>. The Gaussian noise model allows a variety of coefficient selection policies proposed for different applications. Despite its diversity, most of these policies are variants based on the general idea: assuming that the low energy coefficients represent noise, the image is reconstructed with only a subset of the largest coefficients according to a threshold set.

Wavelet shrinkage refers to the reconstruction of noise-free signal (or function) from the shrunk coefficients, based on the idea that the magnitude of wavelet coefficients is related to the measure of smoothness of a signal (or function)<sup>2.26</sup>. The key question in this process is the decision rules which determine the amount of shrinkage for each wavelet coefficient. The choice of decision rules used in this step depends on the approximation requirements and the application. The rest of this section covers three families of decision rules used in the implementation of performance evaluation – namely, the SU, MSE, and SSR decision rules. A more detailed discussion of these rules and other techniques can be found in Appendix E.

### **2.3.1.1 SU Decision Rule: Wavelet Shrinkage with the Universal Threshold**

Decision rules of wavelet shrinkage usually comprise of (1) a shrinkage (or estimation) rule and (2) threshold selection. The classical exemplars of shrinkage rules with global thresholds are the hard thresholding rule and

---

<sup>2.25</sup>See Donoho and Johnstone [37], and Chang et al. [20].

<sup>2.26</sup>See Vidakovic [111].

the soft thresholding rule. The ‘keep-and-kill’ procedure leads to a hard thresholding estimator. Given a threshold  $\lambda_{\mathcal{T}}$ , the estimator of the noise-free coefficient is given by

$$\tilde{\omega}_{\eta}|_{\lambda_{\mathcal{T}}} = \begin{cases} \omega_{\eta} & \text{if } |\omega_{\eta}| > \lambda_{\mathcal{T}} \\ 0 & \text{if } |\omega_{\eta}| \leq \lambda_{\mathcal{T}} \end{cases} \quad (2.3.1)$$

On the other hand, the soft thresholding rule reduces noise by attenuating the wavelet coefficients for all coefficients larger than the threshold. The coefficient estimator  $\tilde{\omega}_{\eta}|_{\lambda_{\mathcal{T}}}$  is given by

$$\tilde{\omega}_{\eta}|_{\lambda_{\mathcal{T}}} = \begin{cases} \omega_{\eta} - \lambda_{\mathcal{T}} & \text{if } \omega_{\eta} \geq \lambda_{\mathcal{T}} \\ \omega_{\eta} + \lambda_{\mathcal{T}} & \text{if } \omega_{\eta} \leq -\lambda_{\mathcal{T}} \\ 0 & \text{if } |\omega_{\eta}| < \lambda_{\mathcal{T}} \end{cases} \quad (2.3.2)$$

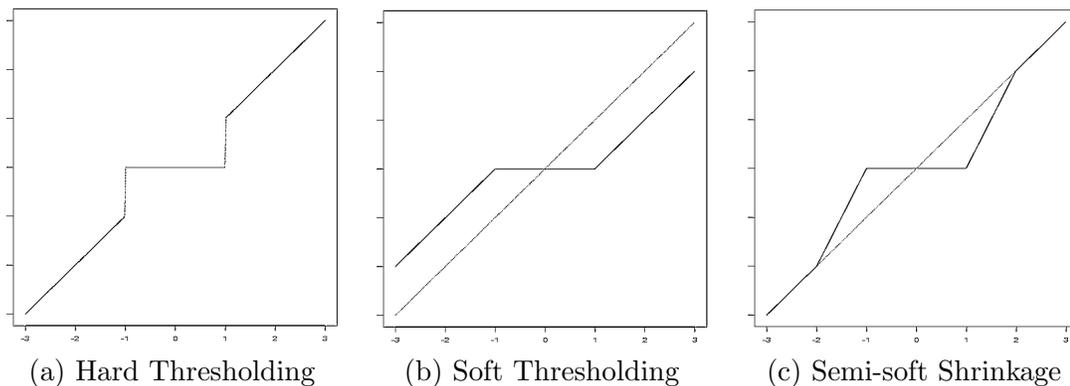


Figure 2.8: Shrinkage Functions: The horizontal axis represents an empirical wavelet coefficient and the vertical axis represents its estimated value after a shrinkage rule is applied. A diagonal line maps the empirical coefficient to itself. The distance between the functions and the diagonal line represents the amount of wavelet shrinkage.

The performance of shrinkage rules depends on the choice of threshold.

Consider the following threshold selection rule:

$$\lambda_{\mathcal{T}}^{Univ} = \sigma \sqrt{(2 \ln N)}, \quad (2.3.3)$$

where  $\sigma$  is the standard deviation of the Gaussian noise and  $N$  is the number of wavelet domain coefficients.  $\lambda_{\mathcal{T}}^{Univ}$  is known as the universal threshold; see Donoho and Johnstone [37], [38]. If it is assumed that the signal is corrupted by white Gaussian noise, the risk<sup>2.27</sup> of the soft threshold rule, given in Eq. (2.3.2), with a universal threshold, defined in Eq. (2.3.3), approaches the minimum risk up to a logarithmic factor; this minimum risk would be attainable only with an oracle that provides the critical information about which coefficients are close to zero; see Donoho and Johnstone [37, 38]; Mallat [74] and Vidakovic [111]. Moreover, the resulting approximation is no more rough than the underlying signal and much of the noise is removed from the observed data, while sharp signal discontinuities remain sharp in the approximation; see Donoho and Johnstone [37, 38] and Donoho et al. [39].

---

<sup>2.27</sup> Consider an observation

$$\mathcal{I} = \mathcal{S} + \epsilon \quad (2.3.4)$$

where  $\mathcal{S}$  is the unknown, noise-free signal, and  $\epsilon$  the additive noise. Consider an approximation (estimator)  $\hat{\mathcal{S}} \in \mathcal{F}$ , where  $\mathcal{F}$  is the class of candidate estimates of  $\mathcal{S}$ . The risk of approximation (or estimation) is the average loss that will be incurred if  $\hat{\mathcal{S}}$  is used to approximate the signal  $\mathcal{S}$

$$\mathbf{R}(\mathcal{S}, \hat{\mathcal{S}}) = \mathbf{E}_{\mathcal{S}} \mathcal{L}(\mathcal{S}, \hat{\mathcal{S}}). \quad (2.3.5)$$

A loss function is assumed to be a nonnegative real-value function such as mean-squared-loss  $\mathcal{L}(\mathcal{S}, \hat{\mathcal{S}}) = \|\hat{\mathcal{S}} - \mathcal{S}\|_{\mathcal{L}_2}^2$ ; see Korostelev and Tsybakov [69]; Krim et al. [70] and Yang [129].

The value of noise-free signal  $\mathcal{S}$  is unknown, and an estimator is sought to approximate the signal. It is preferable to use an estimator  $\hat{\mathcal{S}}$  with a small value of  $\mathbf{R}(\mathcal{S}, \hat{\mathcal{S}})$  for all values of  $\mathcal{S}$ ; that is, for any noise-free signal from the signal space, the estimator  $\hat{\mathcal{S}}$  incurs a small expected loss. By the same token, among the estimators  $\{\hat{\mathcal{S}}_1 \dots \hat{\mathcal{S}}_n\}$ , the estimator  $\hat{\mathcal{S}}_i$  is preferred to  $\hat{\mathcal{S}}_j$ , if  $\mathbf{R}(\mathcal{S}, \hat{\mathcal{S}}_i) < \mathbf{R}(\mathcal{S}, \hat{\mathcal{S}}_j)$ , for  $1 \leq j \leq n$ . For details, see Barnett [12]; Casella and Berger [17]; Korostelev and Tsybakov [69]; Stuart et al. [106] and Efromovich [40].

In practice, the variance  $\sigma^2$  is usually unknown. In this case, the standard deviation  $\sigma$  can be estimated from the wavelet coefficients at the finest scale. Assume that most of the signal oscillations captured by the observed wavelet coefficients at these subbands are noise. The standard deviation  $\sigma$  of noise may be estimated<sup>2.28</sup> using the median absolute deviation<sup>2.29</sup>(MAD) of the coefficients at the finest level<sup>2.30</sup>. Let  $\hat{\eta}$  be the subset of indices for the finest level of wavelet expansion, the estimated noise standard deviation

$$\tilde{\sigma} = \frac{MAD(\{\omega_{\hat{\eta}}\})}{0.6745} \quad (2.3.8)$$

### 2.3.1.2 MSE Decision Rule: Wavelet Shrinkage with Minimum Mean Squared Error (MSE)

A mean squared error (MSE) approach<sup>2.31</sup> estimates the noise-free component of each wavelet coefficient by minimizing the mean squared error

<sup>2.28</sup>For details, see Donoho [36] and Donoho and Johnstone [37].

<sup>2.29</sup> Given a sequence  $\mathcal{X} = \{X_i\}_{i=1}^n$  where for all  $i$  and  $j$ ,  $1 \leq i < j \leq n$ ,  $X_i \leq X_j$ , Let

$$\tilde{\mathcal{X}} = \begin{cases} X_{(\frac{n+1}{2})} & \text{if } n \text{ is odd} \\ \frac{1}{2}(X_{(\frac{n}{2})} + X_{(1+\frac{n}{2})}) & \text{if } n \text{ is even} \end{cases} \quad (2.3.6)$$

be the median of  $\mathcal{X}$ . The median absolute deviation of sequence  $\mathcal{X}$  is defined as  $MAD(\mathcal{X}) = \tilde{\mathcal{Y}}$ , where  $\mathcal{Y} = \{Y_i = |X_i - \tilde{\mathcal{X}}|\}_{i=1}^n$ .

Now consider a set of  $N$  independent Gaussian random variables,  $\mathcal{Z} = \{z_n\} \stackrel{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2)$ . Let  $\Phi$  be the cumulative normal distribution. A robust estimate of the standard deviation is given by

$$\tilde{\sigma} = \frac{1}{\Phi^{-1}(0.75)} MAD(\mathcal{Z}) \quad (2.3.7)$$

where  $\Phi^{-1}(0.75) = 0.6745$ . See Huber [56].

<sup>2.30</sup>For details, see Mallat [74]; Pizurica [88].

<sup>2.31</sup>see Gregg and Nowak [51], and Nowak [83].

(MSE) of estimation of the coefficients. Each coefficient is shrunk toward zero according to a shrinkage factor<sup>2.32</sup>:

$$\tilde{\omega}_\eta = \alpha_\eta \omega_\eta. \quad (2.3.9)$$

The optimal shrinkage factor<sup>2.33</sup> in the minimum MSE sense is given by

$$\alpha_\eta^{MSE} = \frac{\Upsilon \omega_\eta^2}{\Upsilon \omega_\eta^2 + \text{var}(\omega_\eta)}. \quad (2.3.10)$$

In practice, this optimal shrinkage factor is unknown without the perfect knowledge of the noise-free component  $\Upsilon \omega_\eta$  and the variance  $\text{var}(\omega_\eta)$  of the empirical coefficient. Gregg and Nowak [51], and Nowak [83] propose an approximation<sup>2.34</sup> of the optimal shrinkage factor:

$$\tilde{\alpha}_\eta^{MSE} = \begin{cases} \frac{\omega_\eta^2 - \text{var}(\omega_\eta)}{\omega_\eta^2} & \text{if } \omega_\eta^2 > \text{var}(\omega_\eta); \\ 0 & \text{otherwise.} \end{cases} \quad (2.3.11)$$

The shrinkage factor given in Eq. (2.3.11), bounded over the interval between zero and one, attenuates the absolute magnitude of the empirical coefficients toward zero. The variance of the coefficient can be estimated by  $\tau \sigma^2$  where  $\tau \geq 1$  and  $\sigma$  can be estimated from the image background as detailed in Appendix C.

---

<sup>2.32</sup>These factors are called filtering weights in Gregg and Nowak [51], and Nowak [83].

<sup>2.33</sup>For details, see Appendix E.

<sup>2.34</sup>For details, see Appendix E.

### 2.3.1.3 SSR Decision Rule: Semisoft Shrinkage with Rician based Thresholds

The estimator of the noise parameter  $\sigma$  plays an important role in both the SU and the MSE decision rules. While the latter decision rule estimates the noise parameter  $\sigma$  based on the Rician properties of image noise, the former adopts the robust estimator of the parameter given in Eq. (2.3.8) that assumes Gaussian noise in the data. However, in the case of magnitude MR images, noise captured in the finest scale  $\{\omega_{\hat{\eta}}\}$  is a signal-dependent Rician variable. For this reason, the assumption of Gaussian noise made by this estimator is violated. The application of the SU decision rule on an MR image for noise enhancement indeed represents an extreme case where the Rician nature of image noise is completely ignored. For comparison purposes, the rest of this section presents an alternative decision rule with global thresholds that addresses the Rician properties of image noise. The details can be found in Appendix E. The variance of image noise is signal dependent, and therefore the assumption that standard deviation is constant is violated. Given the signal-dependent nature of image noise, it is infeasible to determine the variance of the noise component of the observed signal at a particular image site. However, it is possible to estimate the range over which the variance of image noise may fall. This idea motivates the SSR decision rule based on semisoft shrinkage.

The semisoft shrinkage rule<sup>2.35</sup> is a generalized shrinkage scheme:

---

<sup>2.35</sup>For details, see Appendix E and Gao and Bruce [46].

$$\tilde{\omega}_\eta \Big|_{\lambda_{\mathcal{T}_1}, \lambda_{\mathcal{T}_2}} = \begin{cases} 0, & \text{if } \omega_{j,\mu} \leq \lambda_{\mathcal{T}_1}, \\ \mathbf{sgn}(\omega_{j,\mu}) \frac{\lambda_{\mathcal{T}_2}(|\omega_{j,\mu}| - \lambda_{\mathcal{T}_1})}{\lambda_{\mathcal{T}_2} - \lambda_{\mathcal{T}_1}}, & \text{if } \lambda_{\mathcal{T}_1} < \omega_{j,\mu} \leq \lambda_{\mathcal{T}_2}, \\ \omega_{j,\mu}, & \text{if } \omega_{j,\mu} > \lambda_{\mathcal{T}_2}. \end{cases} \quad (2.3.12)$$

where  $\mathbf{sgn}$  is the signum function<sup>2.36</sup>.

This scheme reduces to hard thresholding with  $\lambda_{\mathcal{T}_1} = \lambda_{\mathcal{T}_2}$  and to soft thresholding with  $\lambda_{\mathcal{T}_2} = \infty$ . In general, semi-soft shrinkage does not attenuate every coefficient, as shown in Fig. 2.8. The distance of a coefficient from the origin measures its significance. Any coefficient at a distance within one  $\lambda_{\mathcal{T}_1}$  from the origin is shrunk to zero. Conversely, large coefficients with an absolute magnitude greater than  $\lambda_{\mathcal{T}_2}$  from the origin are not shrunk.

As Appendix E shows, noise variance is bounded approximately over the interval:

$$\left[ \sigma \sqrt{2 - \frac{\pi}{2}}, \sigma \sqrt{2} \right].$$

These theoretical bounds motivate the following decision rule referred to as the *SSR* decision rule, or *SSR* rule for short. It is a variant of the semisoft shrinkage rule defined in Eq. (2.3.12) with the the following threshold selection rules:

$$\lambda_{\mathcal{T}_k} = \hat{\sigma}_k \sqrt{2 \ln N} \quad \text{for } k = 1, 2. \quad (2.3.14)$$

---

<sup>2.36</sup>The signum function is defined as the following:

$$\mathbf{sgn}(z) = \begin{cases} -1 & : z < 0 \\ 0 & : z = 0 \\ 1 & : z > 0 \end{cases} \quad (2.3.13)$$

see Borowski and Borwein [13].

where

$$\hat{\sigma}_1 = \hat{\sigma} \sqrt{2 - \frac{\pi}{2}}; \quad (2.3.15)$$

$$\hat{\sigma}_2 = \hat{\sigma} \sqrt{2}. \quad (2.3.16)$$

Eq. (2.3.14) can be seen as a modified version of the universal threshold selection rule in Eq. (2.3.3), but incorporating the theoretical properties of Rican noise as discussed in the foregoing. The motivation behind this formulation is to compare the performance of this rule and the  $SU$  rule – both of which are shrinkage rules using a global threshold set but that differ in a significant way with regards to the assumption of noise class. The parameter  $\sigma$  from Eq. (2.2.1) and Eq. (2.2.3) is in general unknown, but an accurate estimate,  $\hat{\sigma}$ , can be obtained from the signal-free data in the image background<sup>2,37</sup>.

### 2.3.2 Enhancement Step: Adaptive Coarse Scale Signal Enhancement

The smoothing step attenuates the noise-component captured in the detail images according to the approximation model. The image reconstructed from the coefficients given by this step is called the smoothed image. It should have the target features over the high SNR regions significantly enhanced due to noise suppression. Over the low SNR regions, however, the approximation model fails to provide a reliable approximation of the data. Surviving noise, along with other confounding features such as background structures, remains active over these regions. The discussion in Section 2.2.2

---

<sup>2,37</sup>For details see Appendix C.

suggests that these features may be modeled as coarse-scale signal fluctuations in a squared magnitude image. The tasks of the enhancement step are not only to remove these signal fluctuations due to surviving confounding features, but, of equal importance, to enhance image contrast.

A medical image is usually characterized by a number of piece-wise smoothly varying regions separated by sharp discontinuities. In a multiresolution analysis, the coarse approximation represents the smooth part of the image whereas the discontinuities are captured by the wavelet coefficients. Rapid shifts in intensity across regions contribute to image contrast essential for feature detection and structure finding. In the wavelet domain, these shifts contribute to the energy in the detail levels of representation; see Appendix A. Enhancement of the contrast can therefore be achieved by amplifying these intensity gradients. One way to accomplish this is to reduce the magnitude of the smooth parts while preserving the energy in the oscillating content captured in the detail spaces. This idea is illustrated in Fig. 2.9 with a 1D example without loss of generality.

In order to suppress background structures as well as remaining noise in the low-signal intensity proximity of the heart, the enhancement scheme in this step is necessarily adaptive to the signal intensity. The adaptive factor which controls the amount of magnitude to be reduced should be a continuous function of signal intensity to avoid creating spurious gradients.

Consider the following operator parameterized by a threshold  $\mathcal{T}$  on an

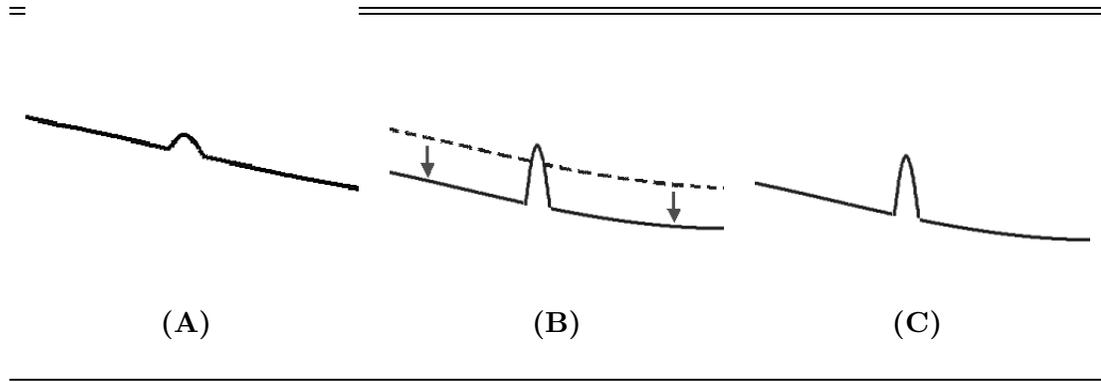


Figure 2.9: Contrast Improvement by Means of Gradient Amplification: A 1D Illustration of the Concept. (A) shows a signal as a superimposition of a smooth part and a high frequency component. A reduction of the average magnitude of the smooth part in (B) results in a higher gradient difference and consequently better contrast between the two components in (C). Multiresolution analysis of the signal provides a means to achieve this effect.

image  $I$ :

$$\Phi I = \begin{cases} \sqrt{\frac{I-\mathcal{T}}{I+\mathcal{T}}}(I + \mathcal{T}) & \text{if } \mathcal{T} > 0 \\ I & \text{if } \mathcal{T} = 0. \end{cases} \quad (2.3.17)$$

The square root term is an adaptive factor. Fig 2.10 illustrates the general behavior of the operator. The operator returns the original image,  $\Phi I = I$ , with a threshold  $\mathcal{T} = 0$ . Given a positive threshold, i.e.,  $\mathcal{T} > 0$ , the operator reduces the signal intensity of  $I$ . The numerator of the adaptive factor is reduced to zero when  $I = \mathcal{T}$  and to a negative value over the regions where  $I < \mathcal{T}$ . Due to the non-linearity of the square root operator, the adaptive factor rises rapidly for signal intensities that exceed the threshold and tends to 1 as  $I \gg \mathcal{T}$ .

Let  $(\Phi I)^{\mathcal{S}}$  denote the squared image of  $\Phi I$  and  $I^{\mathcal{S}}$  the squared image of  $I$ . The adaptive scheme given by Eq. (2.3.17) under the non-linear trans-

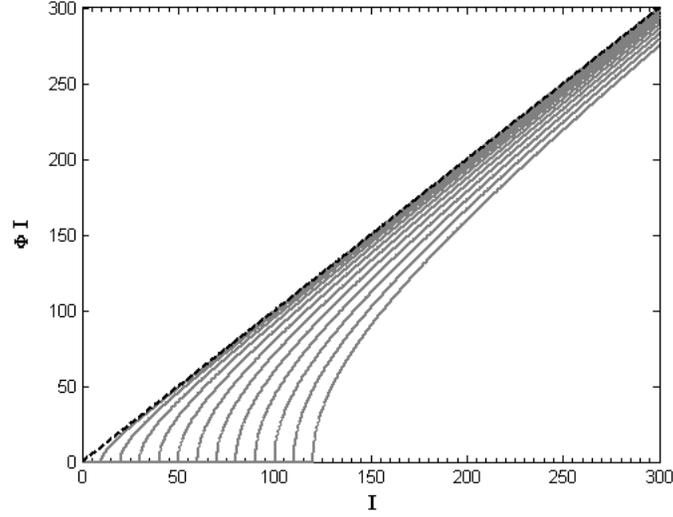


Figure 2.10: Behavior of Operator  $\Phi$  on Signal  $I$ : Each line in the diagram represents the real part of  $\Phi I$  associated with a threshold  $\tau$  set from 0 to 120 in a step of 10. The value of  $\mathcal{T}$  can be recovered from the intersection with the horizontal axis. The dashed line represents  $\Phi I = I$ , i.e.,  $\mathcal{T} = 0$  where the original signals  $I$  are returned.

formation becomes

$$(\Phi I)^{\mathcal{S}} = I^{\mathcal{S}} - \mathcal{T}^2 \quad \text{for } \mathcal{T} > 0 \quad (2.3.18)$$

In the wavelet domain,  $(\Phi I)^{\mathcal{S}}$  is projected onto the approximation and detail spaces:

$$(\Phi I)^{\mathcal{S}} = \underbrace{\sum_{\mu} \mathcal{S}_{j_0, \mu}^{(\Phi I)^{\mathcal{S}}} \phi_{j_0, \mu}^{2D}}_{\mathcal{P}_v(\Phi I)^{\mathcal{S}}} + \underbrace{\sum_{\eta} \omega_{\eta}^{(\Phi I)^{\mathcal{S}}} \psi_{\eta}}_{\mathcal{R}_w(\Phi I)^{\mathcal{S}}} \quad (2.3.19)$$

$$= \mathcal{P}_v(\Phi I)^{\mathcal{S}} + \mathcal{R}_w(\Phi I)^{\mathcal{S}} \quad (2.3.20)$$

where  $\mathcal{S}^{(\Phi I)^{\mathcal{S}}}$  and  $\omega^{(\Phi I)^{\mathcal{S}}}$  are the approximation coefficients and the wavelet coefficients of  $(\Phi I)^{\mathcal{S}}$  respectively. Raising the data to a higher power amplifies the image oscillation; consequently, image features in a squared image are

projected onto a subspace at the scale at least as high as the corresponding scale in the decomposition of the magnitude image. The wavelet representation of  $(\Phi I)^{\mathcal{S}}$  preserves the image features in the detail spaces in the sense that those features in  $\mathcal{R}_w \Phi I$  stay in the detail spaces in the decomposition of the squared image  $\mathcal{R}_w (\Phi I)^{\mathcal{S}}$ . Because the approximation coefficients represent the average intensity in their local neighborhoods within the compact support, the “micro-features” of an image can be preserved and enhanced by reducing the average intensity of the image while preserving the integrity of the details.

The problem with the enhancement step may be solved by applying the signal adaptive operator on only the coarse scale approximation of the image without altering the detail representation of the squared image. This approach has been emerging from various threads of the foregoing analysis leading to the following working assumptions:

1. Both the surviving noise and the background structures can be modeled as some coarse scale (or long term) oscillations. From the perspective of adaptive enhancement, both are seen as noise captured in the coarse scale representation.
2. Image contrast can be enhanced by reducing the average intensity in the local neighborhoods of the coefficients.

Consider applying the enhancement operator specified in Eq. (2.3.18) to  $\mathcal{S}^{(\Phi I)^{\mathcal{S}}}$ , the set of approximation coefficients, exclusively. The resulting representation yields the following modified (squared) image:

$$\widetilde{I}^{\mathcal{S}} = \mathcal{P}_v (\Phi I)^{\mathcal{S}} + \mathcal{R}_w I^{\mathcal{S}}. \quad (2.3.21)$$

The first term  $\mathcal{P}_v \Phi I^{\mathcal{S}} = \sum_{\mu} \mathcal{S}_{j_0, \mu}^{\Phi I^{\mathcal{S}}} \phi_{j_0, \mu}^{2D}$  is the projection of  $\Phi I^{\mathcal{S}}$  on the approximation space. In an orthonormal wavelet system, each approximation

coefficient in a  $J$ -scale decomposition can be expressed using Eq. (2.3.18) as follows:

$$\mathcal{S}_{J,\mu}^{(\Phi I)^S} = \left\langle (\Phi I)^S, \phi_{J,\mu}^{2D} \right\rangle \quad (2.3.22)$$

$$= \sum_{\mu} (I^S - \mathcal{T}^2) \phi_{J,\mu}^{2D} \quad (2.3.23)$$

$$= \sum_{\mu} I^S \phi_{J_0,\mu}^{2D} - \sum_{\mu} \mathcal{T}^2 \phi_{J,\mu}^{2D}. \quad (2.3.24)$$

$$= \mathcal{S}_{J,\mu}^{I^S} - \sum_{\mu} \mathcal{T}^2 \phi_{J,\mu}^{2D}, \quad (2.3.25)$$

where  $\mathcal{S}^{I^S}$  is the corresponding approximation coefficient of the squared image. Using the fact that  $\sum_{\mu} \phi_{J,\mu} = 2^J$  for a  $J$ -scale decomposition<sup>2,38</sup>, Eq. (2.3.25) is therefore

$$\mathcal{S}_{J,\mu}^{(\Phi I)^S} = \mathcal{S}_{J,\mu}^{I^S} - 2^J \mathcal{T}^2 \quad (2.3.26)$$

Define the attenuation factor

$$\tau = 2^J \mathcal{T}^2, \quad (2.3.27)$$

For every approximation coefficient  $\mathcal{S}_{J,\mu}^{I^S}$ , the attenuated coefficient is

$$\mathcal{S}_{J,\mu}^{(\Phi I)^S} = \mathcal{S}_{J,\mu}^{I^S} - \tau \quad (2.3.28)$$

The attenuation operation given in Eq. (2.3.28) reduces the average intensity over the local neighborhoods in the squared smoothed image. The attenuated approximation is re-integrated in the synthesis step with the details at each scale to yield an enhanced squared image  $\widetilde{I^S}$ , according to Eq. (2.3.21).

---

<sup>2,38</sup>See Burrus et al. [15].

Attenuation applied to a squared coarse-scale approximation achieves intensity adaptive dampening of signal oscillation. As a result, the signal energy over high intensity regions is preserved, whereas the low intensity oscillations are attenuated. The non-positive parts of the modified image represent those image features being suppressed, such as those features in the very low intensity regions under the full force of attenuation. The enhanced magnitude image is therefore obtained by the following thresholding operation:

$$I_{se} = \sqrt{\widetilde{I^S} 1_{\{\widetilde{I^S} > 0\}}} \quad (2.3.29)$$

The enhanced image depends on the threshold set for the adaptive enhancement rule. As mentioned, the scheme adaptively reduces signal intensity of the coarse scale approximation over low-intensity regions. With a threshold set at a level which discriminates these structures and the boundaries of the heart, the adaptive scheme suppresses most of these structures in the proximity of the heart. This critical range is a characteristic of MRI scanning. It does not change significantly over a temporal sequence.

The remaining noise after the smoothing step contributes to some small fluctuations in the approximation and wavelet coefficients of the squared magnitude image. A study<sup>2,39</sup> of the distribution of the scaling function coefficients of squared magnitude images suggests that a threshold be set at  $(2^{J+1} + 6)\sigma^2$  to discriminate between coefficients with the support of the scaling functions in the signal regions and those in the Rayleigh (low-intensity) regions. To suppress the spurious fluctuations in the coefficients over low intensity regions, we set a threshold, which should be proportional

---

<sup>2,39</sup>See Wu et al. [127].

to the variance of the noise, at or above this level. Consider a threshold set to  $\zeta\sigma$ . Where  $\zeta = \sqrt{2}$ , attenuation factor  $\tau_n = 2^{J+1}\sigma^2$  is equivalent to the threshold set for removing bias from the approximation coefficients<sup>2.40</sup>. For most practical choices of  $J$ , setting  $\zeta$  to two is adequate to discriminate between the approximation coefficients with support inside signal-free regions and those with support over signal regions of the image. A higher  $\zeta$  raises the thresholding level to offset larger fluctuations in the coefficients. Our experimental findings indicate that the background noise can be more thoroughly cleaned up with  $\zeta$  set to three. In general, the result is not sensitive to minor variations in the range between two and five.

## 2.4 Summary

This chapter, with the details covered in the accompanying appendices, presents an analysis of the problem of adaptive enhancement of magnitude MR images, providing the essential observations and hypotheses that guide the computational mechanism developed to achieve the objectives of the SE framework. Rician noise deviates in behavior from Gaussian noise over low signal intensity ranges, but, as the SNR increases, its distribution converges to a limiting distribution, which belongs to the Gaussian family. A meaningful distinction can be made between image areas of high signal intensity and those of low signal intensity regions in terms of the signal behavior of confounding features in these areas. Over high signal intensity regions, signal fluctuations due to noise can be attenuated using Gaussian-based enhancement techniques, whereas the confounding structures in the low intensity regions are captured in the coarse scale approximation of the

---

<sup>2.40</sup>This is the wavelet domain equivalence of the bias correction scheme given in Eq. (2.2.8). See Gregg and Nowak [51]; Nowak [83].

squared representation in the wavelet domain. Moreover image contrast can be significantly enhanced by attenuating the average signal intensity of the smoothly varying component of the signal while preserving the integrity of its fine scale structures. In the light of this observation, the solution to the problem of adaptive enhancement is necessarily adaptive to both signal intensity and image structures. An orthogonal wavelet system provides a convenient and efficient framework of signal analysis, transformation, and synthesis to achieve the objectives of adaptive enhancement.

The wavelet-based framework, called the SE framework, summarized in Fig. 2.1, achieves adaptive enhancement of an MR image in two steps, each of which applies the basic procedures of analysis–transformation–synthesis to different image representations. The smoothing step attenuates the high-frequency fluctuations due to Gaussian-like noise in a magnitude image. Noise reduction techniques based on a Gaussian noise model can be adopted for this purpose; the choice of technique is not dictated by the framework but rather by the requirements that arise in the application domain. For a performance evaluation of the framework, three decision rules of wavelet shrinkage, namely the SU, MSE, and SSR rules, are implemented in this study. The enhancement step applies the adaptive enhancement procedures on the squared smoothed signal to suppress the confounding structures in the coarse scale approximation, as well as to enhance image contrast. The transformation rule described in Eq. (2.3.28) and the synthesis rule in Eq. (2.3.29) in this step achieve the essential transformation required for the adaptive enhancement of magnitude MR images.

## Chapter 3

# Experimental Results

### 3.1 Methodology of Performance Evaluation

This chapter provides experimental evidence for the performance of the smoothing-enhancement (SE) scheme applied to magnitude MR images. The aspects of performance under scrutiny should be closely related to the objectives of adaptive enhancement, that is, (1) reduction of image noise; (2) attenuation of background structures in the immediate periphery of the target organ; (3) preservation of target features; and (4) enhancement of image contrast. Meaningful performance evaluation<sup>3.1</sup> of the algorithm can be obtained through a comparison of its output with (1) the original input and (2) with the output of other enhancement procedures in both quantitative and qualitative terms.

To address the questions raised in this study, seven different enhance-

---

<sup>3.1</sup>Another method is to compare the results of the test algorithms with some “gold standard” or “the ground truth”. In the case of medical images of animal/human tissues and organs, the “ground truth” is usually defined by experts of the field. Due to the unavailability of expert-defined data at the time of experiments, this research does not follow this avenue for the test cases that involve medical images.

ment procedures falling into three main groups are applied to a set of test images for comparative evaluation of performance. Three SE proce-

Table 3.1: Enhancement Procedures Implemented for Performance Evaluation

		Wavelet Domain	Wavelet Domain	Spatial Domain
		Gaussian *	Rician *	
Wavelet Domain	Smoothing Enhancement	SE - SU	SE - MSE SE - SSR	
Wavelet Domain	Smoothing only	S - SU	S -MSE	
Spatial Domain	Smoothing only			Gaussian Anisotropic Diffusion

\* Noise model used for  $\sigma$  estimation.

dures, collectively referred to as the SE-group, shown in the first row in Table 3.1, are implemented. They follow the two-step enhancement scheme (see Fig. 2.1) but differ in the wavelet-based smoothing techniques used in the first step. (1) The SE-MSE procedure filters the wavelet coefficients of signals with an MSE shrinkage rule as defined in Eq. (2.3.9) with the shrinkage factor given by Eq. (2.3.11). (2) The SE-SSR procedure employs the semisoft shrinkage rule given in Eq.(2.3.12) with thresholds  $\lambda_{T_1}$  and  $\lambda_{T_2}$  as defined in Eq. (2.3.14)-Eq. (2.3.16). Both of these shrinkage rules use a Rician-based estimator<sup>3.2</sup> of the image noise parameter  $\sigma$ . (3) The SE-SU

<sup>3.2</sup>For details, See Appendix C.

implements the wavelet shrinkage rule<sup>3.3</sup> with the universal threshold as defined in Eq. (2.3.3). The threshold is based on the *MAD* estimator of  $\sigma$  as given in Eq. (2.3.8) under the assumption of Gaussian noise. The inclusion of the last algorithm provides some empirical evidence for the performance of an algorithm which completely ignores the Rician nature of MR image noise.

A major focus of this study is the contribution of the enhancement step to the adaptive enhancement of a magnitude MRI. The S-group of procedures implements only the smoothing step without further enhancement, that is, the second step of the SE framework is turned off. The S-SU algorithm implements the wavelet shrinkage rule with a universal threshold and the S-MSE algorithm with the MSE shrinkage rule. As its counterpart in the SE-group, the S-SU algorithm represents the smoothing procedure working under the assumption of Gaussian noise.

The performance of the SE-procedures is also compared with that of spatial domain enhancement algorithms. Two widely-employed enhancement algorithms of this class, Gaussian filtering and anisotropic diffusion filtering, are chosen, constituting the third group of procedures – collectively referred to as the Diffusion Group. Gaussian filtering<sup>3.4</sup> smooths signal oscillations through linear diffusion of signal intensity with a Gaussian kernel. This is a classical smoothing technique widely used explicitly for image enhancement or implicitly embedded in higher-level processing procedures.

---

<sup>3.3</sup>Both hard and soft thresholding rules are used in the experiments. Since the results are not sensitive to the choice of the rules, only the results from the implementation using the soft thresholding rules are reported.

<sup>3.4</sup>For details, see Trucco and Verri [108].

To better preserve coherent features, the second algorithm<sup>3.5</sup> employs an anisotropic diffusion kernel to guide the diffusion process so that the full force of smoothing is directed along the edges and is kept from crossing the intensity gradients of image features.

The output images of these procedures are compared according to both quantitative and qualitative criteria with respect to the objectives of the problem of adaptive enhancement. Quantitatively, the output images are evaluated with respect to (1) contrast enhancement and (2) the fidelity (goodness) of approximation. Contrast enhancement refers to the improvement of image contrast resulting from an enhancement procedure applied to an image. Visual contrast is defined as the ratio  $\frac{\max - \min}{\max + \min}$  where max is the maximum of the grading intensity and min the minimum of two testing gratings; see Pratt [94]. Let  $\mathcal{F}$  be the foreground and correspondingly  $\mathcal{B}$  the background of a region of interest – a region selected for comparison. The contrast of an image  $\Xi$  is measured by

$$C_{\Xi} = \frac{\bar{L}_{\Xi, \mathcal{F}} - \bar{L}_{\Xi, \mathcal{B}}}{\bar{L}_{\Xi, \mathcal{F}} + \bar{L}_{\Xi, \mathcal{B}}} \quad (3.1.1)$$

where  $\bar{L}_{\Xi, \Omega}$  is the average pixel intensity of the region  $\Omega$ . Given the input  $I_{in}$  and the output  $I_{out}$  of an enhancement procedure, enhancement improvement resulting from an operation may be measured in terms of a contrast improvement ratio (CIR) defined as follows:

$$CIR = \frac{\mathcal{C}_{out} - \mathcal{C}_{in}}{\mathcal{C}_{out}} \times 100\%, \quad (3.1.2)$$

---

<sup>3.5</sup>For the details on anisotropic diffusion filtering, see Weickert [118, 119].

where  $\mathcal{C}_{in}$  measures the image contrast of an input image  $I_{in}$  and  $\mathcal{C}_{out}$  that of the corresponding output  $I_{out}$ . This is a measurement metric used in the literature.

Implicit in Eq. (3.1.1) and Eq. (3.1.2) is the assumption of comparison that the operation preserves the average image intensity of the image over the region of interest. The measurement may be misleading when this assumption does not hold. In an extreme case, a reduction in the average signal intensity of an image may create a spurious contrast improvement due to a smaller normalization factor  $\bar{L}_{\mathcal{F}} + \bar{L}_{\mathcal{B}}$  in the denominator of Eq. (3.1.1) without any significant change in the intensity difference across the boundary of the foreground objects. In the case where the scaling coefficients are attenuated in any fashion such as bias correction, CIR as defined in Eq (3.1.1) - Eq. (3.1.2) is biased in favor of the output. To correct this bias, an alternative measurement metric is employed – called adjusted contrast measurement where both input and output images use the same normalization factor:

$$\mathcal{C}_{\Xi}^{adj} = \frac{\bar{L}_{\Xi,\mathcal{F}} - \bar{L}_{\Xi,\mathcal{B}}}{\bar{L}_{in,\mathcal{F}} + \bar{L}_{in,\mathcal{B}}} \quad \Xi \in \{\text{in, out}\} \quad (3.1.3)$$

and correspondingly the adjusted contrast improvement ratio

$$CIR^{adj} = \frac{\mathcal{C}_{out}^{adj} - \mathcal{C}_{in}^{adj}}{\mathcal{C}_{out}^{adj}} \times 100\%, \quad (3.1.4)$$

Both  $CIR$  and  $CIR^{adj}$  are reported. The interpretation of the empirical findings primarily focuses upon the latter metric. The former, however, is

useful as a reference for comparison since previous research<sup>3.6</sup> uses this metric to study performance.

Both  $CIR$  and  $CIR^{adj}$  measure the improvement in terms of the difference in image contrast between images. Note that the term “improvement” in this context does not necessarily imply that the contrast of the output image is better than that of the input image. As summarized in the following table, a negative value indicates a worsened contrast in the output image, and zero improvement indicates no change.

Contrast of output relative to input	Better	No Change	Worse
$CIR$ ; $CIR^{adj}$	$> 0$	$0$	$< 0$

Let  $\mathcal{I}_S$  be the noise-free image. The goodness of approximation is measured in terms of two metrics: the absolute error measurement with respect to the  $l_1$ -norm Pratt [94]

$$\|\mathcal{I}_S - \mathcal{I}_{out}\|_{l_1} = \frac{1}{N} \sum |\mathcal{I}_S - \mathcal{I}_{out}|, \quad (3.1.5)$$

where  $N$  is the data size, and the Signal-to-Noise Ratio (SNR) with respect to the  $l_2$ -norm Pratt [94]

$$SNR = -10 \log(\xi_{l_2}) \quad \text{where} \quad \xi_{l_2} = \frac{\sum |\mathcal{I}_S - \mathcal{I}_{out}|^2}{\sum |\mathcal{I}_S|^2} \quad (3.1.6)$$

According to Eq. (3.1.5) and Eq. (3.1.6), errors of approximation are mea-

---

<sup>3.6</sup>For example, Gregg and Nowak [51] and Nowak [83].

sured with respect to the noise-free signal  $\mathcal{I}_S$  which is in general an unknown quality of a real cardiac image. Instead, synthetic images are used to evaluate this aspect of performance<sup>3.7</sup>.

Qualitatively, the output image may be evaluated by human inspection with respect to the major performance criteria: (1) attenuation of noise; (2) preservation of the target structures; and (3) reduction of background structures. For this purpose, three representations (or views) of a set of image data, either an input image or an output of an enhancement procedure, are presented to make these image qualities discernible. Fig. 3.1 shows an example for a cardiac image. Referred to as the ‘data view’ in (A) is an

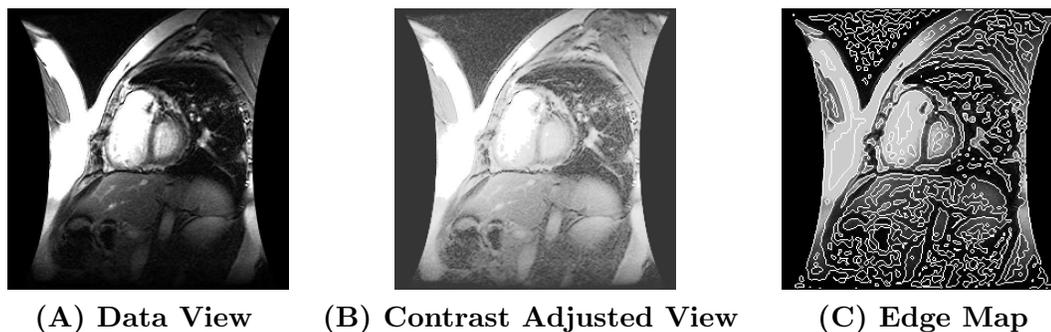


Figure 3.1: Image Representations for Performance Evaluation.

image obtained from the image acquisition processes (in the case of input) or from the enhancement procedures (in the case of output) without any further processing. With this representation, many features visible to computer algorithms, including noise, become indiscernible to a human viewer; thus,

---

<sup>3.7</sup>To isolate errors of approximation from the enhancement improvement, the output of the SE-group is adjusted for the intensity shift that contributes to the enhancement measurement. This intensity shift is computed from the output resulting from the application of the enhancement step to the signal-free image. This adjustment is not necessary for the output of other groups of procedures which do not change the average intensity of the image.

this view does not facilitate judgment on the difference in quality between two images. To illustrate the problem, consider the seemingly homogeneous black region at the top and the regions on either side of image (A) in Fig. 3.1. Unlike the signal-free background at the top, the dark regions on both sides are not part of the image. This difference is indiscernible since image noise is invisible. Contrast adjustment, such as histogram equalization<sup>3.8</sup>, is applied to the data view to produce a ‘contrast adjusted image’ or ‘adjusted image’ for short – as illustrated in image (B) of Fig. 3.1. This image provides a better view of image features such as image noise, the texture of organs and the background structures behind the heart in the thoracic cavity. Illustrated in image (C) of the figure is a map of edges superimposed on the data view of the input image. This map represents the structures (including spurious ones) detectable by edge operators based on the zero crossings of an edge Laplacian<sup>3.9</sup> of image data.

## 3.2 Performance Evaluation

The test algorithms are run on both synthetic images and medical images. The evaluation metrics are computed from both the input and output of the test algorithms. In the case of medical images, a contrast adjusted image and an edge map are also generated for both the input and output of the test algorithms. In order not to distract attention from the discussion of major findings, these images are presented in Part II of the appendix.

---

<sup>3.8</sup>For details, see Pratt [94], and Watt and Policarpo [116].

<sup>3.9</sup>For details, see Pratt [94].

### 3.2.1 Synthetic Images

Each synthetic image used in the experiments is a magnitude image constructed from a pair of identical (noise-free) images corrupted by Gaussian noise with zero mean and identical variance across images. As shown in

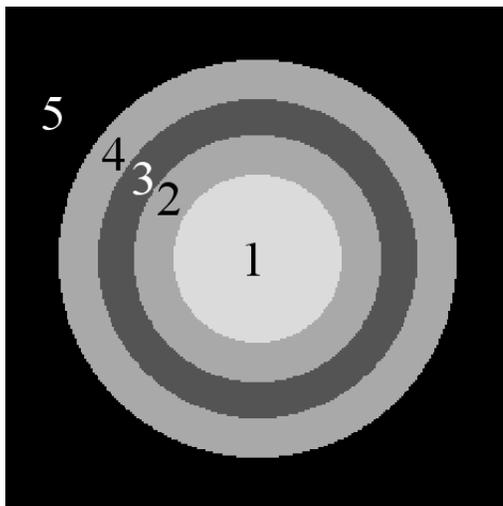


Figure 3.2: Noise-free Image with Region Indices.

Fig. 3.2, the noise-free image is a piece-wise step function in the form of concentric rings, each with homogeneous signal intensity. To facilitate discussion, the homogeneous regions of the image are numbered in increasing order from the center of the concentric rings to the background of the image as shown in Fig. 3.2. This is designed to simulate the essential parts of a CMR image, (a) region 1 represents the region inside a heart chamber; (b) region 2 the heart wall; (c) region 3 the thoracic cavity; (d) region 4 the thoracic wall; and (e) region 5 the background outside the human body. A sequence of experiments is carried out on noisy input generated with different realizations of Gaussian noise with different values of  $\sigma$  selected from

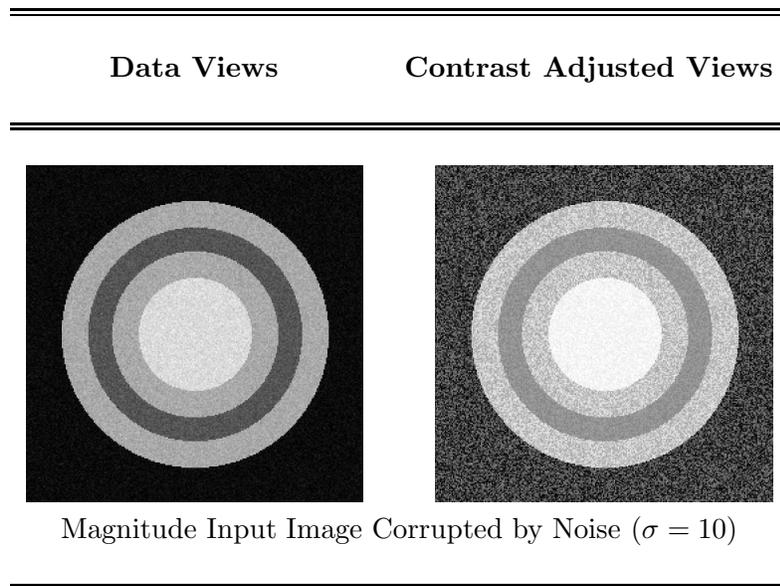


Figure 3.3: Synthetic Image: A Specimen of Noisy Input.

the reference range<sup>3.10</sup>. An example of noisy input is shown in Fig. 3.3.

First, fifty noisy input images are generated with noise parameter  $\sigma = 10$  for performance evaluation. Consider the improvement on image contrast. The findings are consistent across different regions selected from various parts of the images for comparison. The findings shown in Fig. 3.4 result from experiments using a region across Ring 1 and Ring 2, a region of obvious significance since its background represents the heart wall and the foreground the cardiac chamber. The image in Fig. 3.4(A) is an example of a noisy input image superimposed with the selected region. The diagrams on the right of the figure display the performance of the algorithms with respect to the performance metrics; each bar represents the range of measurement obtained from fifty repeated trials, bounded at the top and bottom

<sup>3.10</sup>The reference range is between 3 and 10; for details, see Appendix C.

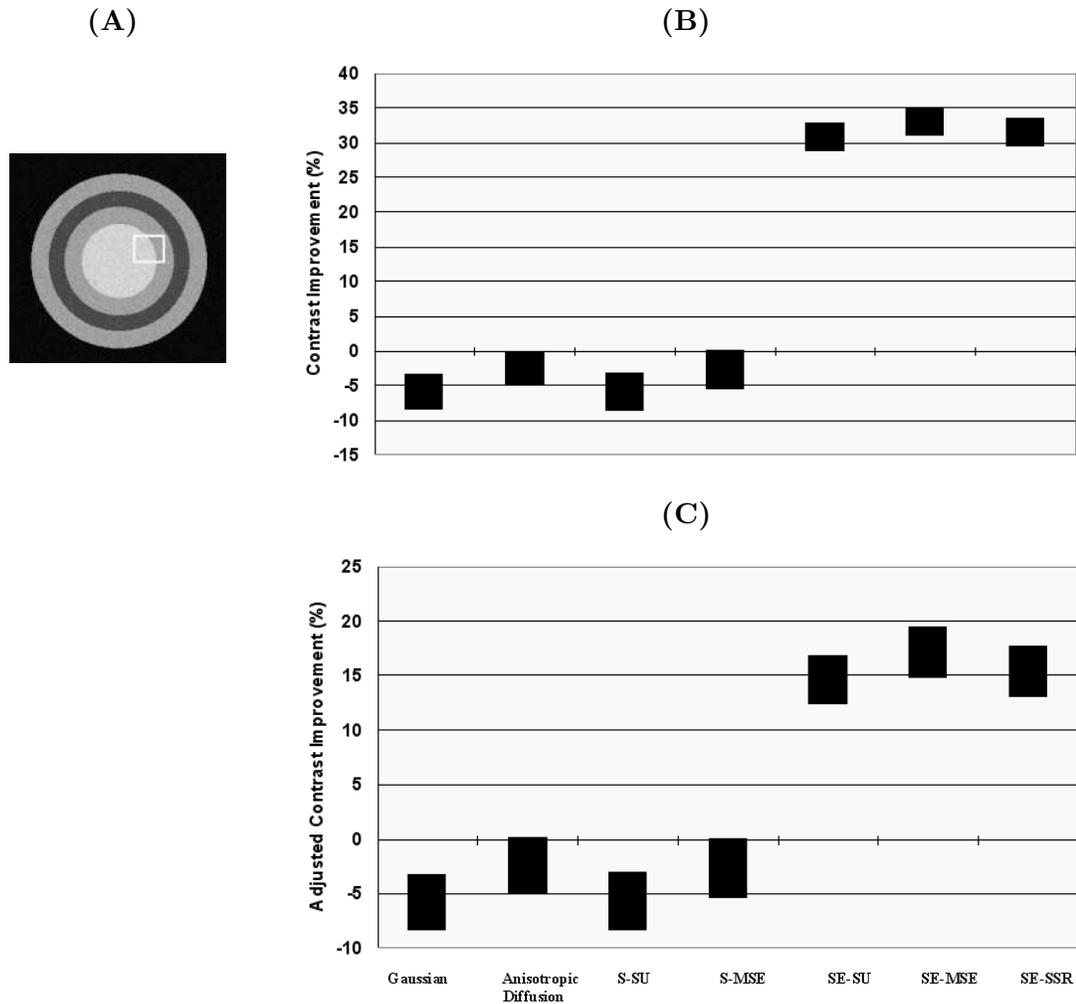


Figure 3.4: Synthetic Image with Noise ( $\sigma = 10$ ). (A) The region of interest (ROI) selected for comparison, (B) Contrast improvement ratio  $CIR$  and (C) Adjusted contrast Improvement ratio  $CIR^{adj}$ . The height of the bars represents the distance between the maximum and minimum of the measurement.

respectively by the maximum and minimum contrast improvement ratio.

Significant differences are observed between SE-group algorithms and the other groups of algorithms. All SE-procedures provide significant (positive) improvement in the contrast over those without the enhancement step. Signal smoothing in the other groups, however, reduces noise only at the expense of image contrast. The negative ratios attained by both the diffusion and S-groups attest to this trade-off. The intra-group difference indicates that the anisotropic diffusion filtering and MSE filtering provide the best results in their respective groups.

Similar inter-group differences are observed in approximation error with respect to the  $l_1$  norm between the SE-group and the other two groups. The procedures in the first group, as shown in Fig. 3.5(A), produce much lower absolute error than those in the other groups. It is however the performance comparison in terms of signal-to-noise ratios with respect to the  $l_2$ -norm that provides informative evidence about the role of the Rician noise model in adaptive enhancement. The SE-MSE and SE-SSR procedure out-perform all the other algorithms whereas the output from the SE-SU algorithm achieves the lowest SNR. Unlike the *MSE* and *SSR* scheme, the *SU* scheme completely ignores the stochastic properties of Rician noise. The threshold based on the median of absolute deviation (MAD) estimate of  $\sigma$  leads to suboptimal shrinkage of coefficients, causing a higher discrepancy between the smoothed image and the noise-free signal.

The foregoing results represent the general pattern that holds for a broad spectrum of noise with  $\sigma$  ranging from 3 - 10. Fig. G.1 and Fig. G.2 in

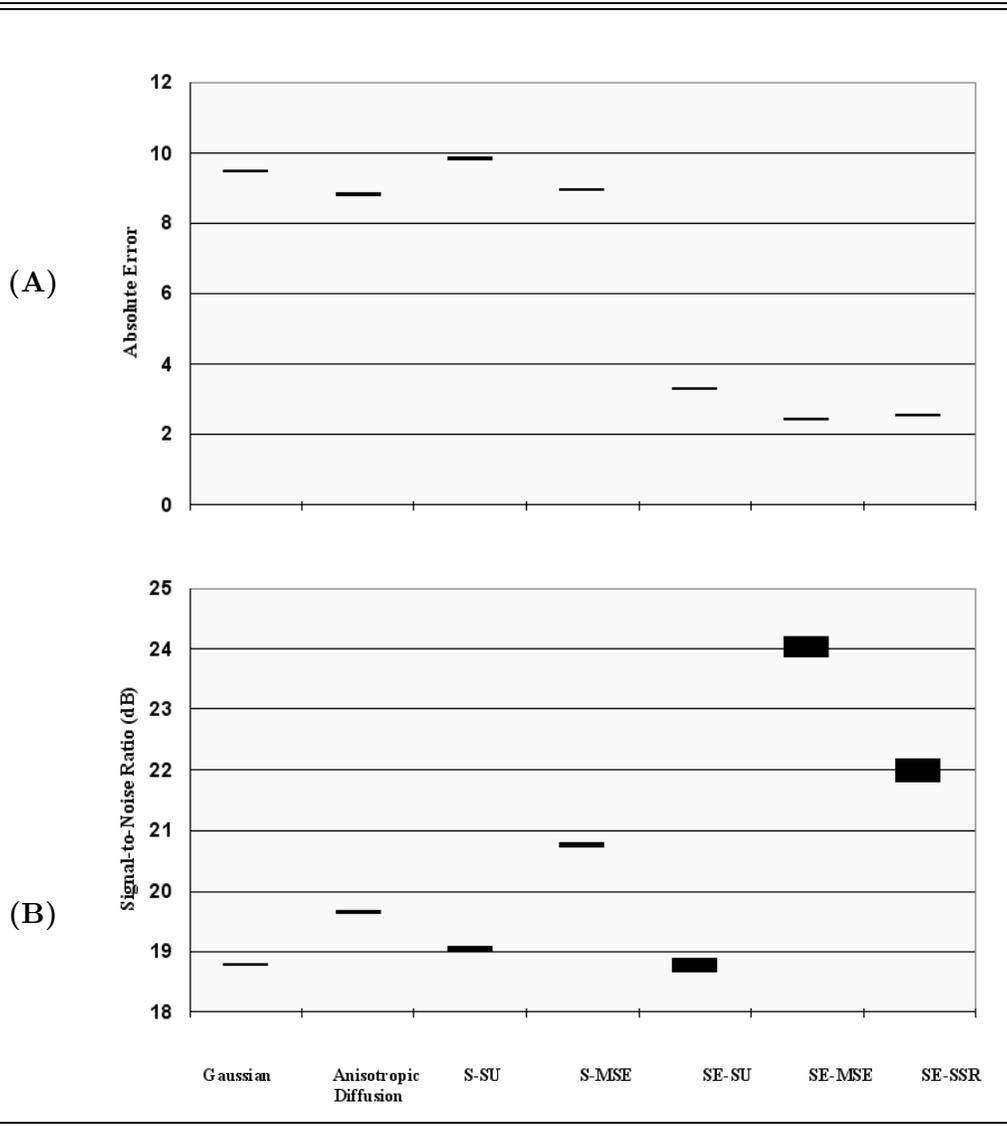


Figure 3.5: Synthetic Image with Noise ( $\sigma = 10$ ) : Absolute Error with respect to the  $l_1$  norm and SNR with respect to the  $l_2$  norm.

Part II of the appendix summarize the results of experiments using noise with different values of  $\sigma$ . The SE algorithms stand out as the only group that achieves significant improvement in image contrast. Estimation errors increase with the noise parameter  $\sigma$ . In terms of the  $l_1$ -norm, the absolute error of the SE-group is significantly lower than the error of the other groups. The inter-group difference from these other groups widens as  $\sigma$  is increased. Over the whole range, the SE-MSE and SE-SSR scheme achieve a higher SNR than the other algorithms.

### 3.2.2 Clinical Magnitude MR images

Different sequences of medical images<sup>3,11</sup> are used for performance evaluation. The results are consistent over all test cases. This section presents the findings from a subset of test cases to highlight the important aspects of the comparative performance of the algorithms. Summarized in Table 3.2

Selected Test Cases		Parameters		Associated Output Figures in Part II of the Appendix
ID	Type	$\sigma$	$\mathcal{T}$	
SRS0000-10	Cardiac	5.85	30	Fig. H.1 – Fig. H.3
SRS0000-70	Cardiac	5.85	35	Fig. H.4 – Fig. H.6
Eman	Cardiac	8.70	26	Fig. H.10 – Fig. H.12
SRS0006-24	Cardiac	4.21	50	Fig. H.13 – Fig. H.15
BG-3	Brain	8.80	44	Fig. H.16 – Fig. H.18
CA-2	Cardiovascular	7.34	51	Fig. H.19 – Fig. H.21

Table 3.2: Selected Test Cases for Performance Evaluation

is a set of six test cases selected for discussion. The first four test cases are cardiac images from three different sequences. The first two cases se-

---

<sup>3,11</sup>The cardiac and the brain data sets, acquired using a GE Medical Systems Genesis Signa MRI system, with all the information about the identities of the patents removed, were provided by the Hospital for Sick Children in Toronto.

lected from the SRS0000 sequence represent two cross-sections of the heart: SRS0000-10 captures the upper section (toward the base) of the heart and SRS0000-70 the middle section (at the level of the ventricles). The Eman sequence and SRS0006 sequence represent two extreme cases. Among all the test image sequences, the Eman sequence is characterized by its low average signal intensity with the highest noise parameter  $\sigma$  whereas the SRS0006 sequence consists of images with high mean signal intensity and relatively low  $\sigma$ . While the test images from the SRS0000 sequence are captured at the beginning of systole (contraction phase), the other two are the snapshots in the diastole (relaxing phase) of the cardiac cycle. To show the applicability of the scheme to other classes of medical MR images, this section also includes two non-cardiac images: BG-3 is an MRI scan of a brain and CA-2, an MRA<sup>3,12</sup> image of carotid arteries, is a cross-section of a neck.

Test Cases	Selected Regions	Reference Figures
SRS0000-10	Heart: atria, aortic root, & arteries	Fig. 3.6
SRS0000-70	Left ventricle	Fig. 3.7
Eman	Heart: ventricles with papillary muscles	Fig. 3.8
SRS0006-24	Heart: atria, aortic root, & arteries	Fig. 3.9
BG-3	Superior colliculus	Fig. 3.10
CA-2	Carotid artery	Fig. 3.11

Table 3.3: Performance Evaluation: Improvement in Image Contrast.

As previously mentioned, improvement in image contrast is measured in terms of two metrics, namely,  $CIR$  and  $CIR^{adj}$ , over selected regions of

---

<sup>3,12</sup>Magnetic resonance angiography (MRA) is a branch of MRI technology that provides detailed images of blood vessels. It provides an important diagnostic tool for detection and diagnosis of heart disorder, stroke and blood vessel diseases.

interest. Table 3.3 lists the regions of interests and the output reference figures of the test cases. Evaluation over other regions provides similar and consistent results. Each of the reference figures consists of four parts. The configuration of sub-figures on the left shows the region of interest in the data view in part (A) and the foreground and background partition created manually over the window of selection. The foreground and background of the image are displayed in part (B) and part (C) respectively in each figure. To allow easy interpretation, the foreground of the image in part (B) is superimposed over a completely “blacked-out” background. Similarly, the foreground is set to white in part (C) to reveal the background of the image.

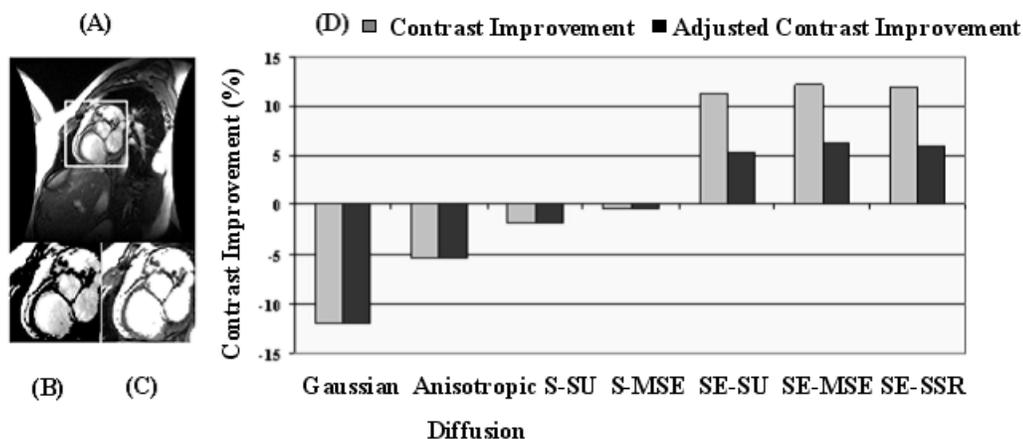


Figure 3.6: Cardiac Image: SRS0000-10. (A) region of interest; (B) foreground partition; (C) background partition; and (D) contrast improvement comparison.

Part (D) on the right shows the results of the performance evaluation with respect to the improvement of image contrast. Both  $CIR$  (the gray bars) and  $CIR^{adj}$  (the black bars) are reported. Scanning from right to left, one can see the improvement in the performance of the algorithms: the three

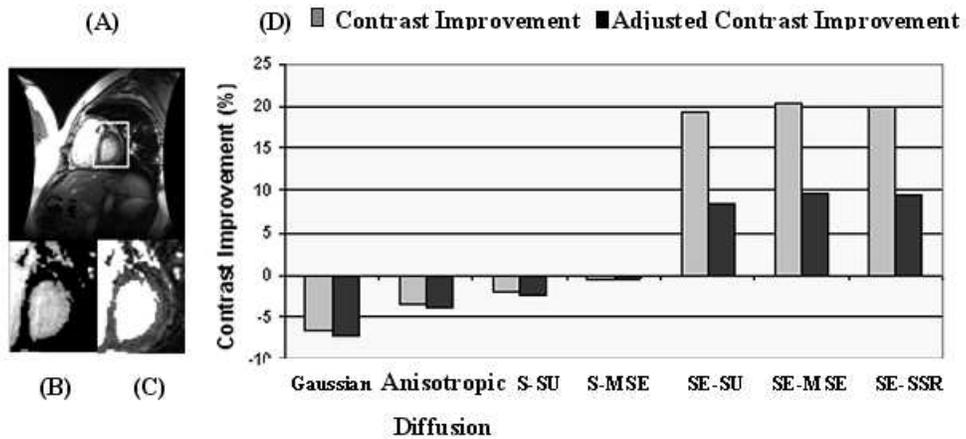


Figure 3.7: Cardiac Image: SRS0000-70. (A) region of interest; (B) foreground partition; (C) background partition; and (D) contrast improvement comparison.

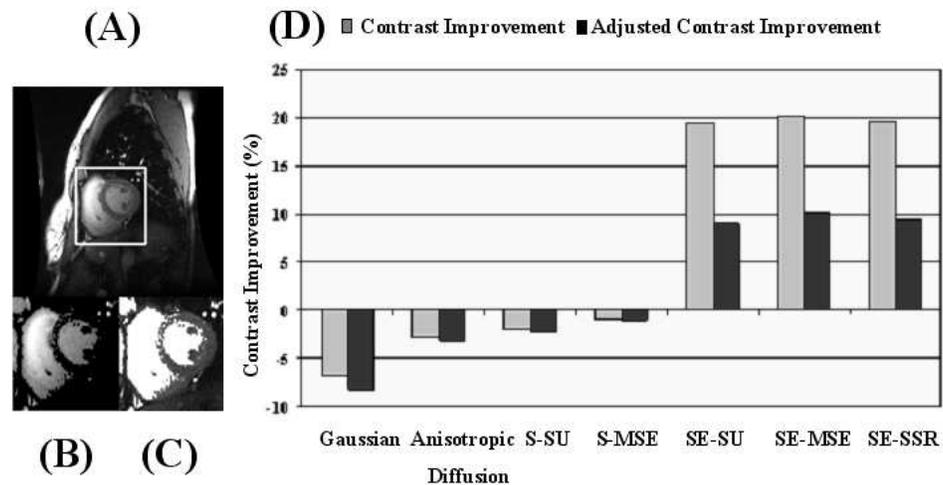


Figure 3.8: Cardiac Image: Eman-83. (A) region of interest; (B) foreground partition; (C) background partition; and (D) contrast improvement comparison.

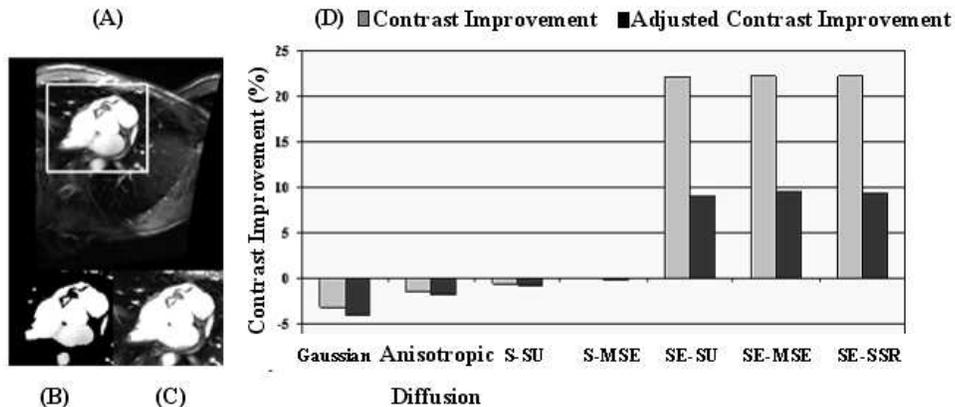


Figure 3.9: Cardiac Image: SRS0006-24. (A) region of interest; (B) foreground background partition; (C) background partition and (D) contrast improvement comparison.

algorithms on the right, namely, the SE-SSR, the SE-MSE, and the SE-SU algorithm, are the SE-group, the next two algorithms are the S-group, and the two on the left are spatial-domain diffusion-based algorithms, namely Gaussian and anisotropic diffusion filtering. Qualitatively,  $CIR$  and  $CIR^{adj}$  agree on the comparative performance of the algorithms. Algorithms in the SE-group are the only ones capable of generating images with better image contrast over the corresponding input image. The other groups remove noise at the expense of image contrast. In general, wavelet-based algorithms show better performance than the spatial domain algorithms in spite of the fact that anisotropic diffusion filtering adapts to image structure to preserve essential signal discontinuities.

Differences in performance are also evident from the qualitative evaluation of the output image<sup>3.13</sup> of the algorithms. The reference figures are

<sup>3.13</sup>The input and output images are shown in Part II of the appendix.

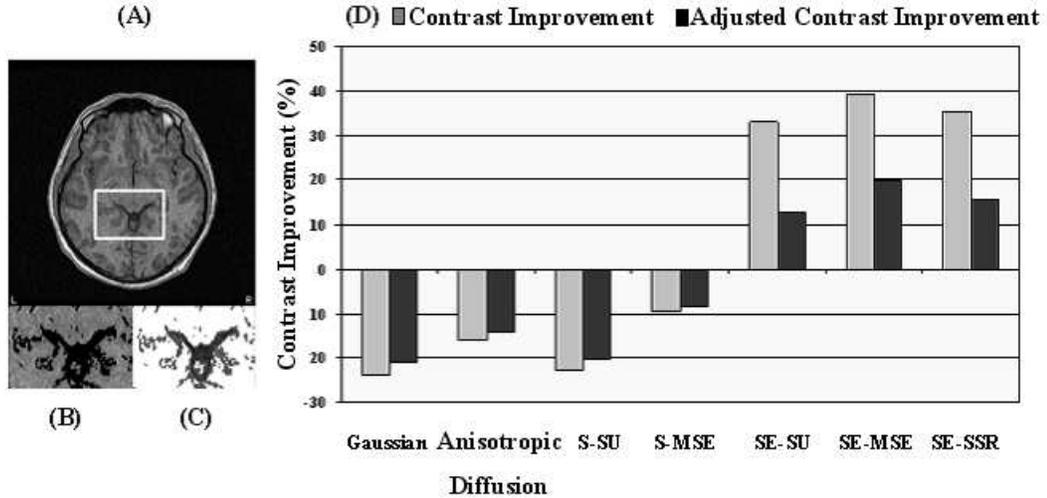


Figure 3.10: Brain Image: BG-3. (A) region of interest; (B) foreground partition; (C) background partition; and (D) contrast improvement comparison.

listed in Table 3.2. The SE-group by virtue of its enhancement scheme is capable of removing much of the noise and low-intensity features of the images. Fig. 3.12 provides an example. To highlight the cross group difference in performance, the figure shows the contrast adjusted view and the edge map of the output of representative algorithms from each group.

Consider the bottom row (C) of images: both views attest to the efficacy of the SE-group in removing most of the low-intensity features around the heart and outside the body. In contrast, a substantial portion of these features survive the smoothing processes in the other two groups. This observation represents a general difference in the capability of enhancement of the target organ across groups, a difference that is found in all test cases in the experiments (See Part II of the appendix.)

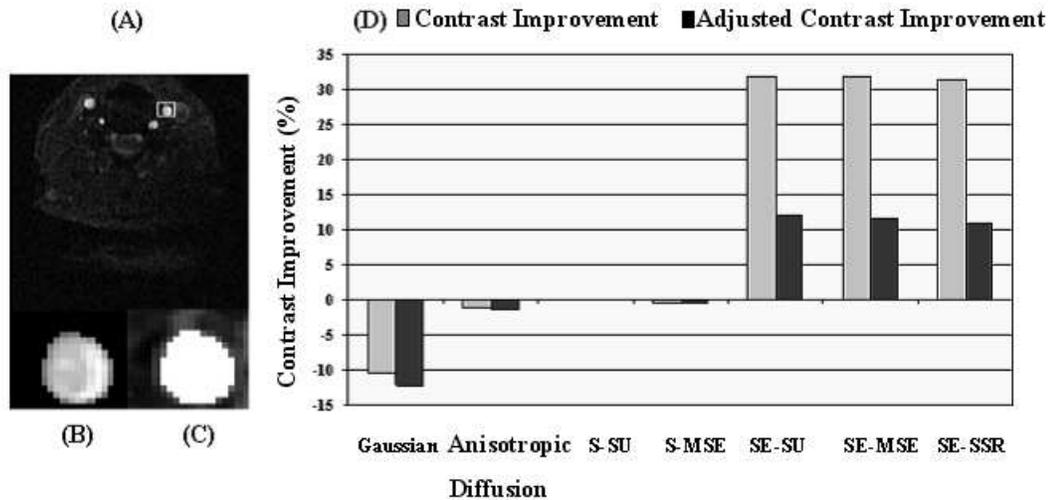


Figure 3.11: Cardiovascular Image: Carotid Arteries 2. (A) region of interest; (B) foreground partition; (C) background partition; and (D) contrast improvement comparison.

Non-cardiac images are selected to demonstrate the applicability of the enhancement framework. CA-2 is a cross-section of a human neck. Carotid arteries constitute the target organs of the cardiovascular MR image. There are four carotid arteries (Fig. 3.13), one pair on each side of the neck, that deliver oxygen carrying blood from the heart to the head and the brain. These carotid arteries are represented by the high-intensity circular objects in a very noisy background. It can be easily observed from the contrast adjusted images in Fig. H.19 - Fig. H.21 in Appendix II that the arteries are embedded and concealed in the background features of the neck. All but the SE-group fail to enhance the image in any significant way. For a quick reference, Fig. 3.14 shows the input and output images, one from each enhancement scheme group. The arteries are much enhanced by the SE-SSR scheme (the output of which is shown in Part (D) of the figure) which removes most of the background structures (also compare the edge maps in

---

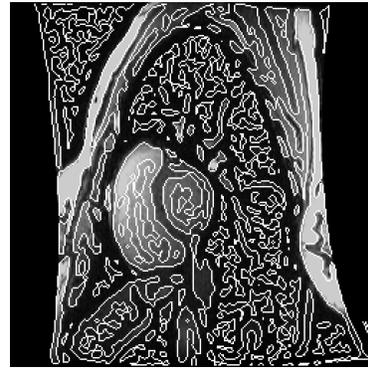
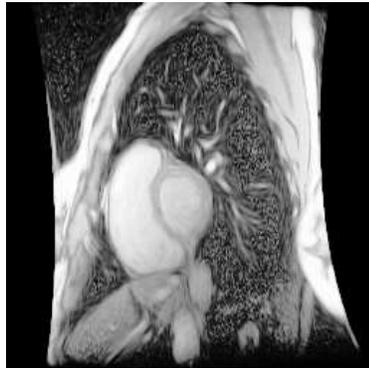
---

Contrast Adjusted Views

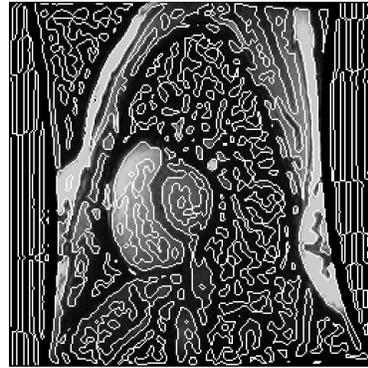
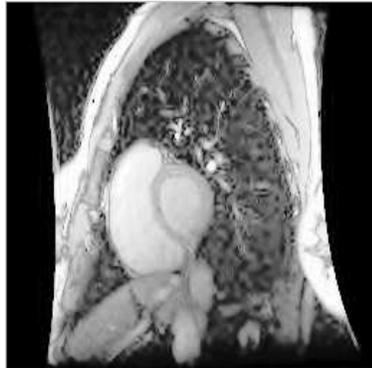
Edge Map

---

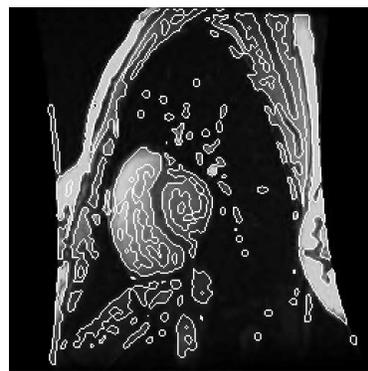
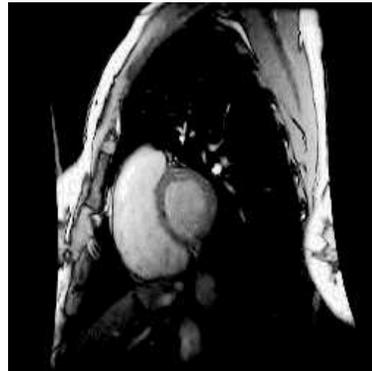
---



(A) Anisotropic Diffusion Filtering



(B) S-MSE Procedure



(C) SE-MSE Procedure

---

---

Figure 3.12: Cardiac Image: Eman-70. Comparative Performance.

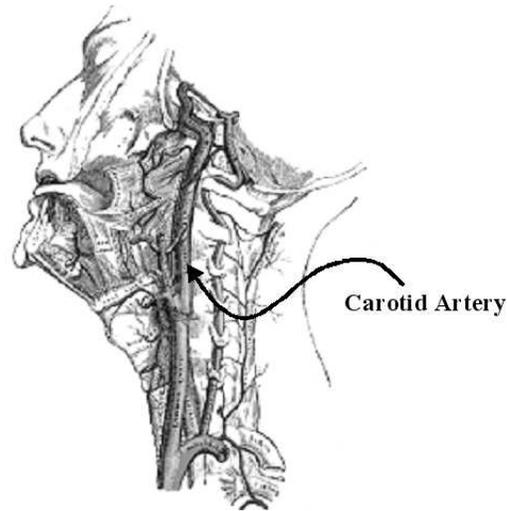


Figure 3.13: Carotid Arteries

Appendix II.)

One major motivation behind the computational framework of adaptive enhancement is the preservation of target structures – in the context of clinical diagnosis, that means anatomical structures of target organs. Techniques based on wavelet analysis and those based on (anisotropic) diffusion represent two of the major approaches that aim at signal recovery with minimum loss of relevant feature information. Both the SE-MSE and SE-SSR scheme provide the best approximation to the noise-free synthetic images in terms of error measures with respect to the  $l_1$  and the  $l_2$  norms. The two examples shown in Fig. 3.15 and Fig. 3.16 illustrate this quantitative difference in the quality of the enhanced images generated by the test algorithms. The ‘R’ and ‘L’ at the bottom of BG-3 are marked by the scanner to indicate the orientation of the image. In a sense, each of them is comprised of salient and coherent structures of high signal intensity against a signal-free but noisy background. It is natural to expect that the enhancement scheme would

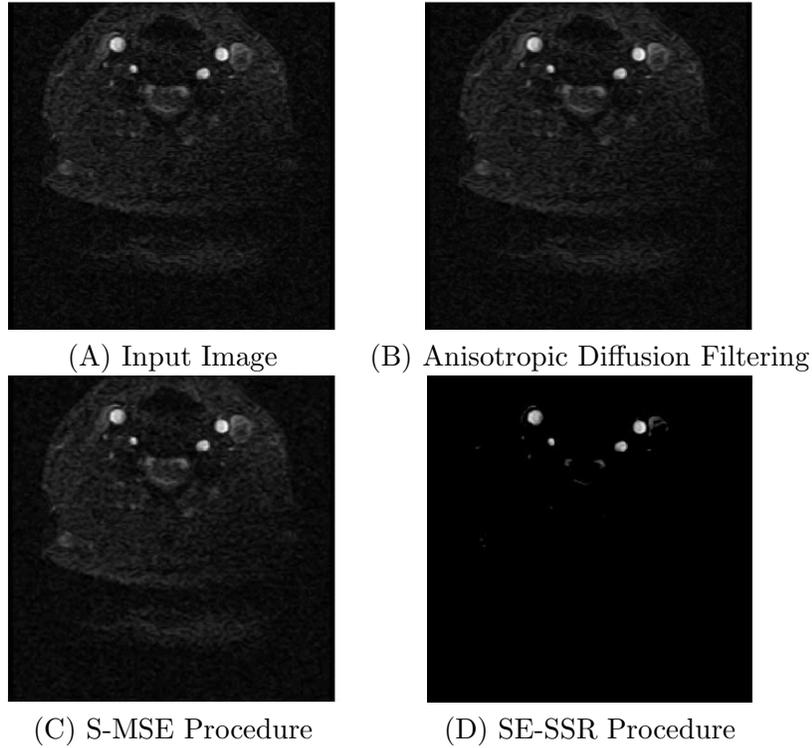
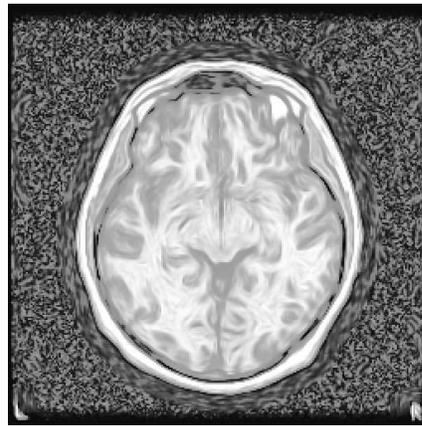


Figure 3.14: Cardiovascular Image: Carotid Arteries 2. Given data from input and output images.

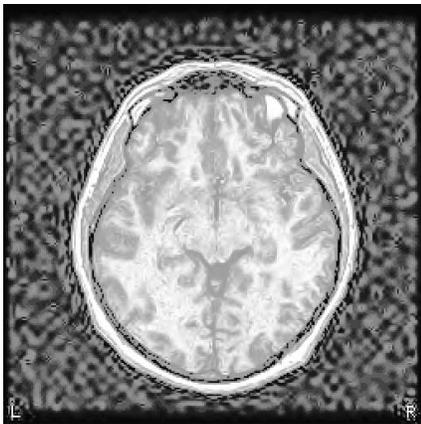
preserve these structures. In this case, as shown in Fig. 3.15, wavelet-based schemes preserve these characters in an obviously much better way than their diffusion-based counterparts. While the markers stay sharp in the output image of schemes in the S-group and the SE-group (as shown in Part (C) and (D) of the figure), the diffusion-based operations blur the marker significantly as shown in Part (B) of the figure. Comparing the different views of the outputs presented in Appendix II, it is easy to see that the SE-group performs much better than the other group in removing noise from the image, enhancing the image contrast and preserving coherent structures



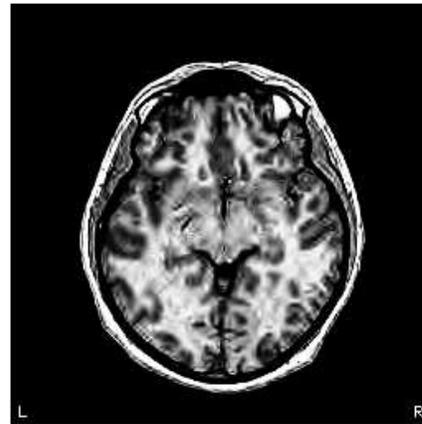
(A) Input Image



(B) Anisotropic Diffusion Filtering



S-MSE Procedure



(D) SE-SSR Procedure

---

---

Figure 3.15: Brain Image: BG-3. Input and output images (contrast adjusted).

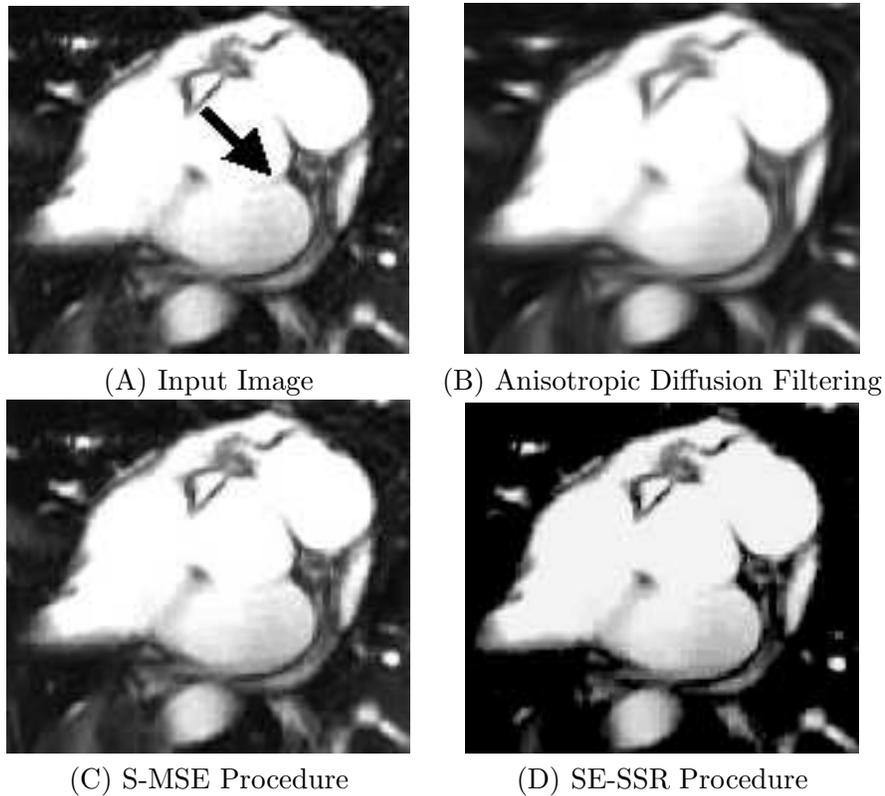


Figure 3.16: Cardiac Image: SRS0006-24. Given data of input and output images. The arrow in (A) points to a valve of the heart.

of the image.

Fig. 3.16 presents the target region of interest in the input and outputs image of SRS0006-24. Image contrast of the target structures of the heart is much improved by the SE-SSR scheme (part (D)) over the input image (part (A)) as well as the output of the scheme in other groups. The image captures part of a heart valve (indicated by the arrow in part (A) of the figure) in motion. The structure is preserved under the S-MSE scheme (Part (C)) and the SE-SSR scheme (Part (D)). The latter indeed achieves some slight

enhancement of the structure over the input data. In contrast, the valve is almost entirely smoothed out by the anisotropic diffusion filter (Part (B)). These examples provide telling evidence as to how well different schemes retain important information of coherent structures.

### **3.3 Summary**

This chapter presents the major findings of comparative evaluation of the enhancement schemes. A set of algorithms in three groups is applied to gain some insight into the performance of the framework proposed in this research. All the experimental findings agree with the hypothesis of this report that the SE-framework of adaptive enhancement achieves the combined effect on the input image: that is, it reduces both image noise and other confounding features outside the organs of interest, preserves the integrity of target structures and enhances their image contrast. For the algorithms in other test groups, this task set imposes conflicting objectives, one can only be achieved at the expense of the others. The non-cardiac test cases demonstrate that the scope of applicability of the framework is not restricted to this special class of images that motivated it.

## Chapter 4

# Conclusion

### 4.1 Discussion

This report presents a computational framework of adaptive enhancement of magnitude MR images for clinical diagnosis. It seeks an algorithmic solution which is capable of reducing image noise and low-intensity structures that obscure relevant information about the anatomical structures of the organ and to enhance image contrast of the target structures. Empirical evaluation of performance produces strong support for the theoretical hypothesis concerning the effect of the scheme on input data. In both quantitative and qualitative terms, the adaptive enhancement schemes implemented in the SE-group are capable of producing a significantly enhanced image in comparison to the output of other schemes. Image noise and low-intensity structures in the proximity of target organs are significantly attenuated and image contrast is enhanced. Turning off the enhancement step fails to produce this effect; thereby strongly supporting the major hypothesis of this report that the improvement of image quality is unquestionably a unique contribution of the adaptive strategy of the framework.

Underlying the enhanced image quality is the synergy of two kinds of adaptivity built into the enhancement scheme. The signal dependence of image noise and its interaction with low intensity structures necessitate the adaptivity of the first kind, namely, the adaptivity to signal intensity, which is a focus of our preceding discussion. Preservation and enhancement of the integrity of coherent structures relies upon the adaptivity of another kind – that is, adaptivity to coherent structures.

The wavelet-domain enhancement schemes are capable of maintaining the coherence of the structures of an image – even in the case where these structures are vaguely visible to the human eye such as a partial heart valve caught in a snapshot of a beating heart. The preservation of coherent structures starts from the first step of the enhancement scheme. An efficient representation of image structures in terms of the scale-based hierarchy of orthonormal decomposition of signal allows better discrimination between random and coherent signal fluctuations. As a consequence, wavelet shrinkage retains the integrity of these structures while neutralizing the effect of spurious signal fluctuations. The adaptive enhancement is realized in the second step through the mechanics of directing the force of signal attenuation toward the low-intensity regions and emphasizing the image gradient along the structural boundaries. Wavelet analysis provides a necessary decomposition of image data to achieve both purposes simultaneously.

A question naturally arises in this context: what is the optimal choice of the basis, i.e., the set of signal expansion functions, for adaptive enhance-

ment. Empirically the scheme of adaptive enhancement is almost basis-independent. The output reported in Chapter 3, being itself generated from different wavelet bases, represents a subset of experiments where different wavelet bases are tested on each case. The results of performance evaluation do not significantly vary with the choice of wavelet bases. It is, however, worth commenting on the Gibbs phenomenon. Like its Fourier counterpart, the wavelet expansion of a signal tends to overshoot near the discontinuities – an observation termed the Gibbs phenomenon. The phenomenon causes reasonable concern for adaptive enhancement, because signal discontinuities are crucial for image enhancement and analysis. It has been shown that wavelet expansion with different orthogonal basis exhibits this phenomenon – the Haar wavelet basis is a rare exception; see Walter and Shen [113]. This property of the Haar wavelet basis, along with its being the only truly symmetric basis for orthonormal expansion, may serve as an explanation for empirical observations reported in many previous studies that the Haar wavelet basis produces a better denoised image than other orthonormal bases do. Evidence also suggests that this set of base functions also allows a better recovery of high-intensity signal.

The framework of adaptive enhancement arises from the context of spatial-temporal analysis of cardiac image sequences. As each input data sequence may involve hundreds of images, computational complexity of processing and analysis is an important concern in algorithmic design in each step of the way of extraction of interesting information from the data set. The computational load of the adaptive enhancement scheme presented in this report is shared by four major components. These are (1) wavelet analysis and synthesis, (2) transformation of representations, (3) enhancement based

on the approximation model in the smoothing step, and (4) the adaptive enhancement scheme in the second step. All but the smoothing step are linear in the input data, i.e., the number of pixels in the image. The complexity of the smoothing step varies with the choice of Gaussian-based enhancement scheme. Wavelet shrinkage techniques such as the SU, MSE and SSR scheme used in the performance evaluation and comparison in this research are linear in the data, and so are the implementations of the schemes as a whole. Thus, in terms of computational load, the schemes in the SE-group have an advantage over the techniques based on an iterative procedure of intensity diffusion, such as anisotropic diffusion filtering. In general, the complexity of the scheme is determined by the smoothing scheme. The choice of the techniques used in this step should therefore balance between the computational requirement and the quality of enhancement it can attain.

## 4.2 Future Work

The adaptive scheme achieves image enhancement by simultaneous attenuation of confounding features and improvement of image contrast. The complexity of the data set is greatly reduced in a way that signals of coherent structures, in particular those corresponding to the anatomical parts of target organs, become more visible – in terms of detectability and recognizability – for the subsequent algorithmic operations. This research provides both quantitative and qualitative measurements of these effects. How this newly gained data quality is translated into better performance of these algorithmic operations – particularly segmentation and localization of the target organs – is an important issue for further exploration.

The focus of this research is that of adaptive enhancement and the feasibility of a solution to the problem. It raises a number of interesting questions to be explored. First, the smoothing technique plays an important role in the scheme of adaptive enhancement. The comparative performance evaluation restricts its choice of techniques used in the smoothing step to a class of shrinkage techniques that is not adaptive to the geometrical structures of image signal. An interesting issue for future exploration is the role of (geometrical) structure adaptivity of smoothing techniques in adaptive enhancement. The gain in performance with the structure adaptive smoothing step must be analyzed in conjunction with its price in extra computation load, since the existing techniques of this class are mostly non-linear in their complexity. Second, the performance of the SE framework is shown to be almost basis independent with respect to the standard orthonormal wavelet bases. The general framework of wavelet analysis still affords other choices for the purpose of adaptive enhancement, such as the translation invariant wavelet basis, non-decimal wavelet decomposition, wavelet packets and the structure adaptive bases such as curvelet and contourlet.

The tuning parameter  $\mathcal{T}$  of the enhancement step should be set in relation to the critical range of intensity which discriminates the boundaries of the target object with the background structures in its low-intensity proximity. As pointed out in Section 2.2.2, this critical range does not change dramatically across images both spatially and temporally. Few thresholds are required for a sequence of cardiac images. This report leaves this parameter to be experimentally determined. Estimation of this factor from image data may be an interesting issue for further research. It is advantageous to release human experts from the task of threshold selection. However, this

advantage must be weighted against the requirement for optimality of parameter selection. The critical range that separates the structure signal from its surround may be too small to be correctly gauged without some idea of the boundary. Nevertheless, semi-automated estimation of this parameter may still be desirable if it can suggest a neighborhood for the starting point for of the optimal estimate for human experts.

# Appendices

## Part I

# Technical Notes

# Appendix A

## Mechanics of Wavelet Analysis

Wavelet analysis, a branch of applied mathematics, gives rise to a repertoire of tools for signal/function analysis with application to almost every discipline of computational science. The formal aspects of the system are however beyond the scope of this report. For the clarity and convenience of exposition, this appendix<sup>A.1</sup> intends to provide a brief review of some essential concepts of the framework with the sole purpose to fix notation and to allow a minimum understanding of the mechanics of the techniques used in this report.

A wavelet system decomposes a signal into a hierarchically organized representations in terms of a set of basis functions called wavelets, each of which is generated from a prototype waveform, called the mother wavelet,  $\psi$ , under dilation and translation. A real function admissible to be a wavelet

---

<sup>A.1</sup>The discussion is based on the following reference, which will not be cited separately, including [28; 97; 8; 31; 25; 75; 76; 77; 130; 50; 109; 105; 107; 15; 104; 74; 24; 3; 32; 95].

must have finite energy and a zero mean. A “dyadic system” generates the basis functions with a power-of-two logarithmic scaling of dilation and translation of the mother-wavelet, leading to a construction of an orthonormal wavelet basis,  $\{\psi_{j,k}\}$ ,

$$\psi_{j,k}(\tau) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{\tau - 2^j k}{2^j}\right) \quad \text{where } j, k \in \mathbb{Z} \quad (\text{A.0.1})$$

Wavelet analysis maps a signal  $\mathbb{S}$  to a set of wavelet coefficients given by

$$\omega_{j,k} = \langle \mathbb{S}, \psi_{j,k} \rangle = \int_{-\infty}^{+\infty} \mathbb{S}(\tau) \psi_{j,k}(\tau) d\tau. \quad (\text{A.0.2})$$

Each member of the coefficient set denoted by  $\{\omega_{j,k}\}$  provides a measurement of the local variations of the signal  $\mathbb{S}$  in a neighborhood of size proportional to the scale indexed by  $j$  around the location indexed by  $k$ .

Given an orthonormal basis  $\{\psi_{j,k}\} = \{\psi_{j,k} | j, k \in \mathbb{Z}\}$ , wavelet synthesis provides the inverse mapping which allows the reconstruction of the signal  $\mathbb{S}$  in terms of the wavelet coefficients:

$$\mathbb{S}(\tau) = \sum_j \sum_k \omega_{j,k} \psi_{j,k}(\tau) \quad (\text{A.0.3})$$

This inverse mapping “pieces together” all the structural details of the image under the double summation operators across scales and locations. A truncated series  $\sum_{j>j_0} \sum_k \omega_{j,k} \psi_{j,k}$  yields a coarse-scale approximation, denoted by  $\mathbb{S}_{j_0}$ , by removing all the details recoverable at finer scales less than  $j_0$ . A hierarchy of approximations  $\{\mathbb{S}_j\}$  at successively coarser scales can be represented in terms of a set of basis functions, called scaling functions,  $\{\phi_{j,k}\}$ , each of which is generated from a father scaling function  $\phi$  under dilation

and translation  $\phi_{j,k}(\tau) = 2^{-\frac{j}{2}} \phi\left(\frac{\tau-2^j k}{2^j}\right)$ . The coarse-scale representation  $\mathbb{S}_j$  of the signal at a scale  $j$  is defined as [3]

$$\mathbb{S}_j(\tau) = \sum_k \mathcal{S}_{j,k} \phi_{j,k}(\tau) \quad (\text{A.0.4})$$

where the approximation coefficients are given by

$$\mathcal{S}_{j,k} = \langle \mathbb{S}, \phi_{j,k} \rangle = \int_{-\infty}^{+\infty} \mathbb{S}(\tau) \phi_{j,k}(\tau) d\tau. \quad (\text{A.0.5})$$

Signal synthesis given by Eq. (A.0.3) can be rewritten as

$$\mathbb{S} = \underbrace{\sum_k \mathcal{S}_{j_0,k} \phi_{j_0,k}}_{\text{Approximation}} + \underbrace{\sum_{j < j_0} \underbrace{\sum_k \omega_{j,k} \psi_{j,k}}_{\text{Details at scale } j}}_{\text{Details at scales } j < j_0}. \quad (\text{A.0.6})$$

The first term is the coarse scale approximation  $\mathbb{S}_{j_0}$  at scale  $j_0$  and the last term represents the details at scales  $j < j_0$ . A sequence of approximations  $\{\mathbb{S}_j\}$  can be generated recursively by adding to the approximation of level  $j$  the details captured at the corresponding level, i.e.,

$$\mathbb{S}_{j-1} = \mathbb{S}_j + \sum_k \omega_{j,k} \psi_{j,k}. \quad (\text{A.0.7})$$

The coefficients  $\mathcal{S}_{j_0,k}$  at scale  $j_0$  in Eq. (A.0.6) can thus be interpreted as an aggregate measurement of the variations captured at the coarsest scales  $j \geq j_0$ . On the other hand, any signal structure measurable at a scale  $j < j_0$  is recoverable from the wavelet coefficients at the corresponding scale. The second term of the left member of Eq. (A.0.6) is a compact expression of the operation which collects all these details in the reconstruction.

Orthonormal wavelet analysis partitions signal energy in multi-scale representations. Signal energy in a scale  $j$ , denoted by  $\mathbb{E}_j$ , is contained in the coefficients at that scale, i.e.,

$$\mathbb{E}_j = \sum_k |\omega_{j,k}|^2. \quad (\text{A.0.8})$$

The total energy of a signal, given by<sup>A.2</sup>  $\mathbb{E} = \|\mathbb{S}\|^2$ , is preserved in the decomposition, that is,

$$\mathbb{E} = \sum_j \mathbb{E}_j \quad (\text{A.0.9})$$

$$= \sum_k |\mathcal{S}_{j_0,k}|^2 + \sum_{j < j_0} \sum_k |\omega_{j,k}|^2. \quad (\text{A.0.10})$$

Consider an orthonormal wavelet basis with compact support. Each wavelet coefficient depends only on the local structures of the signal, capturing signal activities over the local neighborhood within the compact support of the corresponding basis function. Suppose that over the neighborhood of  $\tau_0$ , the signal may be approximated by the polynomial expansion:

$$\mathbb{S} = P + Re \quad (\text{A.0.11})$$

where  $P$  is the polynomial approximation of  $\mathbb{S}$ , and  $Re$  the remainder. Given a wavelet with  $n$  vanishing moments, i.e.,  $\int_{-\infty}^{\infty} \tau^n \psi(\tau) d\tau = 0$ , the associated wavelet coefficients capture only the residual term of the polynomial

---

<sup>A.2</sup>Given  $\mathbf{x}$ ,  $\|\mathbf{x}\|$  denotes the norm of  $\mathbf{x}$ .

approximation of degree  $n - 1$ . Let  $\mathcal{P}_v f$  be the projection of the function  $f$  to the discrete approximation spaces  $\mathcal{V}$  spanned by the scaling functions, and  $\mathcal{R}_w f$  be the projection to the detail signal space  $\mathcal{W}$  spanned by the wavelets. Then Eq. (A.0.6) can be interpreted as

$$\mathbb{S} = \mathcal{P}_v \mathbb{S} + \mathcal{R}_w \mathbb{S} \quad (\text{A.0.12})$$

The coarse-scale approximation is the projection of the polynomial approximation of the signal to the approximation space, and the details representation is the aggregate of the projections of those fine-scale oscillations that cannot be represented by the polynomial approximation. Projected into this detail space spanned by the wavelets are rapidly oscillating structures and signal singularities.

A separable wavelet orthonormal basis for the space of two dimensional signal with finite energy can be constructed from tensor products of a scaling function  $\phi$  associated with one-dimensional discrete spaces of signal decomposition and the corresponding wavelet function  $\psi$ . Generated by a two-dimensional separable scaling function<sup>A.3</sup>,  $\phi^{2D}(\tau)$ , is a basis of discrete approximation space  $\mathcal{V}_j^{2D} = \mathcal{V}_j \otimes \mathcal{V}_j$ . The corresponding wavelet functions<sup>A.4</sup>,  $\{\psi_{j,\mu}^{\mathbf{H}}, \psi_{j,\mu}^{\mathbf{V}}, \psi_{j,\mu}^{\mathbf{D}}\}$ , where  $\mu$  is the equivalence of  $k$  in  $\mathbb{Z}^2$ , constitute an or-

---

<sup>A.3</sup>The separable basis functions associated with a 1D scaling function  $\phi$  is defined as the followings. Let  $\tau = [\tau_1, \tau_2]$  and  $\mu = [\mu_1, \mu_2]$  be vectors of a 2D spatial location. The scaling function is given by

$$\phi^{2D}(\tau) = \phi^{2D}(\tau_1, \tau_2) = \phi(\tau_1)\phi(\tau_2), \quad (\text{A.0.13})$$

and the corresponding set of basis functions that span the approximation space is given as the following:

$$\phi_{j,\mu}^{2D} = 2^{-j} \phi^{2D}(2^{-j}(\tau - 2^j \mu)) \quad (\text{A.0.14})$$

thonormal basis of the  $2D$  discrete detail signal space. Let  $\Psi = \{\mathbf{H}, \mathbf{V}, \mathbf{D}\}$  be the index set for the orientations – horizontal, vertical, and diagonal respectively – of the basis functions that span the detail spaces. Every  $2D$  signal with finite energy,  $\mathbb{S}$ , may be represented in terms of the separable basis

$$\mathbb{S} = \sum_{\mu} \mathcal{S}_{j_0, \mu} \phi_{j_0, \mu}^{2D} + \sum_{j < j_0} \sum_{\mu} \sum_{\kappa \in \Psi} \omega_{j, \mu}^{\kappa} \psi_{j, \mu}^{\kappa} \quad (\text{A.0.19})$$

where

$$\mathcal{S}_{p, q} = \left\langle \mathbb{S}, \phi_{p, q}^{2D} \right\rangle = \int \mathbb{S}(\tau) \phi_{p, q}(\tau) d\tau; \quad \text{and} \quad (\text{A.0.20})$$

$$\omega_{p, q}^r = \left\langle \mathbb{S}, \psi_{p, q}^r \right\rangle = \int \mathbb{S}(\tau) \psi_{p, q}^r(\tau) d\tau. \quad (\text{A.0.21})$$

The separable basis  $\{\psi_{j, \mu}^{\kappa}\}$  captures the signal structures of an image at different scales, locations, and orientations. The details of an image at each scale is therefore represented in three sub-representations or sub-images, each generated by a wavelet of a specific orientation. The energy of the horizontal wavelet,  $\psi^{\mathbf{H}}$ , is largely localized at low horizontal frequency and high vertical frequency; that of the vertical wavelet,  $\psi^{\mathbf{V}}$ , at high horizontal frequencies and low vertical frequencies; that of the diagonal wavelet,  $\psi^{\mathbf{D}}$ , at high frequencies in both directions. These sub-images capture different

---

<sup>A.4</sup>The separable basis function associated with a  $1D$  scaling function  $\phi$  and its associated wavelet function  $\psi$  is defined as the following:

$$\psi^{\mathbf{H}}(\tau) = \phi(\mu_1)\psi(\mu_2); \quad (\text{A.0.15}) \quad \psi^{\mathbf{V}}(\tau) = \psi(\mu_1)\phi(\mu_2); \quad (\text{A.0.17})$$

$$\psi^{\mathbf{V}}(\tau) = \psi(\mu_1)\phi(\mu_2); \quad (\text{A.0.16})$$

Let  $\kappa \in \Psi = \{\mathbf{H}, \mathbf{V}, \mathbf{D}\}$ .

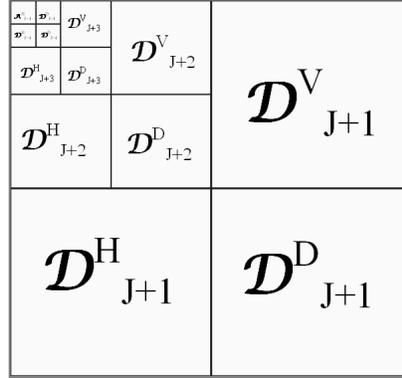
Let  $\tau = [\tau_1, \tau_2]$  and  $\mu = [\mu_1, \mu_2]$  be vectors of a  $2D$  spatial location. The set of basis functions that span the detail signal space is given by

$$\psi_{j, \mu}^{\kappa} = 2^{-j} \psi^{\kappa}(2^{-j}(\tau - 2^j \mu)). \quad (\text{A.0.18})$$

types of signal signatures in terms of variations along different directions. Figure A.1 shows a wavelet pyramid of a medical image. The diagrammatic figure shows how the sub-images are arranged. The one at the upper-left corner is the approximation generated by the  $\phi^{2D}$  at a selected coarsest scale, where the others are the image decompositions in various spaces spanned by the separable wavelets at different scales and orientations.

Multiresolution analysis results in a set of hierarchically related representations given by orthogonal projections of an image onto the approximation and detail spaces. A number of useful properties of wavelet representations are relevant for image analysis. More important ones include the followings:

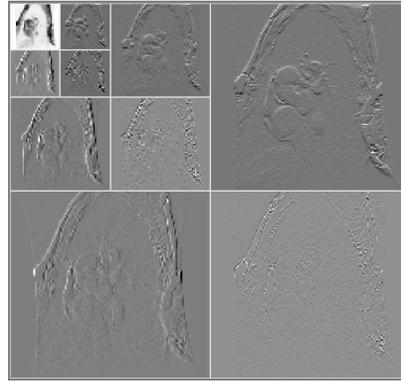
<b>Locality</b>	Each wavelet coefficient represents the localized information about the structure of the signal in both scale and location.
<b>Multi-resolution</b>	Wavelet coefficients represent information about the structures of the signal in a nested set of scales.
<b>Statistical independence</b>	Wavelet coefficients of a real-world image tend to be uncorrelated. For obvious reason, this property is also called decorrelation.
<b>Energy compaction</b>	The energy of the wavelet representation of the signal is concentrated in a sparse number of scale-location coordinates where signal singularities occur within the support of the wavelet basis.
<b>Exponential Decay</b>	The magnitude of wavelet coefficients tends to decay exponentially across scales.



(A)



(B)



(C)

---

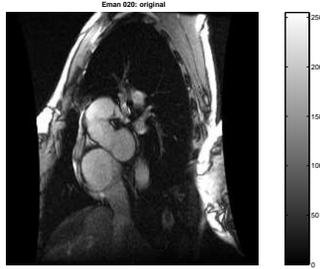


---

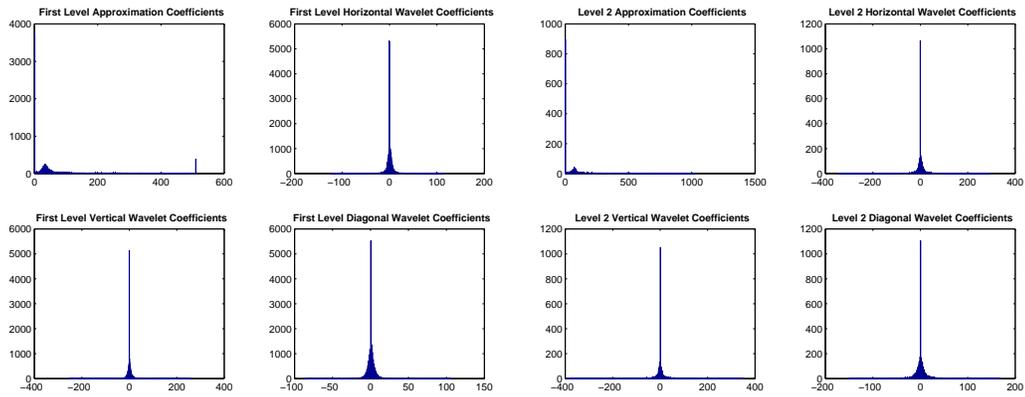
$\mathcal{D}_J^\kappa$ : Detail image generated by wavelets  $\{\psi_{J,\mu}^\kappa\}$  at scale  $2^J$ , and  
 $\mathcal{A}_{J_0}$ : Approximation image generated by scaling functions  $\{\phi_{J_0,\mu}^{2D}\}$  at scale  $2^{J_0}$ , where  $\kappa \in \{\mathbf{H}, \mathbf{V}, \mathbf{D}\}$

Figure A.1: Multiresolution Decomposition using Haar Transform. (A) illustrates the pyramidal organization of sub-images generated by a 2D wavelet decomposition. The sub-image at the upper-left corner is the approximation  $\mathcal{A}_{J+5}$ . (B) is the spatial domain representation of an image. (C) is the wavelet decomposition of the image.

Proven to be pivotal for our purpose is that signal representations in the wavelet domain are characterized by a sparsity of non-negligible wavelet coefficients which are largely uncorrelated. As shown in figure A.2, the spiked peak for zero-valued coefficients is conspicuous from the histogram of the empirical coefficients of a selected MRI at the four finest scales using the Haar Transform. This sparsity of non-negligible coefficients reduces the computational load of analysis. More importantly, small groups of non-negligible coefficients allow a distinct signature for the salient features of the scene.

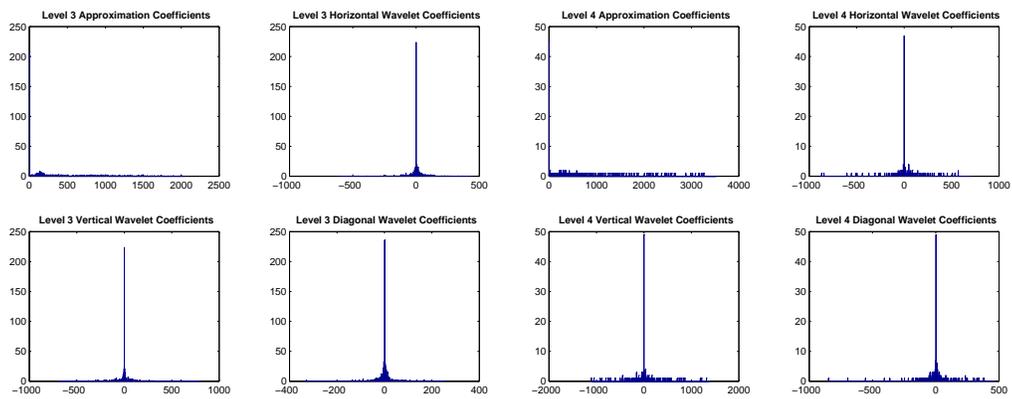


(a) Original



(b) Finest

(c) Fine



(d) Less Fine

(e) Coarse

Figure A.2: Histogram of Wavelet Coefficients using Haar Transform. There are four histograms of the coefficients at each level of resolution — the top-left histogram describes the distribution of the approximation coefficients and the other three histograms describe that of the wavelet coefficients associated with the basis functions of different orientation. Signal analysis in the wavelet domain is characterized by the histograms of wavelet coefficients that are closely clustering around spikes at zero at the fine scales.

## Appendix B

# Statistical Properties of Rician Noise

This appendix provides some further details on the SNR-dependent distribution of Rician noise. Let  $I$  be magnitude image data,  $S$  noise free signal, and  $\sigma^2$  the noise parameter, which is the variance of noise in the complex images. Given the (noise-free) signal and the noise parameter, observed data are distributed Rician with the condition density function given in Eq. (2.2.1):

$$f(I|S, \sigma) = \frac{I}{\sigma^2} \exp \left\{ -\frac{I^2 + S^2}{2\sigma^2} \right\} I_o \left( \frac{IS}{\sigma^2} \right) I_{\{x \geq 0\}}(I)$$

where  $I_o$  is the modified Bessel function of the first kind of zeroth order, and  $I_{\mathcal{A}}$  the indicator function over the set  $\mathcal{A}$ .

## B.1 Rician Noise in Signal Free Regions

First consider  $I_0$ , the modified Bessel function of the first kind of zeroth order is given by [2; 9]:

$$I_0(x) = \frac{1}{\pi} \int_0^{\pi} e^{\pm x \cos\theta} d\theta = \frac{1}{\pi} \int_0^{\pi} \cosh(x \cos\theta) d\theta, \quad (\text{B.1.1})$$

where  $\cosh(\eta) = \frac{1}{2}(e^{\eta} + e^{-\eta})$ . Observe that  $\cosh(0) = 1$ . Thus,

$$I_0(0) = \frac{1}{\pi} \int_0^{\pi} 1 d\theta = \frac{1}{\pi} \left[ \theta \right]_0^{\pi} = \frac{1}{\pi} \pi = 1 \quad (\text{B.1.2})$$

Thus in the absence of signal,  $S = 0 \implies \frac{IS}{\sigma^2} = 0$ . Rician Distribution given in Eq. (2.2.1) can be rewritten as

$$\mathcal{F} = \frac{I}{\sigma^2} \exp \left\{ -\frac{I^2}{2\sigma^2} \right\} 1_{I>0}(I),$$

that is Eq. (2.2.2), which is a Rayleigh distribution[41; 66].

## B.2 Rician Noise in Infinity Signal Regions

Observe that the asymptotic behavior<sup>B.1</sup> of the modified Bessel function for large  $x$

$$I_n(x) \sim \frac{e^x}{\sqrt{2\pi x}} \left[ 1 - \frac{\mu - 1}{8x} \right] \quad (\text{B.2.1})$$

where  $\mu = 4n^2$ ; see [60]. Thus,

$$I_0(x) \sim \frac{e^x}{\sqrt{2\pi x}} \left[ 1 + R_0(x) \right]. \quad (\text{B.2.3})$$

where  $R_0(x)$  is the higher order terms in Eq. (B.2.3), that is,

$$R_0(x) = \frac{1}{8x}. \quad (\text{B.2.4})$$

For a signal significantly stronger than noise, i.e,  $S \gg \sigma$ , the distribution of noise of Rician type can be approximated by substituting Eq. (B.2.3) in (2.2.1)

$$\begin{aligned} \mathcal{F} &\sim \frac{I}{\sigma^2} \exp\left\{-\left(\frac{I^2 + S^2}{2\sigma^2}\right)\right\} \frac{1}{\sqrt{2\pi\frac{IS}{\sigma^2}}} \exp\left\{\frac{IS}{\sigma^2}\right\} \left(1 + R_0\left(\frac{IS}{\sigma^2}\right)\right) \\ &= \frac{I}{\sigma^2} \frac{\sigma}{\sqrt{2\pi IS}} \exp\left\{-\frac{I^2 + S^2 - 2IS}{2\sigma^2}\right\} \left(1 + R_0\left(\frac{IS}{\sigma^2}\right)\right) \\ &= \left\{\sqrt{I} + \sqrt{I}R_0\left(\frac{IS}{\sigma^2}\right)\right\} \frac{1}{\sqrt{2\pi S\sigma^2}} \exp\left\{-\frac{(I - S)^2}{2\sigma^2}\right\}. \end{aligned} \quad (\text{B.2.5})$$

It is obvious from Eq. (B.2.4) and Eq. (B.2.3) that as  $x$  approaches infinity,

$$\lim_{x \rightarrow \infty} R_0(x) = 0, \quad (\text{B.2.6})$$

and thus<sup>B.2</sup>

$$\lim_{x \rightarrow \infty} I_n(x) \sim \frac{e^x}{\sqrt{2\pi x}}. \quad (\text{B.2.7})$$

---

<sup>B.1</sup>A series  $S_n(z)$  is an asymptotic expansion series of a function  $f(z)$  over the interval of arg  $z$ , and denoted by  $f(z) \sim S_n(z)$  if

$$\lim_{|z| \rightarrow \infty} z^n R_n(z) = 0, \quad (\text{B.2.2})$$

where  $R_n(t) = f(t) - S_n(t)$ . Intuitively, the sum of a finite number of terms of the series  $S_n(z)$  provides a good approximation to  $f(z)$ . For details, see [61] and [60].

<sup>B.2</sup>For similar result, see [82].

With the high order terms vanished, the asymptotic approximation of the Rician distribution in Eq. (B.2.5) can be written as the followings:

$$\lim_{\frac{S}{\sigma} \rightarrow \infty} \mathcal{F} \sim \sqrt{\frac{I}{S}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(I-S)^2}{2\sigma^2}\right\}.$$

As the SNR goes to infinity,  $\frac{I}{S}$  approaches 1, and the distribution of Rician noise,

$$\lim_{\frac{S}{\sigma} \rightarrow \infty} \mathcal{F} \sim \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2\sigma^2}(I-S)^2\right\}. \quad (\text{B.2.8})$$

This is Eq. (2.2.3) which is referred to as the limiting distribution of Rician noise.

## Appendix C

### Noise Parameter $\sigma$ :

### Estimation and Relevant

### Range

#### C.1 Estimation

The statistical behavior of image noise in a magnitude MR image can be modeled by a Rician distribution. The distribution is a function of the SNR of the noisy data. Signal/noise analysis based on a Rician model assumes the knowledge of the noise parameter  $\sigma$  in Eq. (2.2.1). In practice, however, it is an unknown quantity. Estimates from empirical data are usually required. For our purpose, we need a reliable estimator for the parameter in our implementation. Secondly, we also need to estimate the typical range of  $\sigma$  for clinical MR Images in general, and cardiac images in particular.

A common approach takes advantages of the signal-free regions of air in medical images. Let  $\mathcal{N}_M$  denote a subset of the data selected from these

regions. In absence of signal, signal fluctuation in  $\mathcal{N}_M$  represents pure noise. As shown in Appendix B, its statistical behavior follows a Rayleigh distribution with a mean given by[41]

$$\mathcal{E}(I) = \sigma \sqrt{\frac{\pi}{2}}, \quad (\text{C.1.1})$$

This implies an estimator for  $\sigma$  as follows:

$$\tilde{\sigma}_{\mathcal{R}} = \overline{\mathcal{N}_M} \sqrt{\frac{2}{\pi}}. \quad (\text{C.1.2})$$

The sample mean  $\overline{\mathcal{N}_M}$  is defined over the set the signal-free data in  $\mathcal{N}_M$  as

$$\overline{\mathcal{N}_M} = \frac{\sum_{X \in \mathcal{N}_M} X}{\mathbf{Card} \mathcal{N}_M}, \quad (\text{C.1.3})$$

where  $\mathbf{Card} \mathcal{X}$  denotes the cardinality of the data set  $\mathcal{X}$ . A similar estimator is defined over the signal-free background of a squared magnitude image. Let  $\mathcal{N}_S$  denote a set of signal-free data of a squared image. According to Eq. (2.2.7), the mean of  $\chi^2$  noise is twice the  $\sigma^2$ . This relation yields an estimator

$$\tilde{\sigma}_{\chi^2} = \sqrt{0.5 \overline{\mathcal{N}_S}}. \quad (\text{C.1.4})$$

where the term

$$\overline{\mathcal{N}_S} = \frac{\sum_{X \in \mathcal{N}_S} X}{\mathbf{Card} \mathcal{N}_S}, \quad (\text{C.1.5})$$

represents the sample mean of signal-free data in  $\mathcal{N}_S$ .

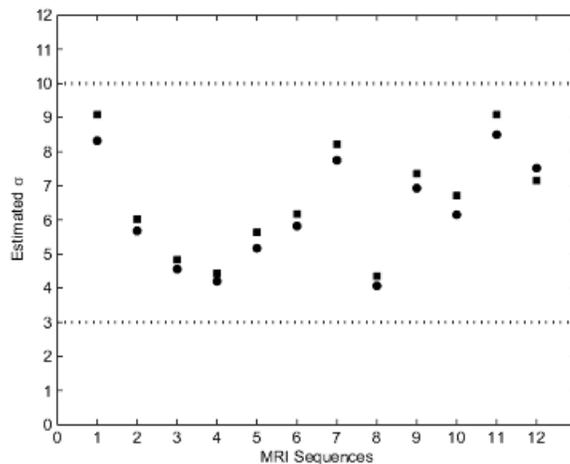


Figure C.1: Estimated  $\sigma$  from Selected Clinical MRI sequences.  $\bullet$  :  $\tilde{\sigma}_{\mathcal{R}}$ .  $\blacksquare$  :  $\tilde{\sigma}_{\chi^2}$ . Seq. 1 - 10: cardiac image sequences; Seq. 11: brain images; and Seq. 12: angiographic images.

## C.2 Relevant Range

Performance analysis with stimulated data generated from a broad range of  $\sigma$ 's show that both  $\tilde{\sigma}_{\mathcal{R}}$  and  $\tilde{\sigma}_{\chi^2}$  yield similar result within small neighborhoods of the true values. To identify the relevant range of  $\sigma$  for clinical MR images, estimation is performed using both estimators on different image sequences, both cardiac and non-cardiac. As shown in Fig. C.1, the noise parameter  $\sigma$  falls within the range between three and ten. This is the reference range that this report uses for the analysis of noise behavior.

## Appendix D

# Moments, Cumulants and Convergence Analysis of Rician Noise

### D.1 Statistical Moments and Cumulants

The moments and cumulants of a distribution are two sequences of descriptive constants that measure the properties of the distribution [67]. The first four members in the sequences provide a useful description about the location and the shape of a distribution. Distributions with a finite number of lower moments or cumulants in common are expected to bear resemblance to each other. In practice, distributions with the first four moments equated result in a remarkably good approximation of one another [67].

Let  $f_X$  be the probability distribution of a random variable  $X$ . The  $r^{th}$

statistical moment from the origin is given by [41]

$$m_{(r)} = \mathbf{E}(X^r) = \int_{-\infty}^{\infty} x^r f_X(x) dx. \quad (\text{D.1.1})$$

The moments of a distribution can be derived from the moment generating function of the distribution. The moment generating function of  $f_X$  is defined as [33; 41]

$$M(t) = \mathbf{E}(\exp\{tX\}) = \int_{-\infty}^{\infty} \exp\{tx\} f_X(x) dx \quad \text{for } M(t) < \infty \quad (\text{D.1.2})$$

The  $r^{\text{th}}$  moment of a distribution  $f_X$  can be defined in terms of its moment generating function as the follows [33; 41]

$$m_{(r)} = M^{(r)}(0) = \left. \frac{d^{(r)} M(t)}{dt^{(r)}} \right|_{t=0} \quad r = 1, 2, 3, \dots \quad (\text{D.1.3})$$

Cumulants are mathematically related to moments. Let  $\kappa_r$  denote the sequence of cumulants of a distribution. The cumulant-generating function of a distribution [33; 41] can be defined in terms of its moment-generating function given in Eq. (D.1.2) as  $\mathcal{K}(t) = \ln M(t)$ , and the sequence of the cumulants are given by [33; 41]

$$\kappa_r = \mathcal{K}^{(r)}(0) = \left. \frac{d^{(r)} \mathcal{K}(0)}{dt^{(r)}} \right|_{t=0}. \quad (\text{D.1.4})$$

There is a close relation between the  $r^{\text{th}}$  cumulant of a distribution and the moments of the distribution. The first four cumulants are given by [67]

$$\kappa_1 = m_1; \quad (\text{D.1.5})$$

$$\kappa_2 = m_2 - (m_1)^2; \quad (\text{D.1.6})$$

$$\kappa_3 = m_3 - 3m_2m_1 + 2(m_1)^3; \quad (\text{D.1.7})$$

$$\kappa_4 = m_4 - 4m_3m_1 - 3(m_2)^2 + 12m_2(m_1)^2 - 6(m_1)^4 \quad (\text{D.1.8})$$

With respect to each of these four cumulants, a comparison between the cumulant of a Rician distribution and the corresponding one of a Gaussian distribution helps to determine the effective range of signal intensity where the behavior of image noise can be reasonably approximated by the limiting distribution associated with infinite signal intensity.

## D.2 Moments and Cumulants of the Gaussian Family

The distribution of a Gaussian random variable  $X \sim \mathcal{N}(\mu, \sigma^2)$  is given by [41]

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \frac{(x - \mu)^2}{\sigma^2} \right\} \quad (\text{D.2.1})$$

The moment generating function<sup>D.1</sup> of a Gaussian distribution is given by

$$M(t) = \exp \left\{ \mu t + \frac{\sigma^2 t^2}{2} \right\}. \quad (\text{D.2.2})$$

Thus the moments can be derived by taking the derivatives of the moment generating function given by Eq. (D.2.2).

$$\frac{dM}{dt} = (\mu + \sigma^2 t) \exp \left\{ \mu t + \frac{\sigma^2 t^2}{2} \right\} \quad (\text{D.2.3})$$

$$m_1 = \left. \frac{dM}{dt} \right|_{t=0} = \mu \quad (\text{D.2.4})$$

According to Eq. (D.1.5), the first cumulant  $\kappa_1 = \mu$  gives the mean.

Now take the derivative of Eq. (D.2.3).

$$\begin{aligned} \frac{d^2 M}{dt^2} &= \sigma^2 \exp \left\{ \mu t + \frac{\sigma^2 t^2}{2} \right\} + (\mu + \sigma^2 t)^2 \exp \left\{ \mu t + \frac{\sigma^2 t^2}{2} \right\} \\ &= \left\{ (\mu + \sigma^2 t)^2 + \sigma^2 \right\} \exp \left\{ \mu t + \frac{\sigma^2 t^2}{2} \right\} \end{aligned} \quad (\text{D.2.5})$$

$$m_2 = \left. \frac{d^2 M}{dt^2} \right|_{t=0} = \mu^2 + \sigma^2 \quad (\text{D.2.6})$$

According to Eq. (D.1.6), the second cumulant  $\kappa_2 = \sigma^2$  gives the variance.

<sup>D.1</sup>Combining Eq. (D.1.2) and Eq. (D.2.1) yields the moment generating function of a Gaussian distribution

$$\begin{aligned} M(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\{tx\} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\left(x^2 - 2(\mu + \sigma^2 t)x + \mu^2\right)\right\} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{\frac{-1}{2\sigma^2}\left[(x^2 - \mu + \sigma^2 t)^2 - (\mu + \sigma^2 t)^2 + \mu^2\right]\right\} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x^2 - \mu + \sigma^2 t)^2}{2\sigma^2}\right\} \exp\left\{-\frac{\mu^2 - (\mu + \sigma^2 t)^2}{2\sigma^2}\right\} dx \\ &= \exp\left\{\frac{2\mu\sigma^2 t + \sigma^4 t^2}{2\sigma^2}\right\} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(x^2 - \mu + \sigma^2 t)^2}{2\sigma^2}\right\} dx \\ &= \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\}. \end{aligned}$$

Repeating this procedure produces the higher moments. The third moment<sup>D.2</sup> and the fourth moment<sup>D.3</sup> are given by

$$m_3 = 3\sigma^2\mu + \mu^3 \quad (\text{D.2.9})$$

$$m_4 = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 \quad (\text{D.2.10})$$

According to Eq. (D.1.7) and Eq (D.1.8), both the third cumulant<sup>D.4</sup>  $\kappa_3$  and the fourth cumulant<sup>D.5</sup>  $\kappa_4$  of a Gaussian distribution are zero.

<sup>D.2</sup>Now take the derivative of Eq. (D.2.5).

$$\begin{aligned} \frac{d^3 M}{dt^3} &= 2\sigma^2(\mu + \sigma^2 t) \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \\ &\quad + \left\{(\mu + \sigma^2 t)^3 + \sigma^2(\mu + \sigma^2 t)\right\} \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \\ &= \left(3\sigma^2(\mu + \sigma^2 t) + (\mu + \sigma^2 t)^3\right) \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \end{aligned} \quad (\text{D.2.7})$$

$$m_3 = \left. \frac{d^3 M}{dt^3} \right|_{t=0} = 3\sigma^2\mu + \mu^3 \quad (\text{D.2.8})$$

<sup>D.3</sup>Now take the derivative of Eq. (D.2.7).

$$\begin{aligned} \frac{d^4 M}{dt^4} &= \left\{3\sigma^4 + 3\sigma^2(\mu + \sigma^2 t)^2\right\} \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \\ &\quad + \left(3\sigma^2(\mu + \sigma^2 t)^2 + (\mu + \sigma^2 t)^4\right) \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \\ &= \left(3\sigma^4 + 6\sigma^2(\mu + \sigma^2 t)^2 + (\mu + \sigma^2 t)^4\right) \exp\left\{\mu t + \frac{\sigma^2 t^2}{2}\right\} \end{aligned} \quad (\text{D.2.11})$$

$$m_4 = \left. \frac{\partial^4 M}{\partial t^4} \right|_{t=0} = 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 \quad (\text{D.2.12})$$

<sup>D.4</sup>According to Eq. (D.1.7), the third cumulant

$$\begin{aligned} \kappa_3 &= 3\sigma^2\mu + \mu^3 - 3(\mu)(\mu^2 + \sigma^2) + 2\mu^3 \\ &= 3\sigma^2\mu + \mu^3 - 3\mu^3 - 3\sigma^2\mu + 2\mu^3 \\ &= 0 \end{aligned} \quad (\text{D.2.13})$$

<sup>D.5</sup>According to Eq. (D.1.8), the fourth cumulant

A Gaussian distribution is completely described by its first two cumulants which give its mean and variance. The third and fourth cumulants of a Gaussian random variable are zero. For any random variable, the distance of its higher cumulants from the origin provides a measurement of how close it is to normality. Standardized cumulants are proposed for this purpose to attain scale invariance of the measurement. The standardized skewness and standardized kurtosis are given by

$$\kappa_3^s = \kappa_3 \kappa_2^{-\frac{3}{2}} \quad (\text{D.2.15})$$

and

$$\kappa_4^s = \kappa_4 \kappa_2^{-2} \quad (\text{D.2.16})$$

respectively [33].

Along this line of reasoning, this report compares the first four cumulants, namely,  $\kappa_1$ ,  $\kappa_2$ ,  $\kappa_3^s$  and  $\kappa_4^s$ , to determine the range of signal intensity where a Gaussian distribution can be used for approximation to the distribution of a Rician variable.

---


$$\begin{aligned}
\kappa_4 &= m_4 - 4m_3m_1 - 3m_2^2 + 12m_2m_1^2 - 6m_1^4 \\
&= 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 - 4(3\sigma^2\mu + \mu^3)(\mu) \\
&\quad - 3(\mu^2 + \sigma^2)^2 + 12(\mu^2)(\mu^2 + \sigma^2) - 6\mu^4 \\
&= 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 - 12\sigma^2\mu^2 \\
&\quad - 4\mu^4 - 3\mu^4 - 6\sigma^2\mu^2 - 3\sigma^4 + 12\mu^4 + 12\sigma^2\mu^2 - 6\mu^4 \\
&= 3\sigma^4 + 6\sigma^2\mu^2 + \mu^4 - (3\sigma^4 + 6\sigma^2\mu^2 + \mu^4) \\
&= 0
\end{aligned} \quad (\text{D.2.14})$$

### D.3 Moments and Cumulants of the Rician Family

According to Eq. (2.2.1) and Eq. (D.1.1), the  $r^{\text{th}}$  statistical moment of a Rician variable given  $S$  and  $\sigma$  is given by

$$\begin{aligned} m_{(r)} &= \int_{-\infty}^{\infty} I^r \frac{I}{\sigma^2} \exp\left(-\frac{I^2 + S^2}{2\sigma^2}\right) I_0\left(\frac{IS}{\sigma^2}\right) 1_{\{x \geq 0\}}(I) dI \\ &= \frac{1}{\sigma^2} \exp\left\{-\frac{S^2}{2\sigma^2}\right\} \int_0^{\infty} I^{r+1} \exp\left\{-\frac{I^2}{2\sigma^2}\right\} I_0\left(\frac{IS}{\sigma^2}\right) dI \end{aligned} \quad (\text{D.3.1})$$

The series expansion of the modified Bessel function of the first kind of  $\nu$  order is given by [2]

$$I_\nu(z) = \left(\frac{z}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{2k}}{k! \Gamma(\nu + k + 1)}.$$

Set  $\nu = 0$  and  $z = \frac{IS}{\sigma^2}$ .

$$I_0\left(\frac{IS}{\sigma^2}\right) = \sum_{k=0}^{\infty} \frac{\left(\frac{IS}{2\sigma^2}\right)^{2k}}{k! \Gamma(k + 1)}. \quad (\text{D.3.2})$$

Thus,

$$\begin{aligned} m_{(r)} &= \frac{1}{\sigma^2} \exp\left\{-\frac{S^2}{2\sigma^2}\right\} \int_0^{\infty} I^{r+1} \exp\left\{-\frac{I^2}{2\sigma^2}\right\} \sum_{k=0}^{\infty} \frac{\left(\frac{IS}{2\sigma^2}\right)^{2k}}{k! \Gamma(k + 1)} dI \\ &= \frac{1}{\sigma^2} \exp\left\{-\frac{S^2}{2\sigma^2}\right\} \sum_{k=0}^{\infty} \left[ \frac{\left(\frac{S}{2\sigma^2}\right)^{2k}}{k! \Gamma(k + 1)} \int_0^{\infty} I^{r+2k+1} \exp\left\{-\frac{I^2}{2\sigma^2}\right\} dI \right]. \end{aligned}$$

Let  $\zeta = \frac{I}{\sqrt{2\sigma}}$ . Then  $d\zeta = \frac{dI}{\sqrt{2\sigma}}$ . Each of the integrals inside the summation can be rewritten as

$$\begin{aligned}
\int_0^{\infty} I^{r+2k+1} \exp\left\{-\frac{I}{2\sigma^2}\right\} dI &= \int_0^{\infty} (\sqrt{2\sigma})^{r+2k+1} \zeta^{r+2k+1} \exp\{-\zeta^2\} \sqrt{2\sigma} d\zeta \\
&= (\sqrt{2\sigma})^{r+2k+2} \int_0^{\infty} \zeta^{r+2k+1} \exp\{-\zeta^2\} d\zeta \\
&= (\sqrt{2\sigma})^{r+2k+2} \frac{1}{2} \Gamma\left(\frac{r}{2} + k + 1\right) \quad (\text{D.3.3})
\end{aligned}$$

The last step uses the Euler's form [9] of the Gamma function, that is,  $\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt$ . Define  $\zeta = \sqrt{t}$ . Then  $dt = 2\zeta d\zeta$ . The Gamma function acquires an alternative form

$$\begin{aligned}
\Gamma(z) &= \int_0^{\infty} e^{-\zeta^2} \zeta^{2(z-1)} 2\zeta d\zeta. \\
&= 2 \int_0^{\infty} e^{-\zeta^2} \zeta^{2z-1} d\zeta
\end{aligned}$$

It follows immediately that the integral is

$$\int_0^{\infty} \zeta^{r+2k+1} \exp\{-\zeta^2\} d\zeta = \int_0^{\infty} \zeta^{2(\frac{r}{2}+k+1)-1} e^{-\zeta^2} d\zeta = \frac{1}{2} \Gamma\left(\frac{r}{2} + k + 1\right) \quad (\text{D.3.4})$$

---

<sup>D.5</sup>The modified Bessel function of the first kind of  $\nu$  order is given by [2]

$$I_{\nu}(z) = \left(\frac{z}{2}\right)^{\nu} \sum_{k=0}^{\infty} \frac{\left(\frac{z}{2}\right)^{2k}}{k! \Gamma(\nu + k + 1)}.$$

Thus, the  $r^{th}$  statistical moment of a Rician distribution is expanded as follows

$$\begin{aligned}
m_{(r)} &= \frac{1}{\sigma^2} \exp \left\{ -\frac{S^2}{2\sigma^2} \right\} \sum_{k=0}^{\infty} \left[ \frac{\left(\frac{S}{2\sigma^2}\right)^{2k}}{k!\Gamma(k+1)} (\sqrt{2}\sigma)^{r+2k+2} \frac{1}{2} \Gamma\left(\frac{r}{2} + k + 1\right) \right] \\
&= \frac{1}{\sigma^2} \exp \left\{ -\frac{S^2}{2\sigma^2} \right\} \frac{1}{2} (2\sigma^2)^{\frac{r}{2}+1} \sum_{k=0}^{\infty} \left[ \frac{1}{k!\Gamma(k+1)} \left(\frac{S}{2\sigma^2}\right)^{2k} (2\sigma^2)^k \Gamma\left(\frac{r}{2} + k + 1\right) \right] \\
&= (2\sigma^2)^{\frac{r}{2}} \exp \left\{ -\frac{S^2}{2\sigma^2} \right\} \sum_{k=0}^{\infty} \left[ \frac{1}{k!\Gamma(k+1)} \left(\frac{S^2}{2\sigma^2}\right)^k \Gamma\left(\frac{r}{2} + k + 1\right) \right] \\
&= (2\sigma^2)^{\frac{r}{2}} \exp \left\{ -\frac{S^2}{2\sigma^2} \right\} \sum_{k=0}^{\infty} \left[ \frac{1}{k!\Gamma(k+1)} \left(\frac{S^2}{2\sigma^2}\right)^k \Gamma\left(\frac{r}{2} + k + 1\right) \frac{\Gamma\left(\frac{r}{2} + 1\right)}{\Gamma\left(\frac{r}{2} + 1\right)} \right] \\
&= (2\sigma^2)^{\frac{r}{2}} \Gamma\left(\frac{r}{2} + 1\right) \exp \left\{ -\frac{S^2}{2\sigma^2} \right\} \sum_{k=0}^{\infty} \left[ \frac{\Gamma\left(\frac{r}{2} + k + 1\right)}{\Gamma(k+1)\Gamma\left(\frac{r}{2} + 1\right)} \frac{1}{k!} \left(\frac{S^2}{2\sigma^2}\right)^k \right] \quad (\text{D.3.5})
\end{aligned}$$

The product of the last two terms in the right member of the equation is the confluent hypergeometric function defined [9] as

$${}_1F_1(a, c; z) = e^z \sum_{k=0}^{\infty} \frac{\Gamma(c-a+k)}{\Gamma(c+k)} \frac{\Gamma(c)}{\Gamma(c-a)} \frac{(-z)^k}{k!} \quad (\text{D.3.6})$$

Set  $a = -\frac{r}{2}$ ,  $c = 1$  and  $z = -\frac{S^2}{2\sigma^2}$ . Eq. (D.3.6) yields

$$\begin{aligned}
{}_1F_1\left(-\frac{r}{2}, 1, -\frac{S^2}{2\sigma^2}\right) &= \exp \left\{ \frac{S^2}{2\sigma^2} \right\} \sum_{k=0}^{\infty} \frac{\Gamma\left(\frac{r}{2} + 1 + k\right)}{\Gamma(k+1)} \frac{\Gamma(1)}{\Gamma\left(\frac{r}{2} + 1\right)} \frac{\left(\frac{S^2}{2\sigma^2}\right)^k}{k!} \\
&= \exp \left\{ \frac{S^2}{2\sigma^2} \right\} \sum_{k=0}^{\infty} \frac{\Gamma\left(\frac{r}{2} + 1 + k\right)}{\Gamma(k+1)\Gamma\left(\frac{r}{2} + 1\right)} \frac{\left(\frac{S^2}{2\sigma^2}\right)^k}{k!} \quad (\text{D.3.7})
\end{aligned}$$

Combining Eq. (D.3.5) and Eq. (D.3.7) yields a function form which gives  $r^{th}$  moments of a Rician distribution associated with  $S$  and  $\sigma^2$  as the follows:

$$m_{(r)} = (2\sigma^2)^{\frac{r}{2}} \Gamma\left(\frac{r}{2} + 1\right) {}_1F_1\left(-\frac{r}{2}, 1; -\frac{S^2}{2\sigma^2}\right) \quad (\text{D.3.8})$$

In summary, Eq. (D.3.8) defines the  $r^{th}$  moments of a Rician distribution.

Eq. (D.1.5) - Eq. (D.1.8) specifies how to compute the first four cumulants of the distribution from the corresponding moments. Finally Eq. (D.2.16) and Eq. (D.2.16) govern the computation of the standardized third and fourth cumulants. According to this set of equations,  $\kappa_1$ ,  $\kappa_2$ ,  $\kappa_3^s$ , and  $\kappa_4^s$  of a Rician distribution can be evaluated numerically.

## D.4 Convergence Analysis of Rician Noise

The convergence range (or the range of convergence) refers to the range of signal intensity  $S$  where a Gaussian distribution is deemed to be a reasonable approximation to a Rician distribution. The infimum<sup>D.6</sup> of this range is referred as the point of convergence, that is, the lowest signal intensity at which the approximation remains reasonable. The cumulants of the Rician distribution associated with the signal intensity  $S$  over the convergence range should approach their Gaussian distribution counterparts. The rest of this appendix presents a convergence analysis which seeks a rough estimation of this range based on the distance between each of corresponding pair of cumulants. For the first and second cumulants, the relative distances are used to measure how close the corresponding cumulants are to each other. The first cumulant of a Gaussian distribution measures the mean. The conditional mean  $\mathbf{E}_{X|S,\sigma}$  of observed data is equal to the true signal  $S$ . The second cumulant measures the variance,  $\sigma^2$ . Let  $\mathcal{R}_{\kappa_i}$  and  $\mathcal{R}_{\kappa_i^s}$  denote the  $i^{th}$  cumulant and the  $i^{th}$  standardized cumulant of a Rician distribution

---

<sup>D.6</sup>Infimum of a set  $\mathcal{A}$ , denoted by  $\inf \mathcal{A}$  is the greatest lowest bound of the set; For details, see [53].

respectively. The relative distances of the first two cumulants are given by

$$D_1|_{S,\sigma} = \frac{|\mathcal{R}_{\kappa_1} - S|}{S} \quad \text{for } S > 0 \quad (\text{D.4.1})$$

$$D_2|_{S,\sigma} = \frac{|\mathcal{R}_{\kappa_2} - \sigma^2|}{\sigma^2}. \quad (\text{D.4.2})$$

Since the higher cumulants of a Gaussian distribution are zero, we use the absolute distances for the standardized third and fourth cumulants – that is,

$$D_3|_{S,\sigma} = |\mathcal{R}_{\kappa_3}^s|; \quad (\text{D.4.3})$$

$$D_4|_{S,\sigma} = |\mathcal{R}_{\kappa_4}^s|. \quad (\text{D.4.4})$$

We may expect the point of convergence  $S^C$  to be located at the lowest signal intensity where for all  $S \geq S^C$   $D_i = 0$ ,  $i = 1, 2, 3, 4$ . In practice,  $D_i$  is not zero for two reasons: (1) the numerical results may depart from the theoretical values and (2) more importantly, a Rician distribution is approximately Gaussian only in the range of infinite signal magnitude. Over the range of a finite signal, a Rician distribution departs from Gaussian, although the difference between the distribution and its approximation distribution diminishes as the signal gets stronger relative to the noise. It is therefore reasonable to expect the point of convergence to be in a neighborhood where the difference between the corresponding cumulants is diminishing in a stable fashion. That is to say, the distance between the cumulants of the two distributions are diminishing to a reasonably small magnitude over the range of signal intensity greater or equal to the point of convergence. In particular, as the signal intensity approaches this point from below, (1) each of these distance measures,  $D_i$ , converges to some small value, and (2) its

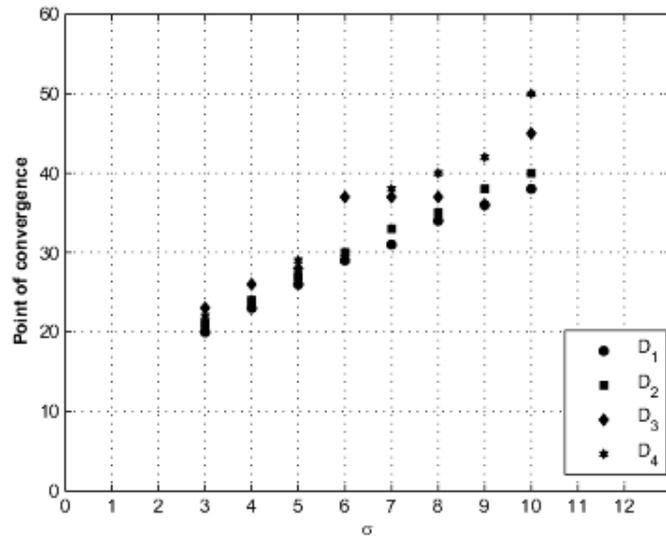
change,  $\frac{\partial D_i}{\partial S}$ , and its rate of change,  $\frac{\partial^2 D_i}{\partial S^2}$ , converge to a point within some considerably small neighborhoods around the asymptote at zero. Moreover, all these measures stay within these neighborhoods over the range of signal intensities above this point.

Figure D.1 summarizes the numerical results<sup>D.7</sup>. Diagram (A) on the left shows  $S_{D_i}^C$  i.e., the signal intensity where each of  $D_i$  starts to converge towards the asymptote at zero. The point of convergence shown in Diagram (B) on the right is given at the intensity where all  $D_i$  converge. That is to say, for a given  $\sigma$ ,  $S^C = \max_{i \in [1 \dots 4]} S_{D_i}^C$ . The relation between  $S^C$  and  $\sigma$  are found to be linear<sup>D.8</sup>, as shown in Diagram B. Obvious is that as  $\sigma$  tends to zero, this linear relation is no longer true. This range, however, has no practical relevance for our purpose. Over a broad reference range where  $\sigma \in [1, 14]$  as shown in the diagram, the linear approximation provides a good description of the data.

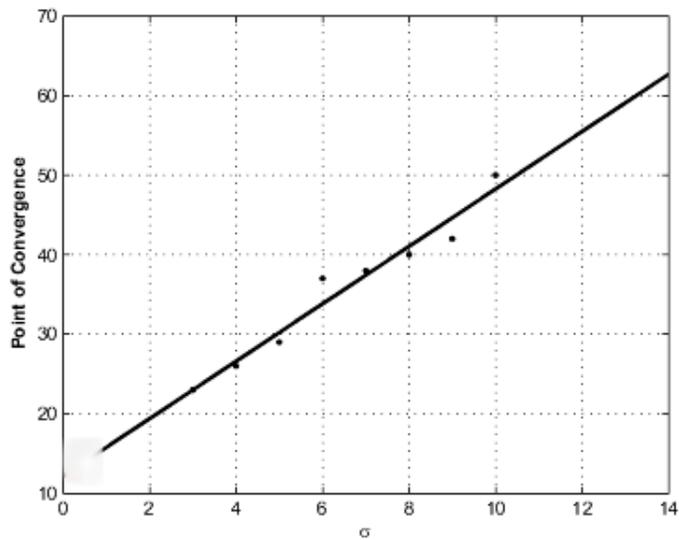
---

<sup>D.7</sup>The details of some selected numerical results are shown in Fig. D.2 - Fig. D.5 at the end of this section.

<sup>D.8</sup>A number of data within and outside the range are obtained for the validation purpose. Polynomials of different orders are used to approximate the relation. The linear approximation provides the predictions which are closest to the validation data.



(A)



(B)

Figure D.1: Point of Convergence of Rician Data. Diagram (A) shows the points of convergence in terms of distance measures with respect to different cumulants at selected  $\sigma$ 's. Diagram (B) shows the linear relation between the point of convergence and  $\sigma$ .

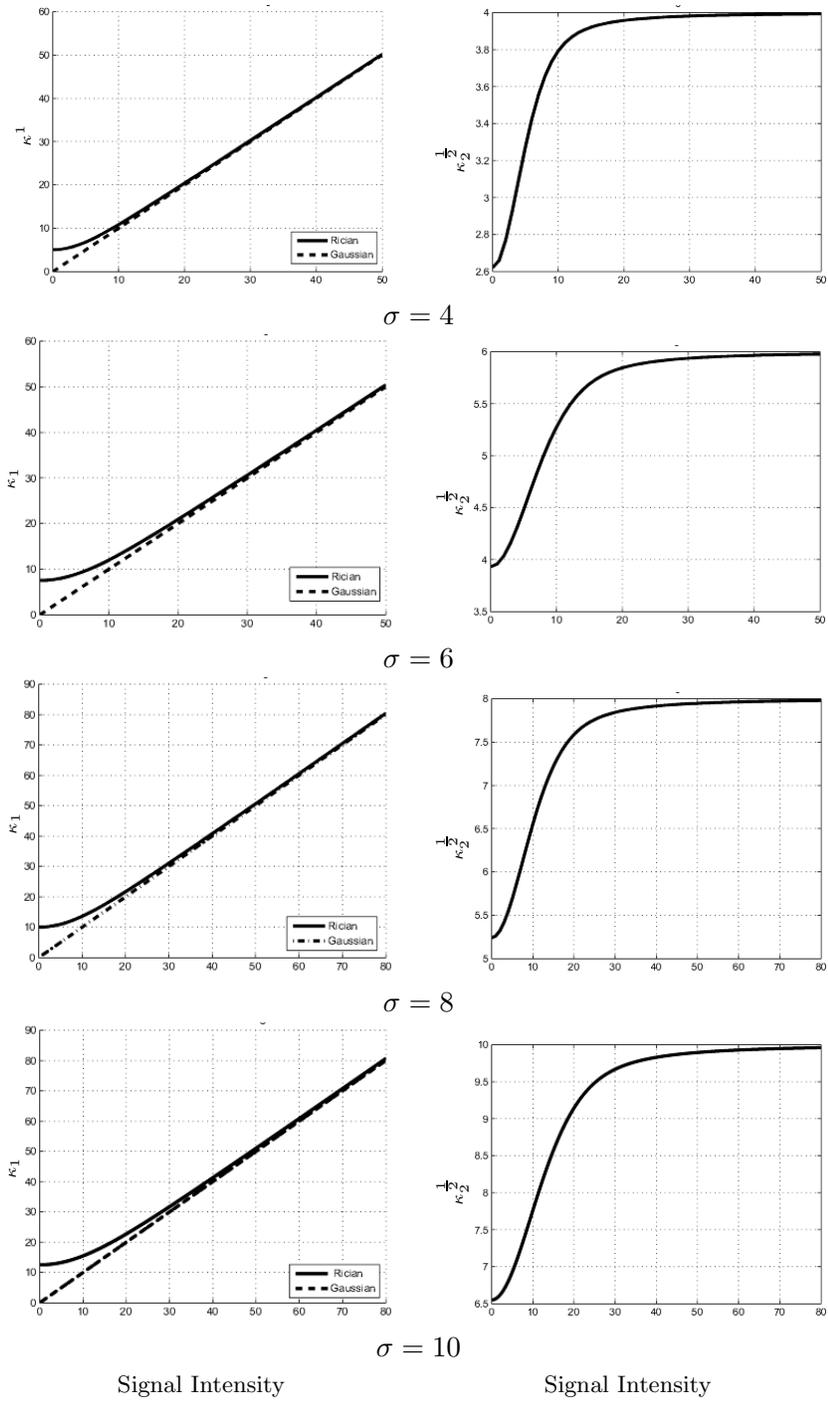


Figure D.2: First and Second Cumulants of Rician Distribution. Right Column: 1<sup>st</sup> cumulant. Dashed line represents the values for a Gaussian variable where  $\kappa_1 = S$ . Left Column: 2<sup>nd</sup> cumulant. In the case of a Gaussian variable,  $\kappa_2^{\frac{1}{2}} = \sigma$ .

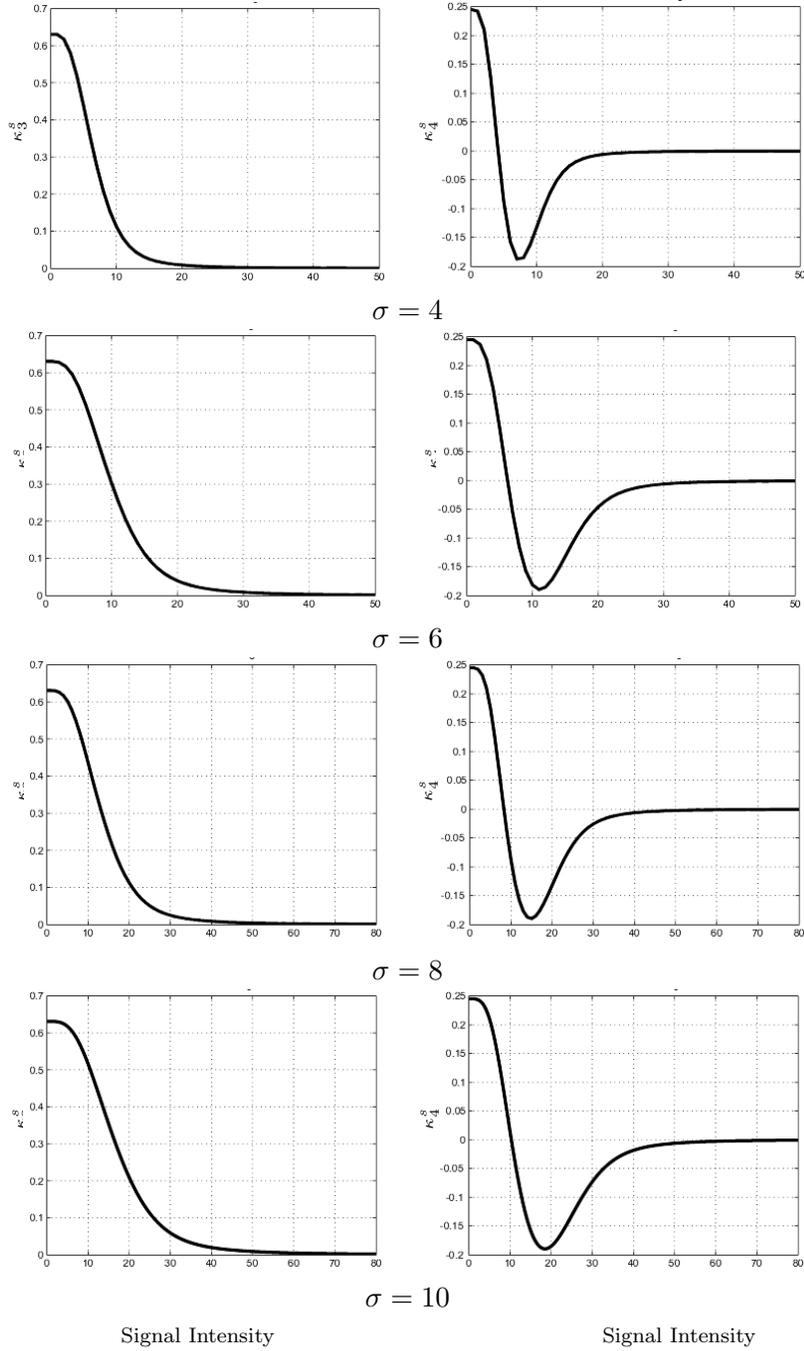


Figure D.3: Standardized Third and Fourth Cumulants of Rician Distribution. Right Column: Standardized third cumulants.  $\kappa_3^s = 0$  for a Gaussian variable. Left Column: Standardized fourth cumulant.  $\kappa_4^s = 0$  for a Gaussian variable.

---

---

 $\sigma = 4$  $\sigma = 6$ 

---

---

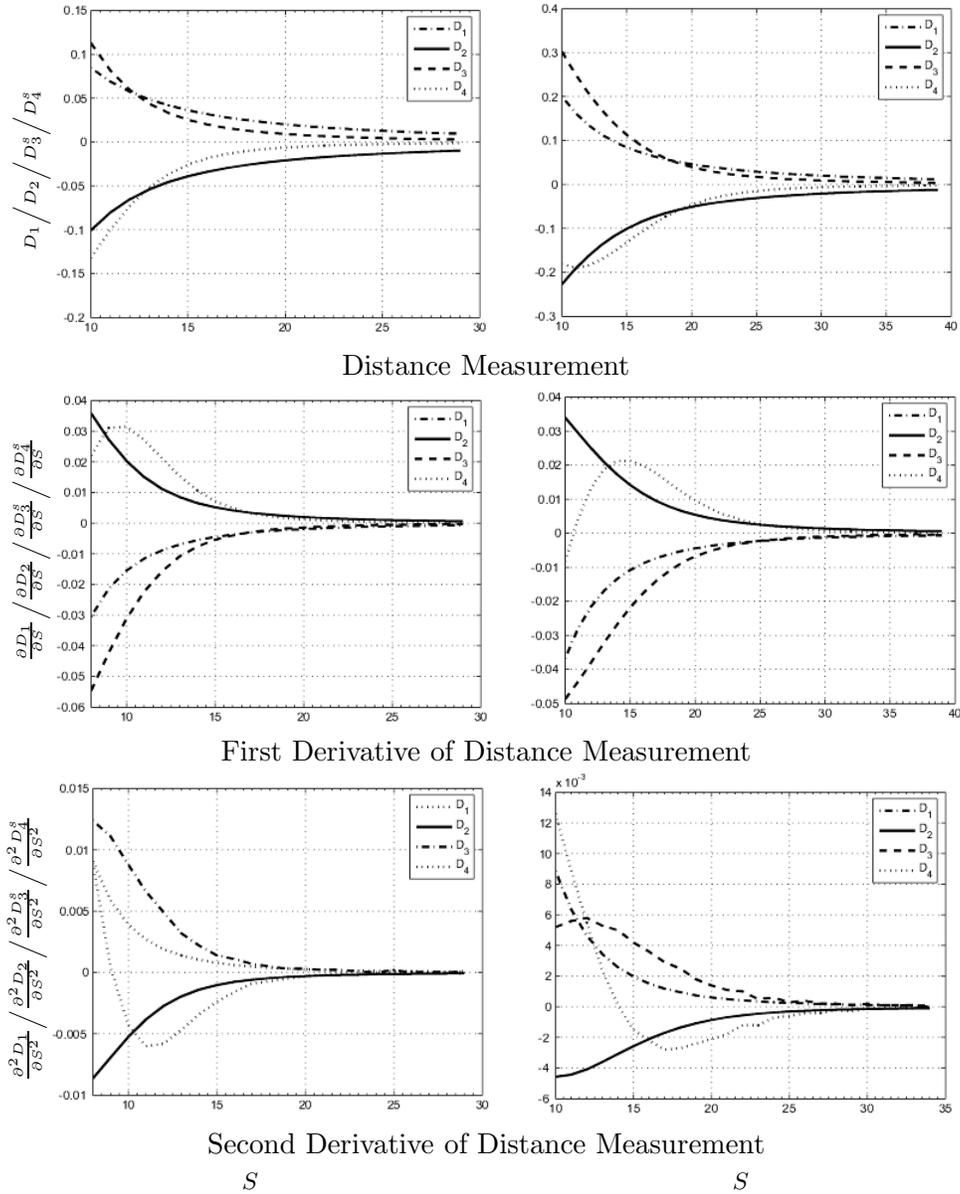


Figure D.4: Convergence of Cumulants of Rician Distribution for  $\sigma = 4$  and  $\sigma = 6$ .

$\sigma = 8$

$\sigma = 10$

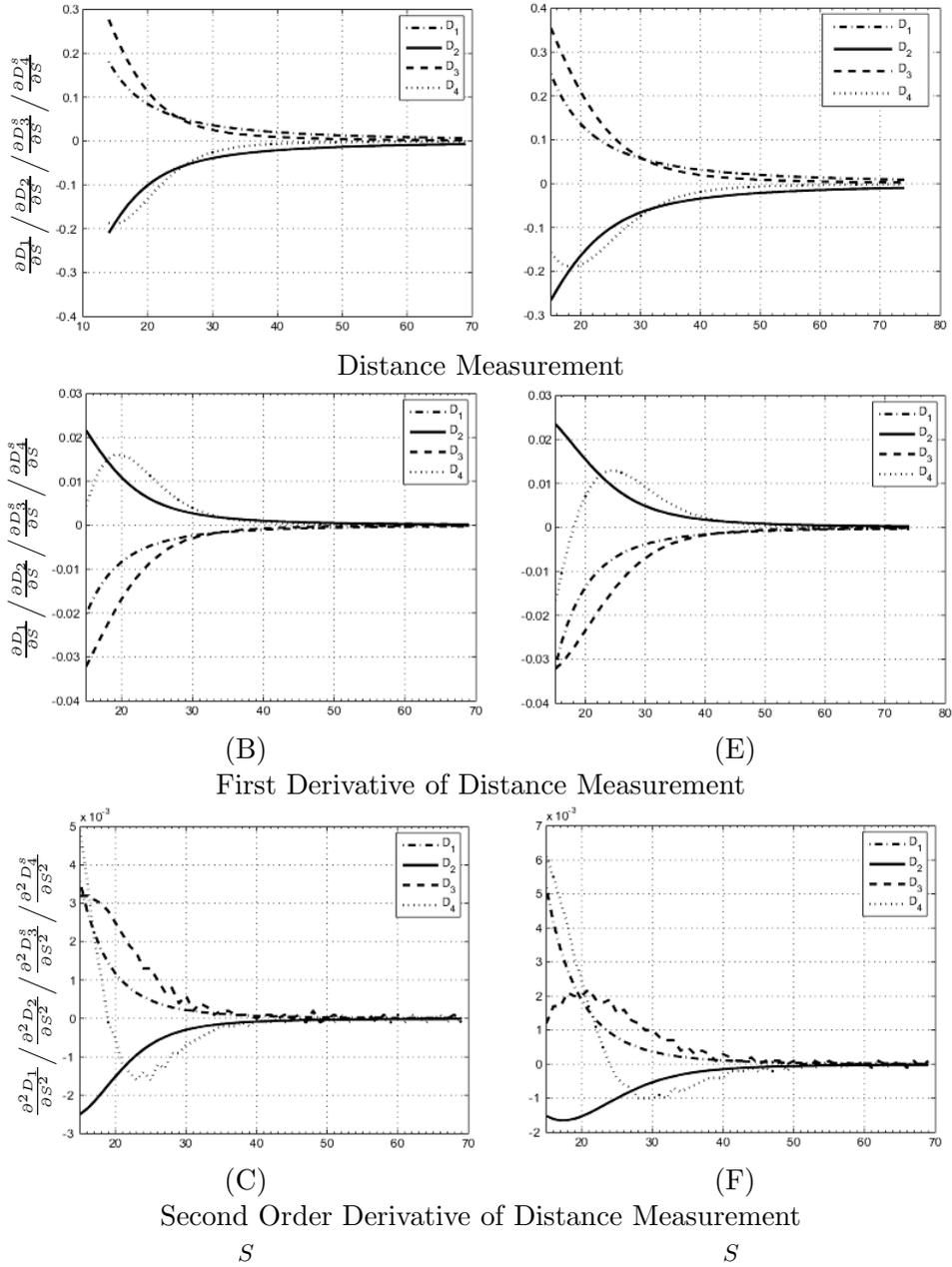


Figure D.5: Convergence of Cumulants of Rician Distribution for  $\sigma = 8$  and  $\sigma = 10$ .

# Appendix E

## Wavelet Shrinkage and Noise Attenuation

### E.1 Introduction

There is a panoply of wavelet-based techniques<sup>E.1</sup> for (Gaussian) noise attenuation. A comprehensive review of these techniques is however beyond the scope of this thesis. Instead, this appendix focuses on the basic ideas of noise attenuation by means of selective reconstruction of image signal, an idea underlying many approaches to wavelet-based noise attenuation.

The general ideas of noise enhancement via selective reconstruction is developed under the assumption of independent Gaussian noise. Consider

---

<sup>E.1</sup>This section focuses on the general framework of wavelet shrinkage which lays a conceptual foundation for different approaches to wavelet-based function approximation and signal smoothing. Many of them incorporate statistical and geometrical modeling of image features in the wavelet domain, such as approximation-theoretic approach[21; 35], Bayesian modeling[23; 26; 27; 44; 71; 72; 90; 93; 103; 102; 110], hypothesis testing[1; 85], singularity detection[57; 75] and spatial context modeling[11; 14; 18; 20; 19; 29; 30; 42; 43; 48; 47; 91; 88; 92; 89; 98]. A general reference can be found in [58; 74; 84; 104]. For approaches to special classes of non-Gaussian noise, see [10; 72; 93; 96].

an approximation model of an image where the observed data,  $\mathcal{I}$ , can be decomposed into two uncorrelated components – (1) noise-free data denoted by  $\Upsilon$  and (2) random noise by  $\xi$ . Our model of the image can be written as

$$\mathcal{I} = \Upsilon + \xi \tag{E.1.1}$$

The noise component  $\xi \sim \mathcal{N}(0, \sigma^2)$  belongs to the Gaussian family of random variables with a mean of zero and a variance of  $\sigma^2$ .

The noise component  $\xi$  is assumed to be a class of high frequency oscillations. This has significant implications for the wavelet representation of the data. The cut-off scale  $J_0$  can be set in such a way that it is reasonable to expect the approximation coefficients  $\{\mathcal{S}_{j_0, \mu}\}$  to be noise free. No such assumption can be made about the wavelet coefficients  $\{\omega_\eta\}$  because wavelets constitute the basis of the detail signal space upon which the high frequency components are projected. The set of wavelet coefficients is thus noisy. The empirical wavelet coefficients<sup>E.2</sup> are the linear combination of the wavelet coefficients  $\Upsilon\omega^\eta$  of the noise-free data and the coefficients  $\xi\omega^\eta$  of the noise component of the observed data. By the linearity of the wavelet transform, the wavelet decomposition is linear with respect to the data components,

---

<sup>E.2</sup>Substitution of the image model given by Eq. (E.1.1) in Eq. (A.0.21) yields the wavelet coefficients for the noisy image  $\{\mathcal{I}\omega_\eta\}$

$$\begin{aligned} \omega_\eta &= \int_{\Omega} \mathcal{I}(\tau) \psi_\eta^*(\tau) d\tau. \\ &= \int_{\Omega} (\Upsilon(\tau) + \xi(\tau)) \psi_\eta^*(\tau) d\tau. \\ &= \int_{\Omega} \Upsilon(\tau) \psi_\eta^*(\tau) d\tau + \int_{\Omega} \xi(\tau) \psi_\eta^*(\tau) d\tau \\ &= \Upsilon\omega_\eta + \xi\omega_\eta \end{aligned} \tag{E.1.2}$$

i.e.,

$$\mathcal{I} = \underbrace{\sum_{\mu} \mathcal{S}_{j_0, \mu} \phi_{j_0, \mu}^{2D}}_{\Upsilon} + \sum_{\eta} \Upsilon_{\omega_{\eta}} \psi_{\eta} \quad + \quad \underbrace{\sum_{\eta} \xi_{\omega_{\eta}} \psi_{\eta}}_{\xi} \quad (\text{E.1.3})$$

The obvious implication of Eq. (E.1.3) is that the noise-free signal  $\Upsilon$  can be recovered by removing the noise-component of the wavelet coefficient set. A diverse number of approaches have been proposed to estimate the noise-free coefficients. Many of these approaches shares a common point of departure in the localization and energy compactness properties<sup>E.3</sup> of multiresolution analysis. The presence of random, uncorrelated noise changes the empirical coefficients: a large number of non-zero but low coefficients that contribute noise and a few large coefficients contributing signal<sup>E.4</sup>. This observation leads to the idea of signal approximation by way of selective wavelet reconstruction.

The noise-free signal can be approximately reconstructed with a selected set of significant coefficients. Pivotal to the approximation is a decision rule by which observed coefficients are modified according to a threshold set which discriminates the significant coefficients from the non-significant ones. The Gaussian noise model allows a variety of coefficient selection policies proposed for different applications. Despite their diversity, most of them are variants on the general idea of assuming that the low energy coefficients represent noise, the image is reconstructed with only a subset of largest coefficients according to some threshold set.

---

<sup>E.3</sup>For these properties, see Appendix A.

<sup>E.4</sup>See [37; 20].

## E.2 SU Decision Rule: Wavelet Shrinkage with the Universal Threshold

Suppose the image model (E.1.1) with additive noise, generated by an *i.i.d.* Gaussian process, with a mean of zero and a variance of  $\sigma^2$ . A series of seminal papers [37; 38; 39] propose a family of decision rules which are asymptotically minimax <sup>E.5</sup> within logarithmic terms over a broad space of images.

First, given an appropriate threshold, two wavelet shrinkage rules stand out as classic exemplars in the literature of wavelet shrinkage. The 'keep-and-kill' procedure leads to a hard thresholding estimator. Given a threshold  $\lambda_{\mathcal{T}}$ , the estimator of the noise-free coefficients is given by

$$\tilde{\omega}_{\eta}|_{\lambda_{\mathcal{T}}} = \begin{cases} \omega_{\eta} & \text{if } |\omega_{\eta}| > \lambda_{\mathcal{T}} \\ 0 & \text{if } |\omega_{\eta}| \leq \lambda_{\mathcal{T}} \end{cases} \quad (\text{E.2.3})$$

---

<sup>E.5</sup>The risk of approximation (or estimation) is the average loss that will be incurred by using an estimator  $\hat{S}$  to obtain an approximation from the observed data; see fn.2.27 and the reference thereof. The minimax risk is defined as

$$\mathbf{R}^{mm}(\mathcal{F}) = \inf_{\hat{S} \in \mathcal{F}} \sup_{\mathcal{S}} \mathbf{R}(\mathcal{S}, \hat{S}), \quad (\text{E.2.1})$$

where  $\mathcal{F}$  is the set of all possible estimators of  $\mathcal{S}$ . The supremum in Eq. (E.2.1) is referred to as the maximal risk of the estimator  $\hat{S}$  on the set  $\mathcal{F}$ . The minimax risk provides a yardstick for how well one can estimate a signal  $\mathcal{S}$  over the class  $\mathcal{F}$ . A minimax estimator is the estimator that achieves the minimax risk. An estimator  $\hat{S}$  of  $\mathcal{S}$  attains an optimal rate of convergence if

$$\sup_{\mathcal{S} \in \mathcal{F}} \mathbf{E} \|\hat{S} - \mathcal{S}\|_{\mathcal{L}_2}^2 \asymp \mathbf{R}^{mm}(\mathcal{F}) \quad (\text{E.2.2})$$

where  $\asymp$  refers to an asymptotic identical relation up to a constant factor. Decision rules of wavelet shrinkage including the hard- and soft-thresholding rule with the universal threshold attains the optimal rate within a logarithmic factor. For details, see [69; 54; 63] and [40].

The soft thresholding rule on the other hand reduces noise by attenuating the wavelet coefficients for all coefficients larger than the threshold. The coefficient estimator is given by

$$\tilde{\omega}_\eta \Big|_{\lambda_{\mathcal{T}}} = \begin{cases} \omega_\eta - \lambda_{\mathcal{T}} & \text{if } \omega_\eta \geq \lambda_{\mathcal{T}} \\ \omega_\eta + \lambda_{\mathcal{T}} & \text{if } \omega_\eta \leq -\lambda_{\mathcal{T}} \\ 0 & \text{if } |\omega_\eta| < \lambda_{\mathcal{T}} \end{cases} \quad (\text{E.2.4})$$

Compromise between hard and soft thresholding rules leads to a range of schemes<sup>E.6</sup> in an attempt to increase the stability and reduce the bias of the estimators for a finite sample. The semisoft shrinkage rule, which use a firm shrinkage function, for example, is a generalized shrinkage scheme [46]:

$$\tilde{\omega}_\eta \Big|_{\lambda_{\mathcal{T}_1}, \lambda_{\mathcal{T}_2}} = \begin{cases} 0, & \text{if } \omega_{j,\mu} \leq \lambda_{\mathcal{T}_1}, \\ \text{sgn}(\omega_{j,\mu}) \frac{\lambda_{\mathcal{T}_2} (|\omega_{j,\mu}| - \lambda_{\mathcal{T}_1})}{\lambda_{\mathcal{T}_2} - \lambda_{\mathcal{T}_1}}, & \text{if } \lambda_{\mathcal{T}_1} < \omega_{j,\mu} \leq \lambda_{\mathcal{T}_2}, \\ \omega_{j,\mu}, & \text{if } \omega_{j,\mu} > \lambda_{\mathcal{T}_2}. \end{cases} \quad (\text{E.2.5})$$

This scheme reduces to hard thresholding with  $\lambda_{\mathcal{T}_1} = \lambda_{\mathcal{T}_2}$  and to soft thresholding with  $\lambda_{\mathcal{T}_2} = \infty$ . In general, semisoft shrinkage does not attenuate every coefficients, as shown in Fig. 2.8(c). The distance of a coefficient from zero in magnitude measures the significance of a coefficient. A coefficient at a distance within one  $\lambda_{\mathcal{T}_1}$  from the origin is shrunk to zero. On the other hand, large coefficients with absolute magnitudes greater than  $\lambda_{\mathcal{T}_2}$  are immune to shrinkage.

Threshold selection is critical. On the one hand, thresholds should be

---

<sup>E.6</sup>These schemes include ‘hyperbola’ rule [110], non-negative garrote rules [45], and others

set above the noise with a high probability. On the other hand, too high of a threshold will remove most of the signal that is supposed to be recovered. Suppose that an oracle gives the ideal threshold  $\lambda_{\mathcal{T}}^{Oracle}$  which defines the significance of coefficients in such a way that the risk of estimation using a decision rule with the ideal threshold is the lowest among all the linear and non-linear rules which attenuates the coefficients. The threshold rules using the universal threshold [38] given by

$$\lambda_{\mathcal{T}}^{Univ} = \sigma \sqrt{(2 \ln N)}, \quad (\text{E.2.6})$$

where  $\sigma$  is the standard derivation of the Gaussian noise and  $N$  is the number of wavelet coefficients  $\omega_{\eta}$ , achieves the lowest bound within an algorithmic factor[37]. In practice, the variance  $\sigma^2$  is usually unknown. In this case, the standard deviation  $\sigma$  can be estimated from the wavelet coefficients at the finest scale, assuming that most of the signal oscillations captured by the observed wavelet coefficients at these subbands are noise. The standard deviation  $\sigma$  of noise may be estimated<sup>E.7</sup> using the median absolute deviation<sup>E.8</sup> (MAD) of the coefficients at the finest level<sup>E.9</sup>. Let  $\hat{\eta}$  be the subset of indices for the finest level of wavelet expansion, the estimated noise standard deviation

$$\tilde{\sigma} = \frac{MAD(\{\omega_{\hat{\eta}}\})}{0.6745} \quad (\text{E.2.7})$$

---

<sup>E.7</sup>For details, see [36; 37].

<sup>E.8</sup>For the definition, see fn. 2.29.

<sup>E.9</sup>For details, see [74; 88].

In general, the universal threshold is selected for a number of good properties. First, the soft-thresholding rule with the threshold achieves a "nearly" minimax risk which is close to the oracular risk up to the factor of  $1 + 2\ln N$ ; see [37]. With a high probability, the universal threshold is set at the asymptotic upper bound of the noise. The resulting approximation is no more rough than the underlying signal and much of the noise is removed from the observed data. Sharp signal discontinuities remain sharp in the approximation. Nevertheless, by using a universal threshold set at the asymptotic upper bound of noise, the algorithm prefers killing (the noise) to preserving (the signal). Other approaches for threshold selection are proposed to improve the quality of approximation, such as data adaptive threshold based on Stein's unbiased risk estimator (SURE) to minimize the estimation risk [38]; or thresholds based on cross-validation functions and generalized cross-validation functions to estimate the mean squared error [58; 59; 74; 81; 122].

### **E.3 MSE Decision Rule: Wavelet Shrinkage with Minimum Mean Squared Error (MSE)**

A mean squared error (MSE) approach is proposed for the high SNR MR images<sup>E.10</sup>. It aims at an estimate of the the noise-free component of each empirical wavelet coefficient that minimizes the mean squared error (MSE) of estimation of the coefficient. Let  $\tilde{\omega}_\eta^{MSE}$  be the estimated wavelet coefficient. Each

---

<sup>E.10</sup>As reviewed in Section 1.2, this approach in its original formulation starts with the assumption that MR images can be classified into high SNR and low SNR images; see [51], and [83].

coefficient is shrunk toward zero according to a shrinkage factor<sup>E.11</sup>:

$$\tilde{\omega}_\eta = \alpha_\eta \omega_\eta \quad (\text{E.3.1})$$

Assume that an empirical wavelet coefficient is an unbiased estimator of the noise-free coefficient,

$$\mathbf{E}(\omega_\eta) = \Upsilon_{\omega_\eta}. \quad (\text{E.3.2})$$

An estimate of the noise-free coefficient with the minimum mean squared error (MSE) is the solution to the following optimization problem

$$\Xi = \mathbf{E} \left[ \left( \Upsilon_{\omega_\eta} - \tilde{\omega}_\eta \right)^2 \right] \longrightarrow \min. \quad (\text{E.3.3})$$

where  $\Upsilon_{\omega_\eta}$  is the noise-free component<sup>E.12</sup> of the coefficients. By the signal model given by Eq. (E.3.1) and the linearity of the expectation operator, the mean squared error of estimation can be expressed as follows:

$$\begin{aligned} \Xi &= \mathbf{E} \left[ \left( \Upsilon_{\omega_\eta} - \tilde{\omega}_\eta \right)^2 \right]; \\ &= \mathbf{E} \left[ \left( \Upsilon_{\omega_\eta} - \alpha_\eta \omega_\eta \right)^2 \right]; \\ &= \mathbf{E} \left[ \Upsilon_{\omega_\eta}^2 - 2\Upsilon_{\omega_\eta} \alpha_\eta \omega_\eta + \left( \alpha_\eta \omega_\eta \right)^2 \right]; \\ &= \Upsilon_{\omega_\eta}^2 - 2\Upsilon_{\omega_\eta} \alpha_\eta \mathbf{E}[\omega_\eta] + \alpha_\eta^2 \mathbf{E}[\omega_\eta^2]. \end{aligned} \quad (\text{E.3.4})$$

The transition to the last step is allowed by the fact that both the noise-free component of a wavelet coefficient  $\Upsilon_{\omega_\eta}$  and the shrinkage factor  $\alpha_\eta$  are non-random quantities. With the assumption given by Eq. (E.3.2), the necessary

---

<sup>E.11</sup>These factors are called filtering weights in [51], and [83].

<sup>E.12</sup>See Eq. (E.1.2).

condition for  $\alpha_\eta$  to be a stationary point gives the following equality:

$$\Upsilon_{\omega_\eta} \mathbf{E}[\omega_\eta] = \alpha_\eta \mathbf{E}[\omega_\eta^2] \quad (\text{E.3.5})$$

With the assumption given by Eq. (E.3.2) that  $\mathbf{E}(\omega_\eta) = \Upsilon_{\omega_\eta}$ , and the well-known properties[64] that

$$\text{var}(x) = \mathbf{E}[x^2] - (\mathbf{E}[x])^2, \quad (\text{E.3.6})$$

Eq (E.3.5) gives the shrinkage factor

$$\begin{aligned} \alpha_\eta^{MSE} &= \frac{\Upsilon_{\omega_\eta^2}}{\mathbf{E}[\omega_\eta^2]} \\ &= \frac{\Upsilon_{\omega_\eta^2}}{\Upsilon_{\omega_\eta^2} + \text{var}(\omega_\eta)} \end{aligned} \quad (\text{E.3.7})$$

The second derivative of  $\Xi$  with respect to  $\alpha_\eta$  is  $\mathbf{E}[\omega_\eta^2] > 0$ . According to the second derivative test,  $\omega_\eta^{MSE} = \alpha_\eta^{MSE} \omega_\eta$ , attains the minimum  $\Xi$ , the mean squared error of estimation of the noise-free coefficient.

This optimal shrinkage factor is unknown in practice without knowledge of the noise-free component  $\Upsilon_{\omega_\eta}$  and the variance  $\text{var}(\omega_\eta)$  of the empirical coefficient. Following from the assumption given in Eq. (E.3.2) and the properties given in Eq. (E.3.6) is that  $\mathbf{E}(\omega_\eta^2) = \Upsilon_{\omega_\eta^2} + \text{var}(\omega_\eta)$ . Thus, the noise-free coefficient can be estimated by

$$\Upsilon_{\omega_\eta^2} \approx \omega_\eta^2 - \text{var}(\omega_\eta). \quad (\text{E.3.8})$$

The optimal (MSE) shrinkage factor given in Eq. (E.3.7) can be approxi-

mated by the following shrinkage rule[51; 83]:

$$\widetilde{\alpha}_\eta^{MSE} = \mathbf{H}\left(\frac{\omega_\eta^2 - \text{var}(\omega_\eta)}{\omega_\eta^2}\right) \quad (\text{E.3.9})$$

where  $\mathbf{H}$  is the Heaviside step function<sup>E.13</sup>. The shrinkage factor given in Eq. (E.3.9), bounded over the interval,  $1 \leq \widetilde{\alpha}_\eta^{MSE} \leq 0$ , attenuates the absolute magnitude of the empirical coefficient toward zero. The variance of the coefficient can be estimated by  $\tau\sigma^2$  where  $\tau \geq 1$  and  $\sigma$  can be estimated from the image background as detailed in Appendix C. According to [83], setting the parameter  $\tau$  to two leads to good results.

It should be emphasized that the MSE properties of the shrinkage rule defined in Eq. (E.3.1) with the shrinkage factor given by Eq. (E.3.9) only holds for the convergence range of magnitude MR images. It is obvious from the foregoing discussion that the optimality of the shrinkage factor as given by Eq. (E.3.9) relies heavily on the assumption given in Eq.(E.3.2) which in general is not valid outside the convergence range.

## E.4 SSR Decision Rule: Semisoft Shrinkage with Rician based Thresholds

The last section of this appendix aims to explore the problem of integrating the properties of Rician noise in a shrinkage rule. For our purpose, it is desirable to have the thresholds selected with some consideration of the the-

---

<sup>E.13</sup> The Heaviside step function is given by

$$\mathbf{H}(x) = \begin{cases} 0 & : x < 0 \\ 1 & : x \geq 0 \end{cases} \quad (\text{E.3.10})$$

For details see [60].

oretical properties of image noise. The second moment of a Rician variable given by [49; 65; 101]:

$$m_{(2)}(I) = S^2 + 2\sigma^2 \quad (\text{E.4.1})$$

is signal dependent. Over the range of convergence, the first moment is converging to  $S$  up to a diminishing bias. Assuming that the bias is negligible within the convergence range, i.e.,  $m_1 = S$ , the variance of Rician noise approaches  $2\sigma^2$ . This bound is also consistent with the experimental findings in [83] which suggests  $2\sigma^2$  is a better choice for estimating the noise-free wavelet coefficients. At a very low intensity range, as  $S$  approaches zero, the variance of Rician noise is given by [41]  $\text{var}(I) = (2 - \frac{\pi}{2})\sigma^2$ . Thus, noise variance is bounded approximately by the following interval:

$$\left[ \sigma \sqrt{2 - \frac{\pi}{2}}, \sigma \sqrt{2} \right].$$

These theoretical bounds motivate the following decision rule which is referred to as the *SSR* decision rule or *SSR* rule for short. Consider the semisoft shrinkage rule defined in Eq. (E.2.5).

$$\lambda_{\mathcal{T}_k} = \hat{\sigma}_k \sqrt{2 \ln N} \quad \text{for } k = 1, 2. \quad (\text{E.4.2})$$

where

$$\hat{\sigma}_1 = \hat{\sigma} \sqrt{2 - \frac{\pi}{2}}; \quad (\text{E.4.3})$$

$$\hat{\sigma}_2 = \hat{\sigma} \sqrt{2}. \quad (\text{E.4.4})$$

The parameter  $\sigma$  from Eq. (2.2.1) and Eq. (2.2.3) is in general unknown, but an accurate estimate  $\hat{\sigma}$  can be obtained from the signal-free data in the

image background<sup>E.14</sup>.

---

<sup>E.14</sup>For details see Appendix C.

## Appendix F

# $\chi^2$ Noise in Squared Magnitude Images

A squared magnitude image  $I^S$  arises mathematically from a sum of squares transformation<sup>F.1</sup> on a pair of images, each with additive Gaussian noise distributed  $\mathcal{N}(0, \sigma^2)$ . The conditional distribution of the resulting data set  $I^S$  is transformed from a Gaussian distribution to a noncentral  $\chi'^2$  distribution[51; 52; 62; 83]:

$$f(I^S | S, \sigma) \sim \sigma^2 \chi'^2(\lambda)$$

with two degrees of freedom ( $\nu = 2$ ) and the non-centrality parameter  $\lambda = \frac{S^2}{\sigma^2}$ . The probability density function of a noncentral  $\chi'^2(\lambda)$  variate is given by

$$f(z | \nu, \lambda) = 2^{-\nu/2} \exp\left\{-\frac{z + \nu}{2}\right\} z^{\frac{\nu-2}{2}} \sum_{r=0}^{\infty} \frac{\lambda^r z^r}{2^{2r} r! \Gamma(\frac{1}{2}\nu + r)} \quad (\text{F.0.1})$$

---

<sup>F.1</sup>see Fig. 2.2.

with mean and variance given as the following:

$$\mathbf{E}\left(\chi_{\nu}^{\prime 2}(\lambda)\right) = \nu + \lambda \quad (\text{F.0.2})$$

$$\text{var}\left(\chi_{\nu}^{\prime 2}(\lambda)\right) = 2(\nu + 2\lambda); \quad (\text{F.0.3})$$

see [41; 62] and [106]. The conditional mean and variance of squared magnitude signals is therefore [83]

$$\mathbf{E}_{I|S^2,\sigma} = S^2 + 2\sigma^2 \quad (\text{F.0.4})$$

$$\text{var}_{I|S^2,\sigma} = 4S^2\sigma^2 + 4\sigma^4 \quad (\text{F.0.5})$$

## Part II

# Output Data Sets of Performance Evaluation

The purpose of this appendix is to show the input/output images of performance evaluation. These data are organized by cases of study. For each test case, the data are presented in a matrix form. Each row shows an image which is either an input or output of an enhancement procedure. Of each of these images, three representations are displayed: (1) the given image on the left; (2) the contrast-adjusted image in the second column; and (3) the edge map on the right.

## Appendix G

# Synthetic Test Cases

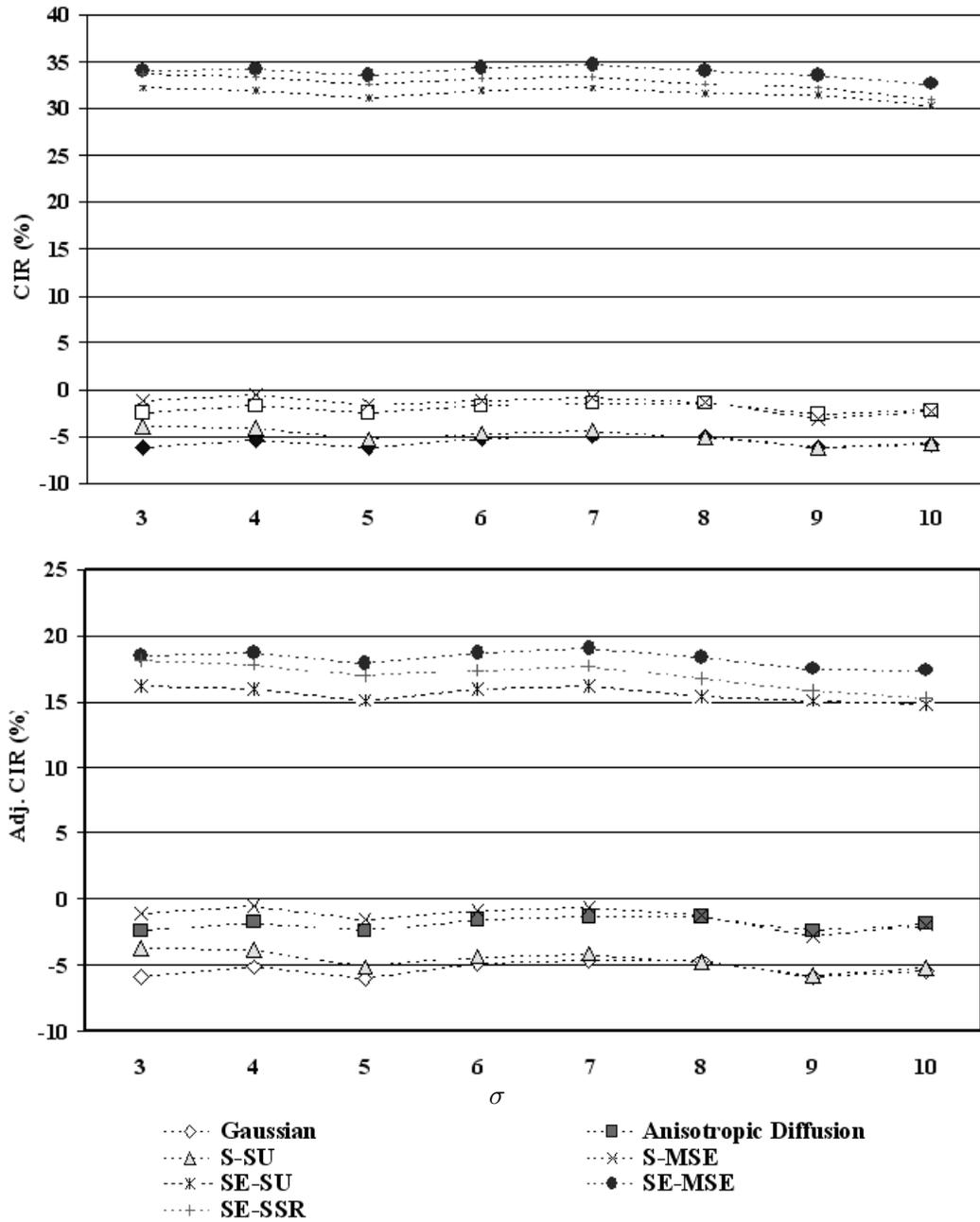


Figure G.1: Synthetic Images. Comparison of contrast improvement across algorithms on test images corrupted by noise with  $\sigma$  over the range from 3 to 10

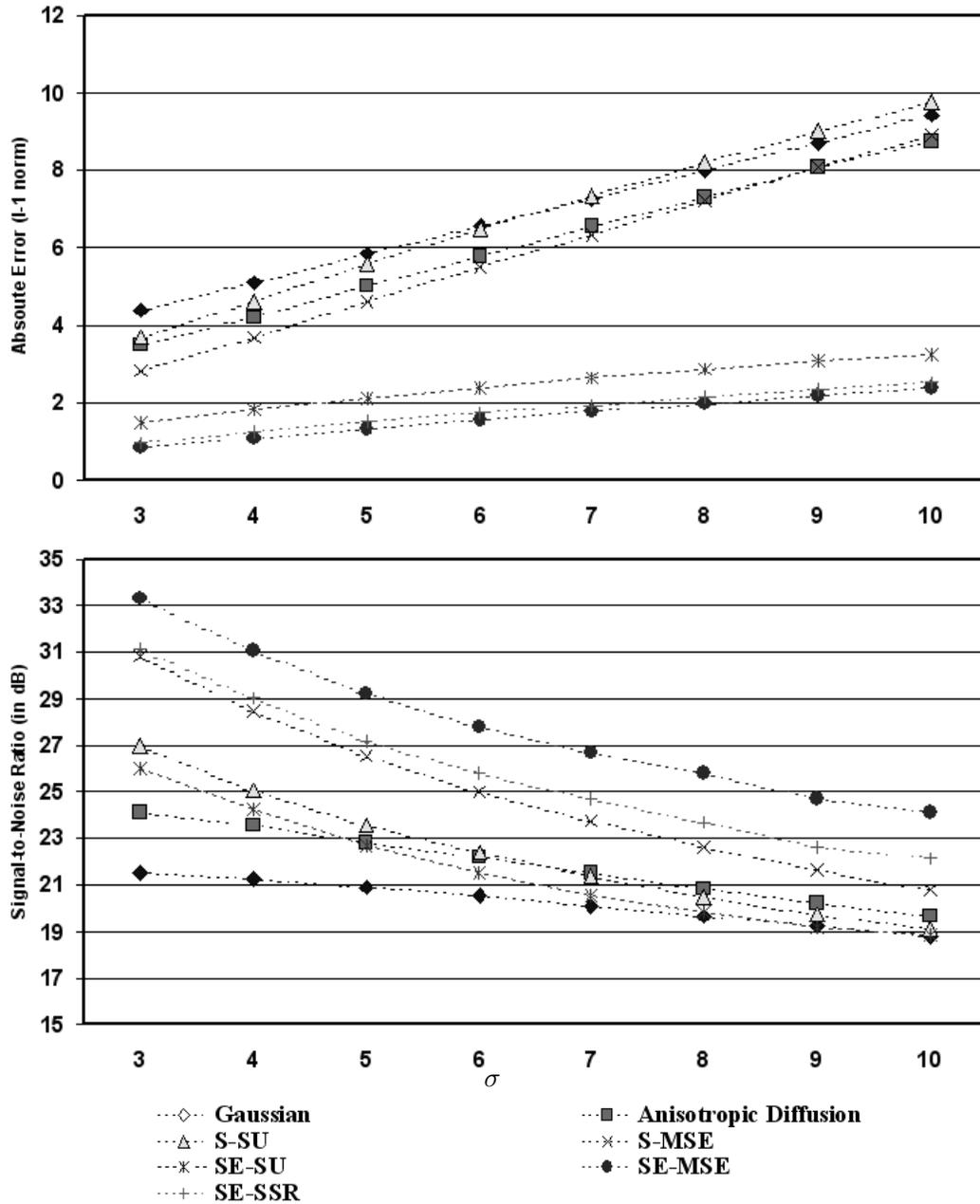


Figure G.2: Synthetic Images. Comparison of absolute error with respect to  $l_1$  norm and SNR with respect to  $l_2$  norm across algorithms on test images corrupted by noise with  $\sigma$  over the range from 3 - 10.

## Appendix H

# Medical Test Cases

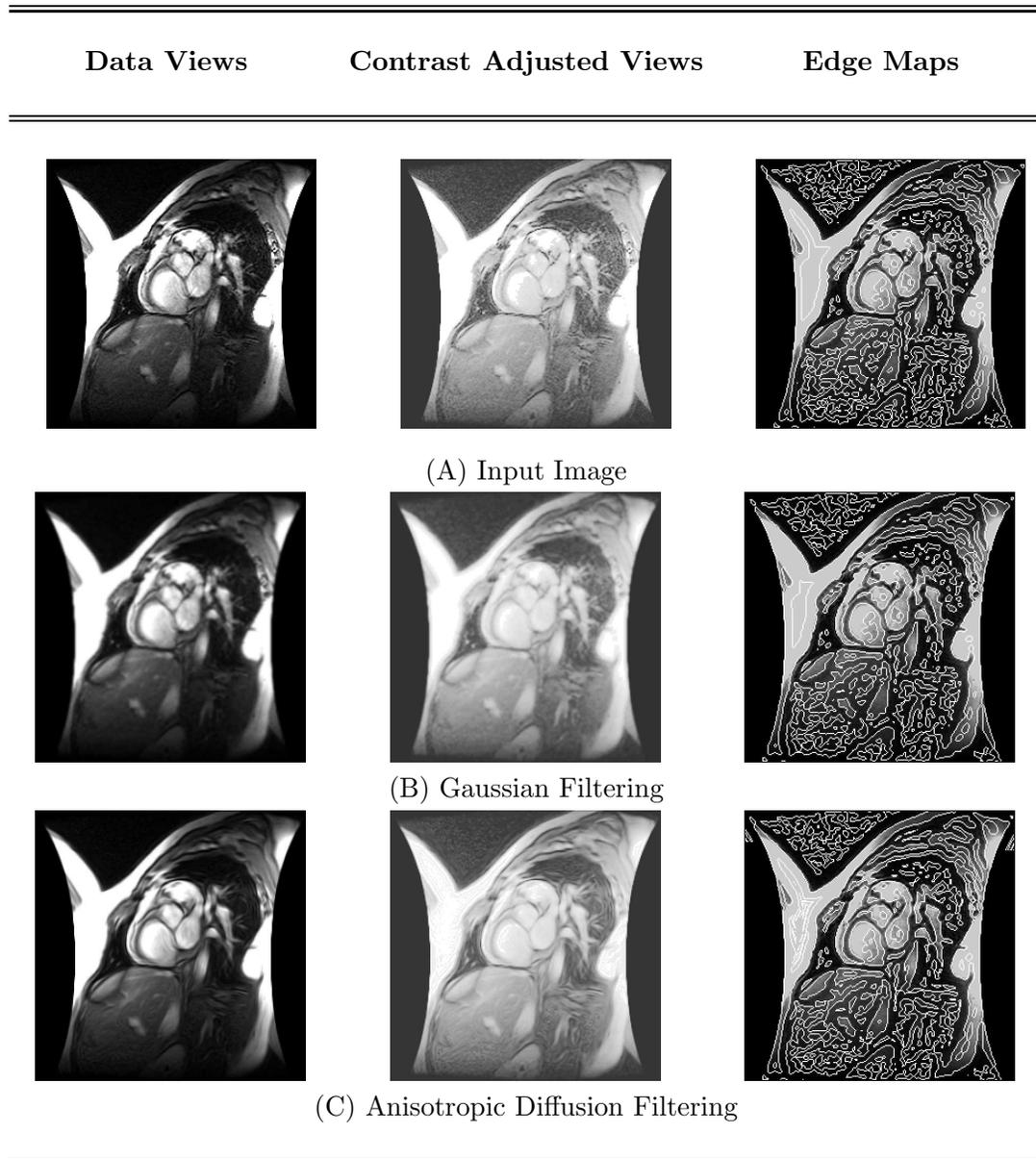


Figure H.1: Cardiac Image: SRS0000-10. Diffusion Group.

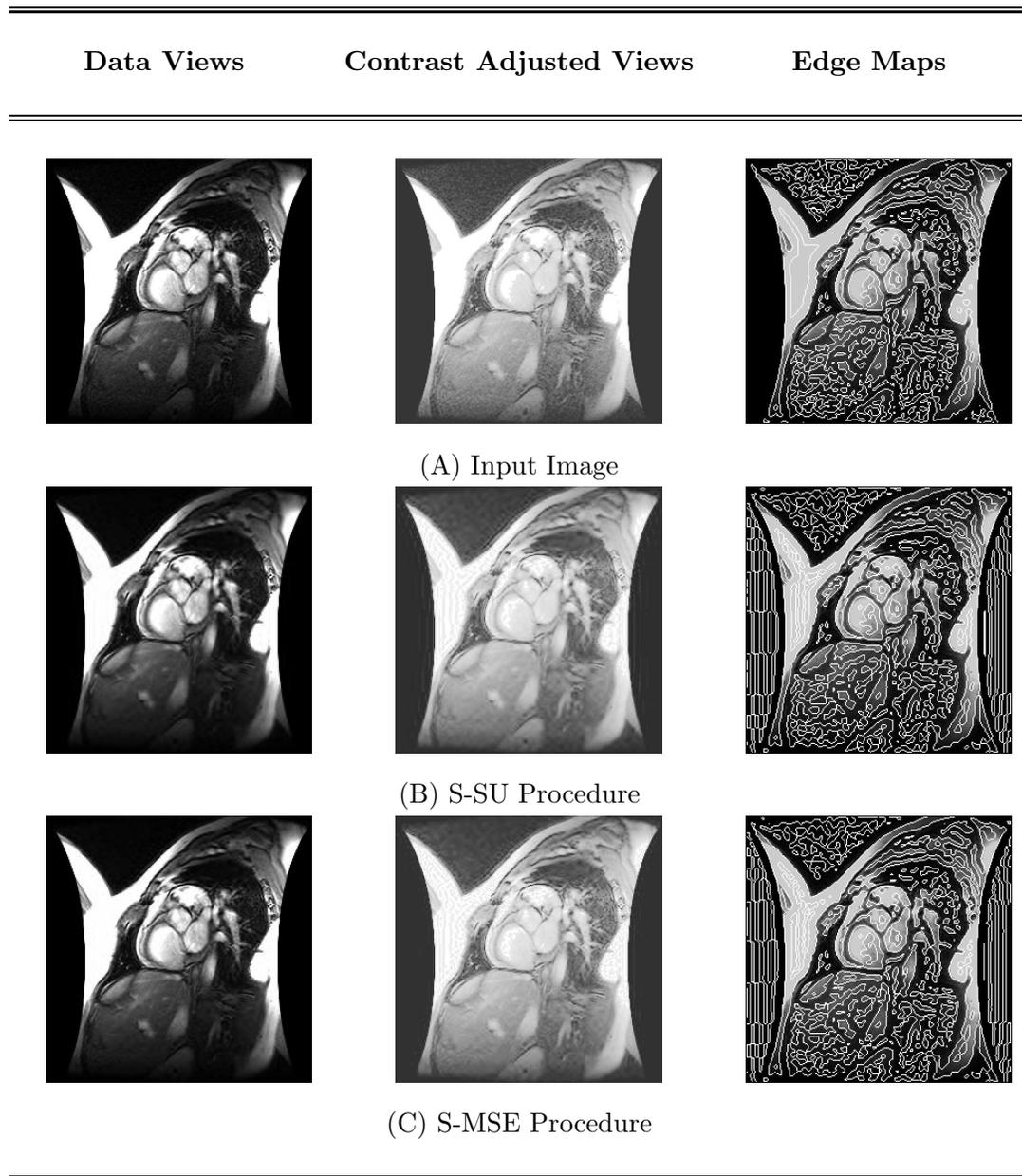


Figure H.2: Cardiac Image: SRS0000-10. S Group.

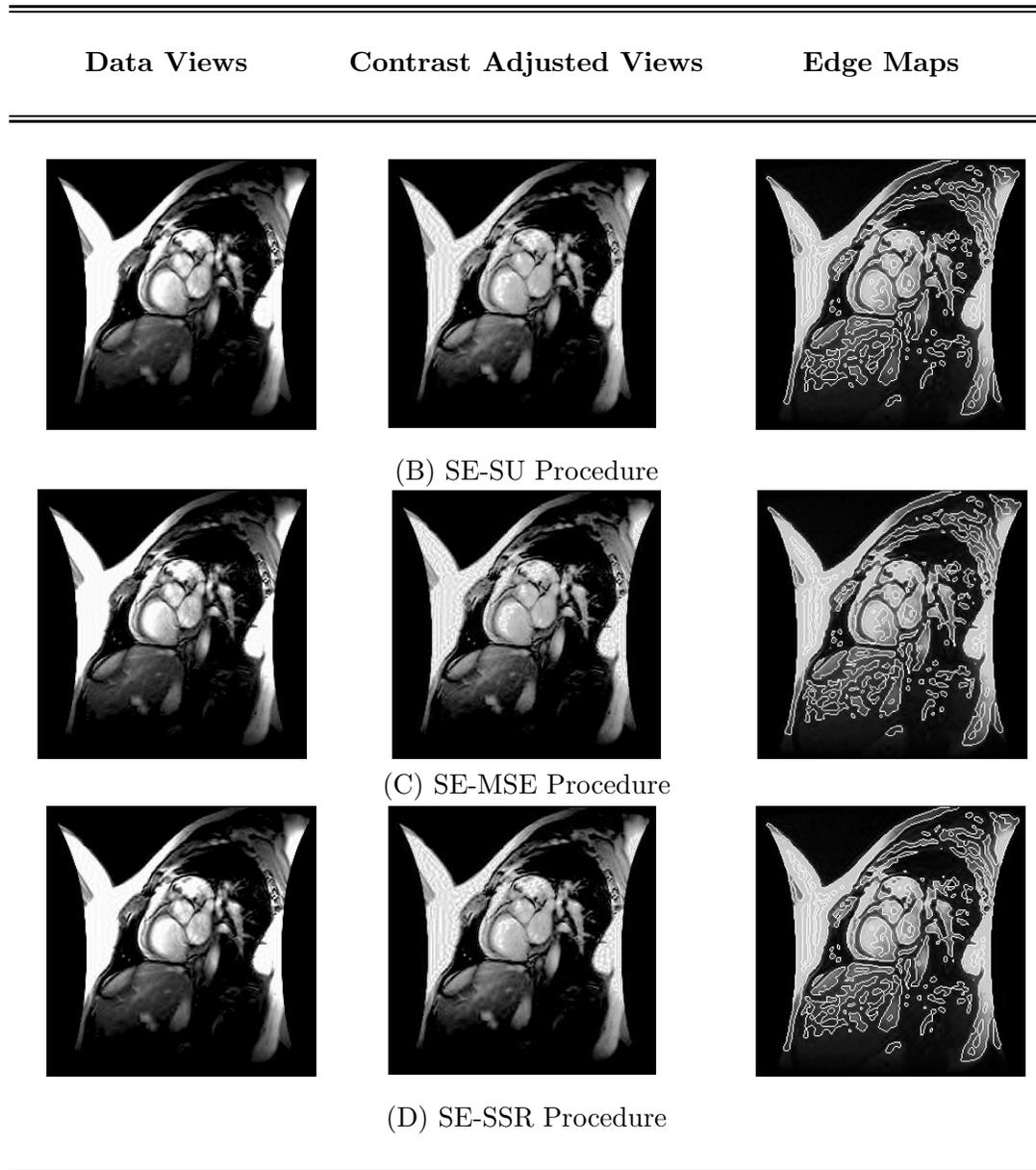


Figure H.3: Cardiac Image: SRS0000-10. SE Group

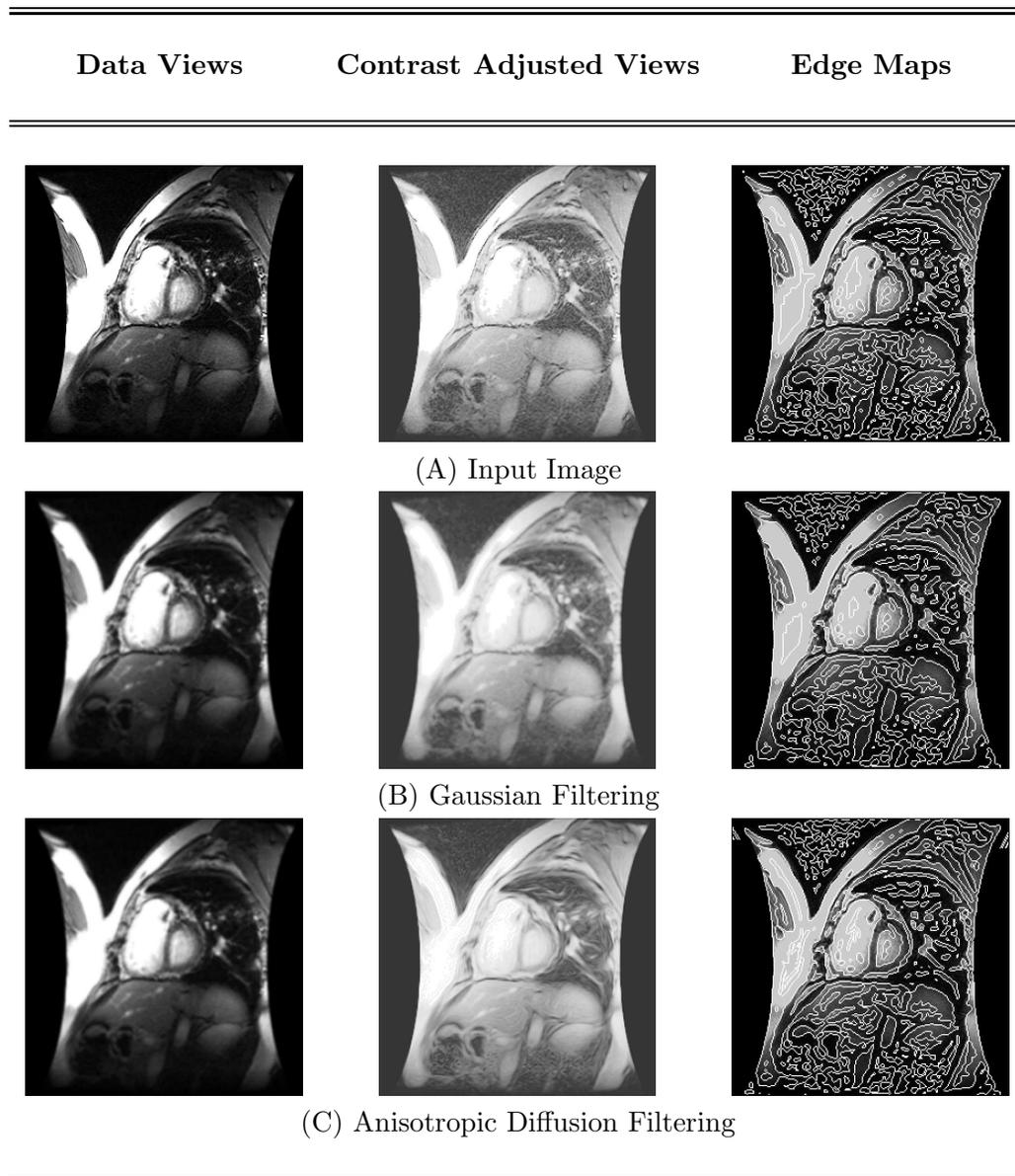


Figure H.4: Cardiac Image: SRS0000-70. Diffusion Group.

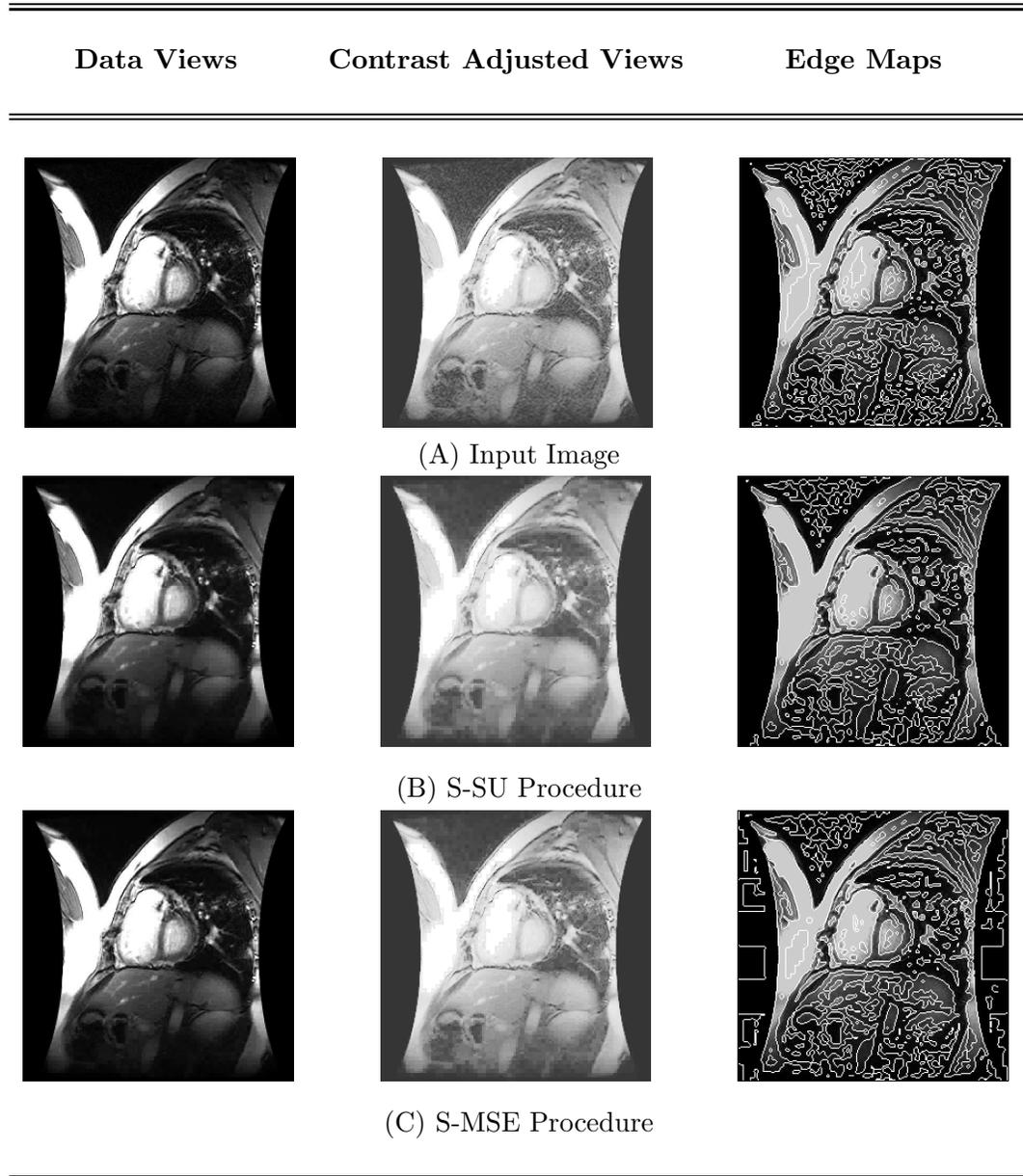


Figure H.5: Cardiac Image: SRS0000-70. S Group.

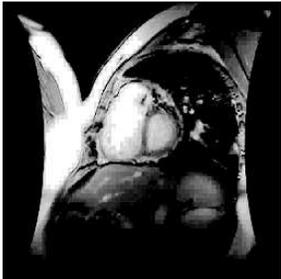
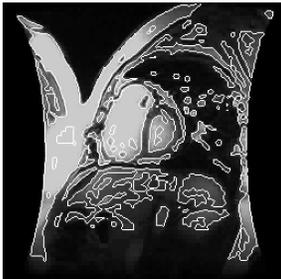
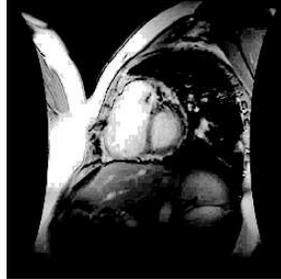
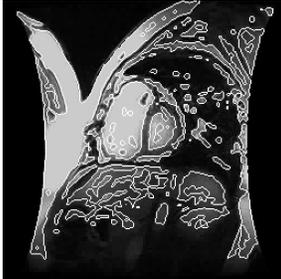
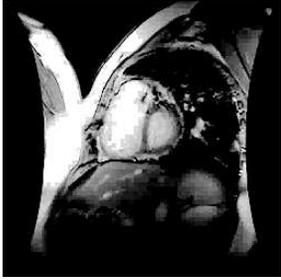
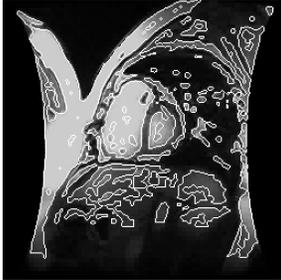
Data Views	Contrast Adjusted Views	Edge Maps
		
(A) SE-SU Procedure		
		
(B) SE-MSE Procedure		
		
(C) SE-SSR Procedure		

Figure H.6: Cardiac Image: SRS0000-70. SE Group.

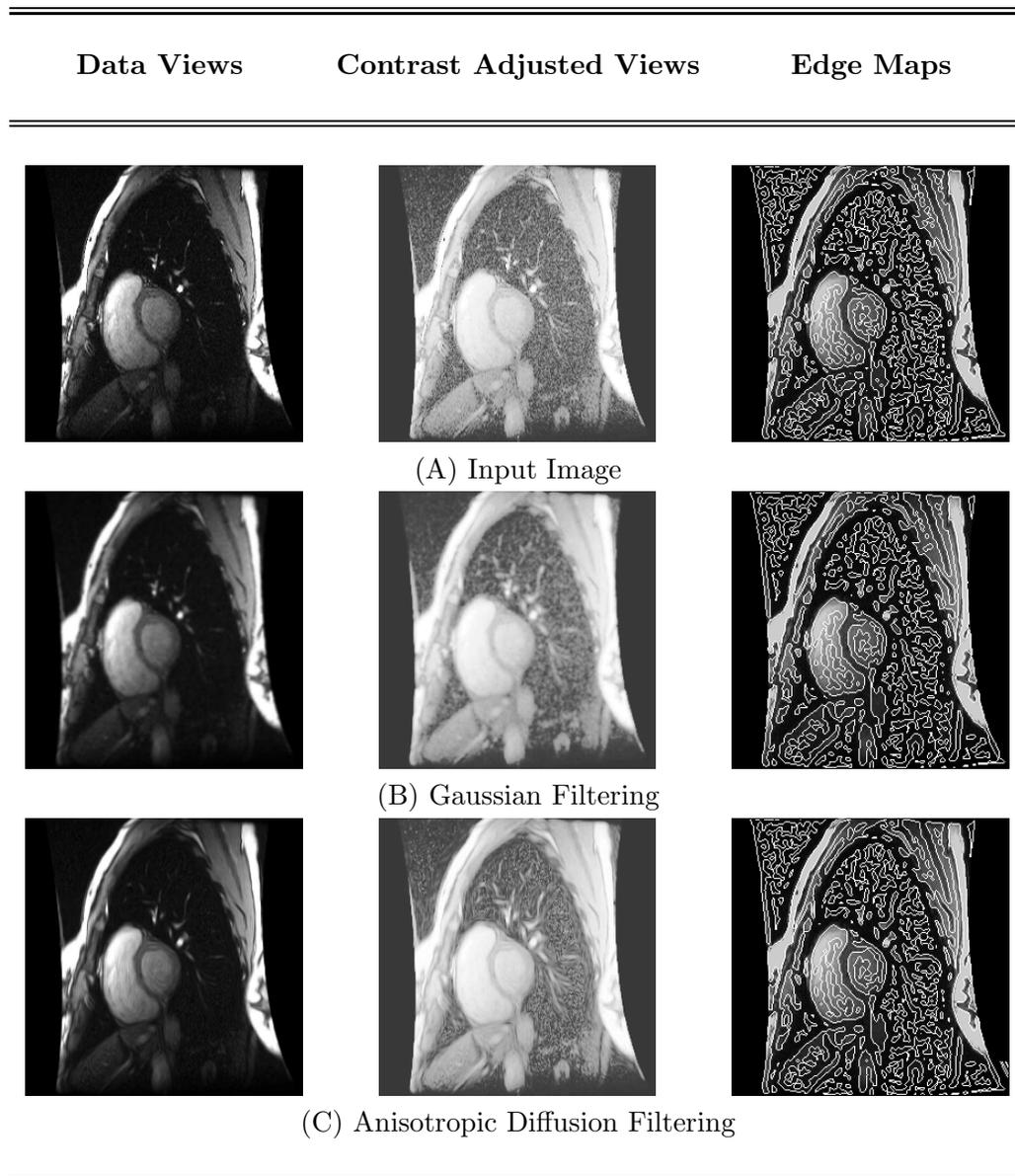


Figure H.7: Cardiac Image: Eman-70. Diffusion Group.

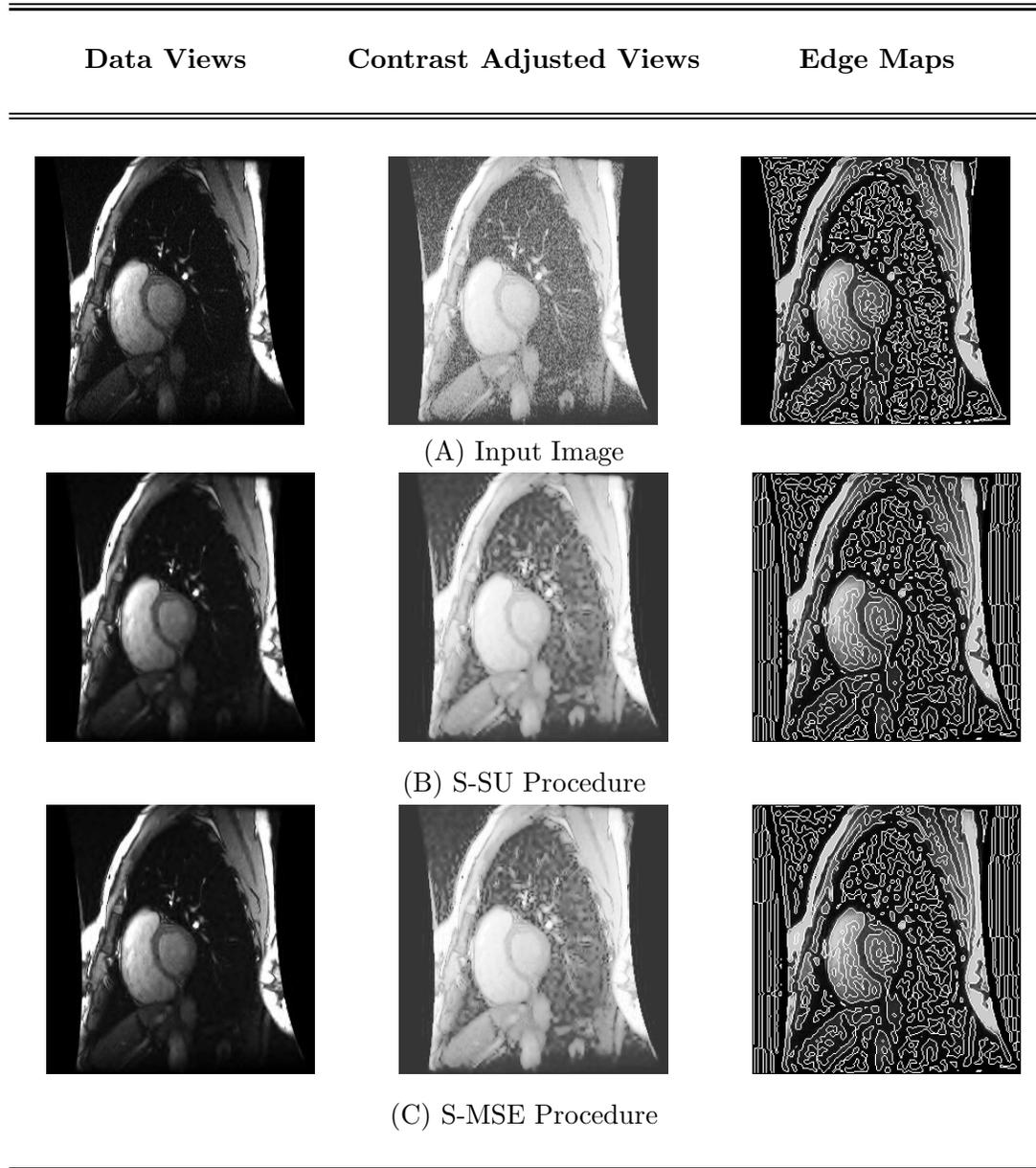


Figure H.8: Cardiac Image: Eman-70. S Group.

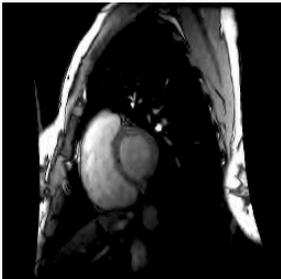
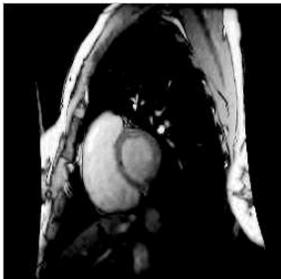
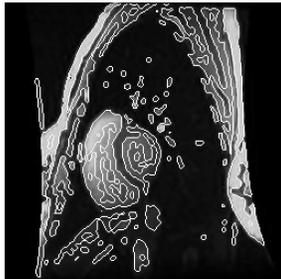
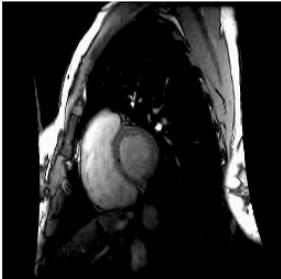
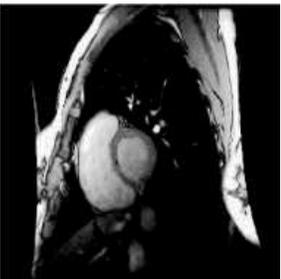
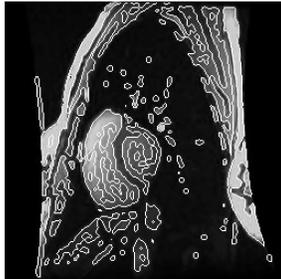
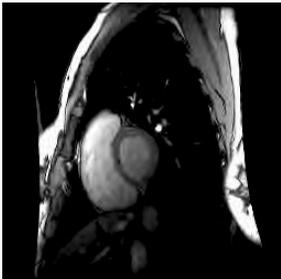
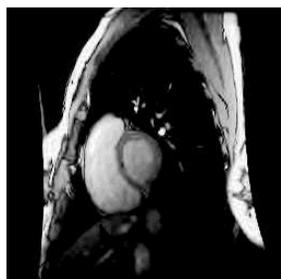
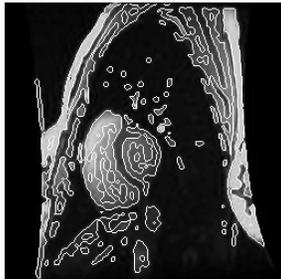
Data Views	Contrast Adjusted Views	Edge Maps
		
(A) SE-SU Procedure		
		
(B) SE-MSE Procedure		
		
(C) SE-SSR Procedure		

Figure H.9: Cardiac Image: Eman-70. SE Group.

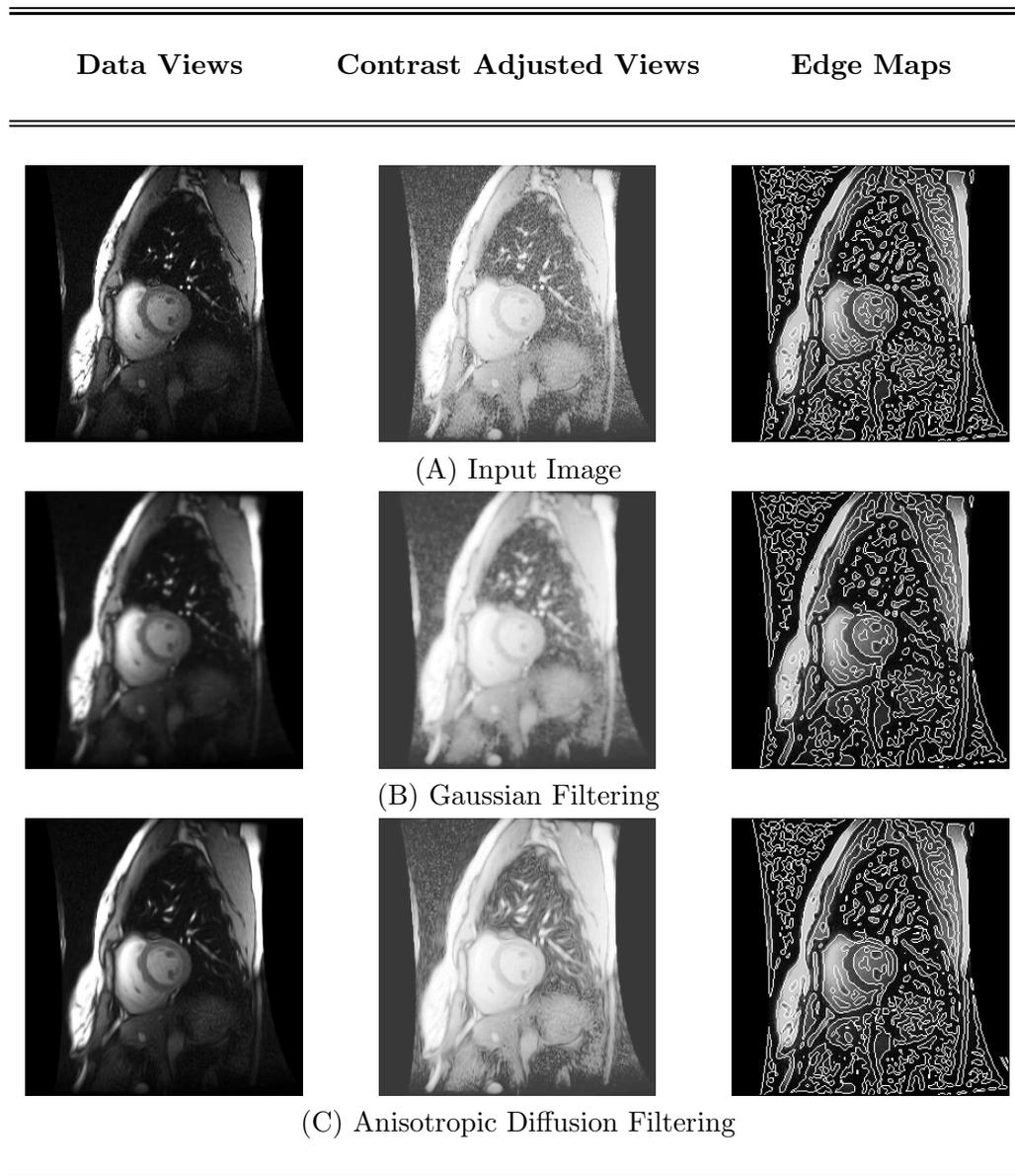


Figure H.10: Cardiac Image: Eman-83. Diffusion Group.

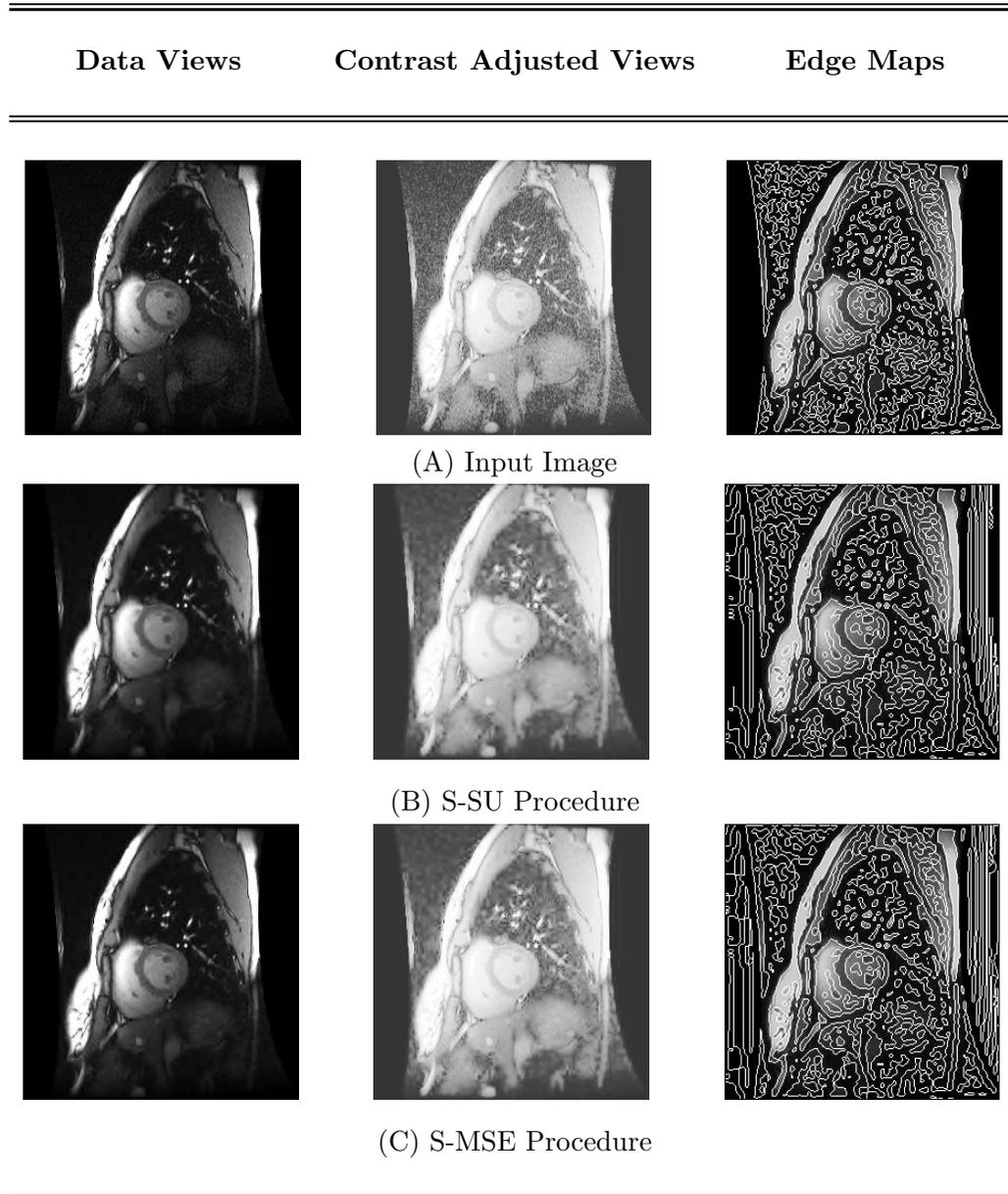
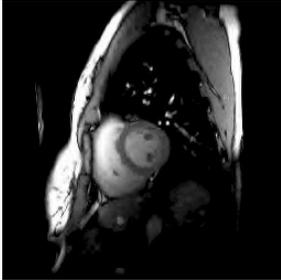
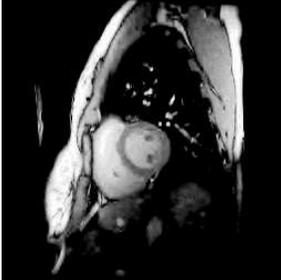
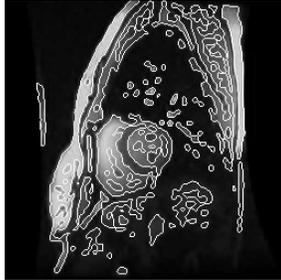
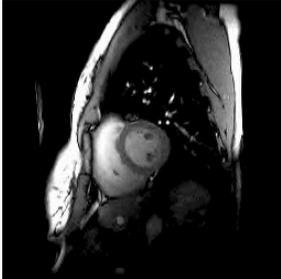
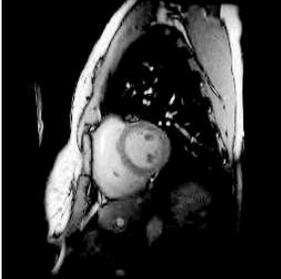
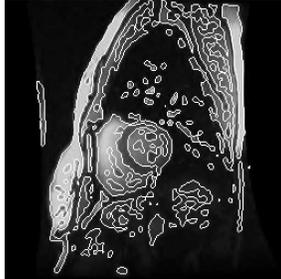
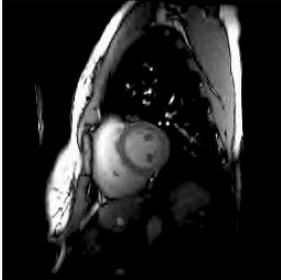
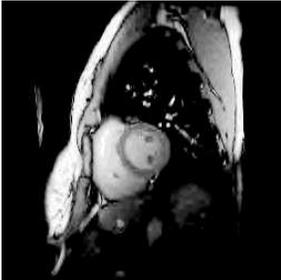
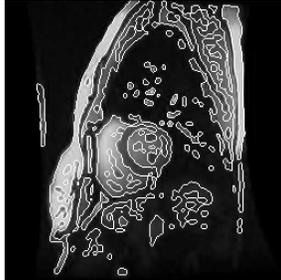


Figure H.11: Cardiac Image: Eman-83. S Group.

Data Views	Contrast Adjusted Views	Edge Maps
		
(A) SE-SU Procedure		
		
(C) SE-MSE Procedure		
		
(C) SE-SSR Procedure		

tbtail

Figure H.12: Cardiac Image: Eman-83. SE Group.

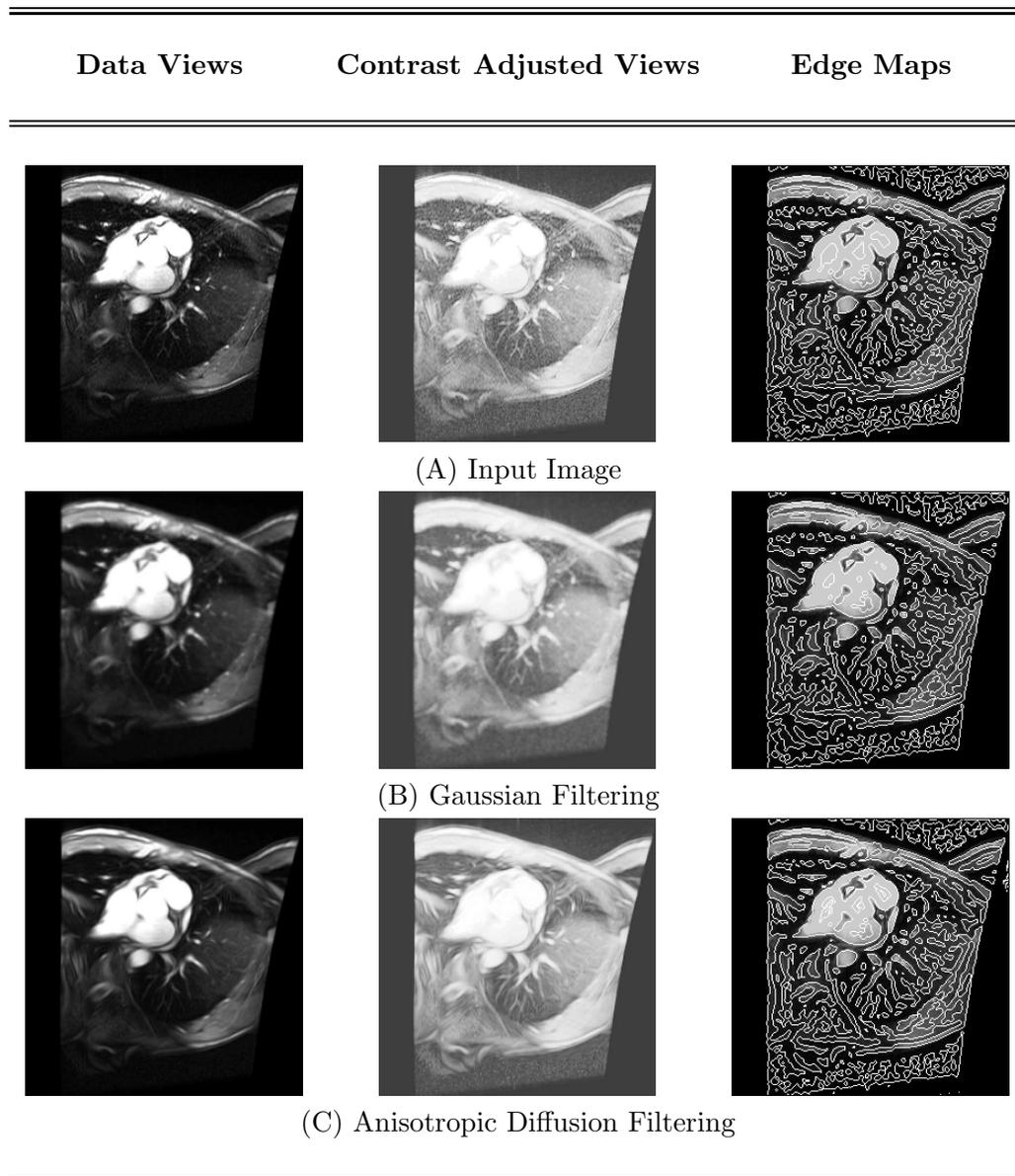


Figure H.13: Cardiac Image: SRS0006-24. Diffusion Group.

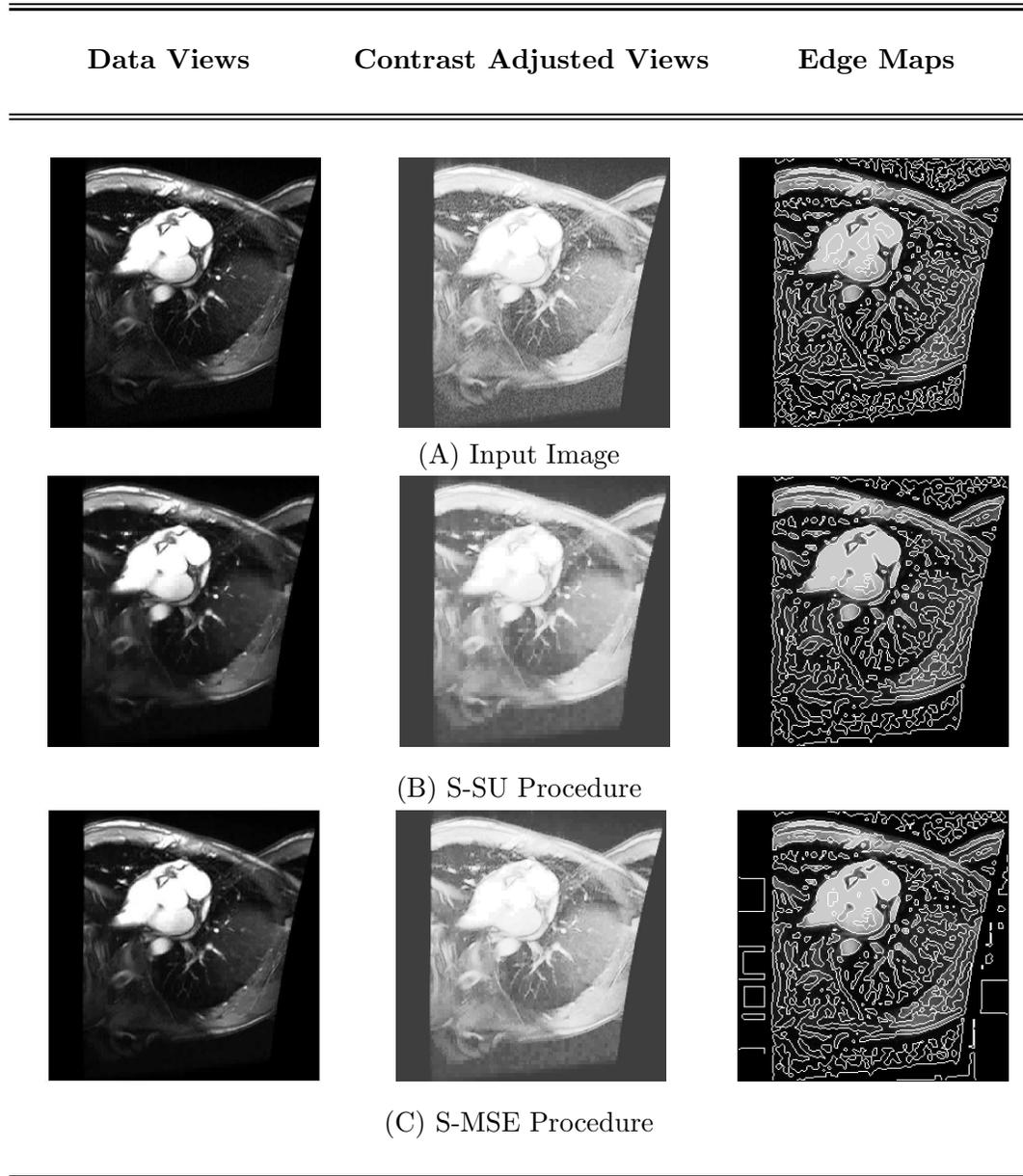


Figure H.14: Cardiac Image: SRS0006-24. S Group.

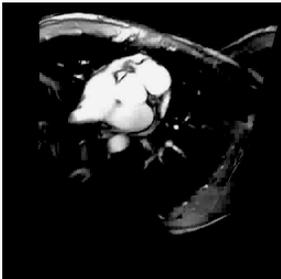
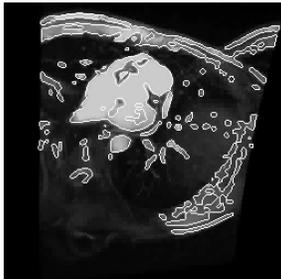
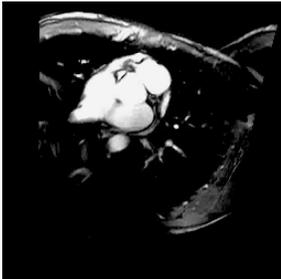
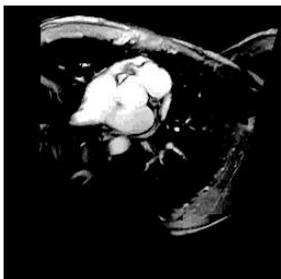
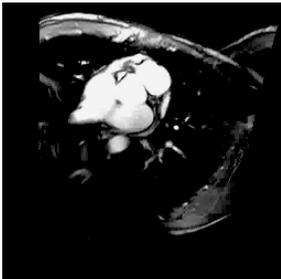
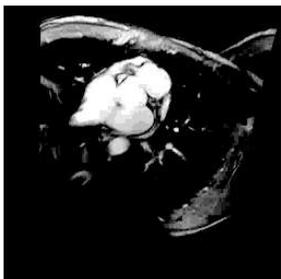
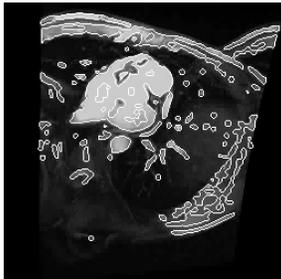
Data Views	Contrast Adjusted Views	Edge Maps
		
(A) SE-SU Procedure		
		
(B) SE-MSE Procedure		
		
(C) SE-SSR Procedure		

Figure H.15: Cardiac Image: SRS0006-24. SE Group

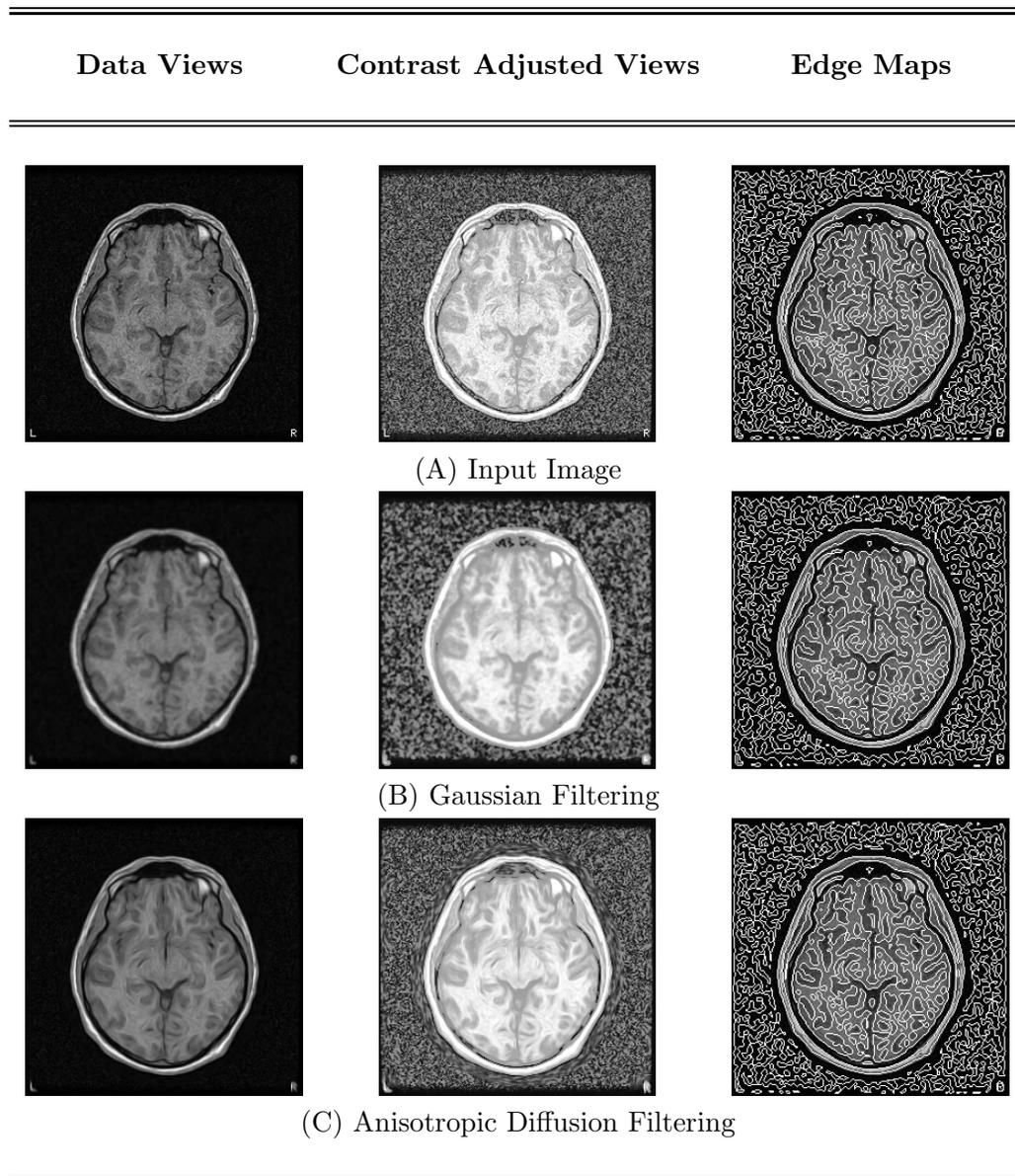


Figure H.16: Brain Image: BG-3. Diffusion Group.

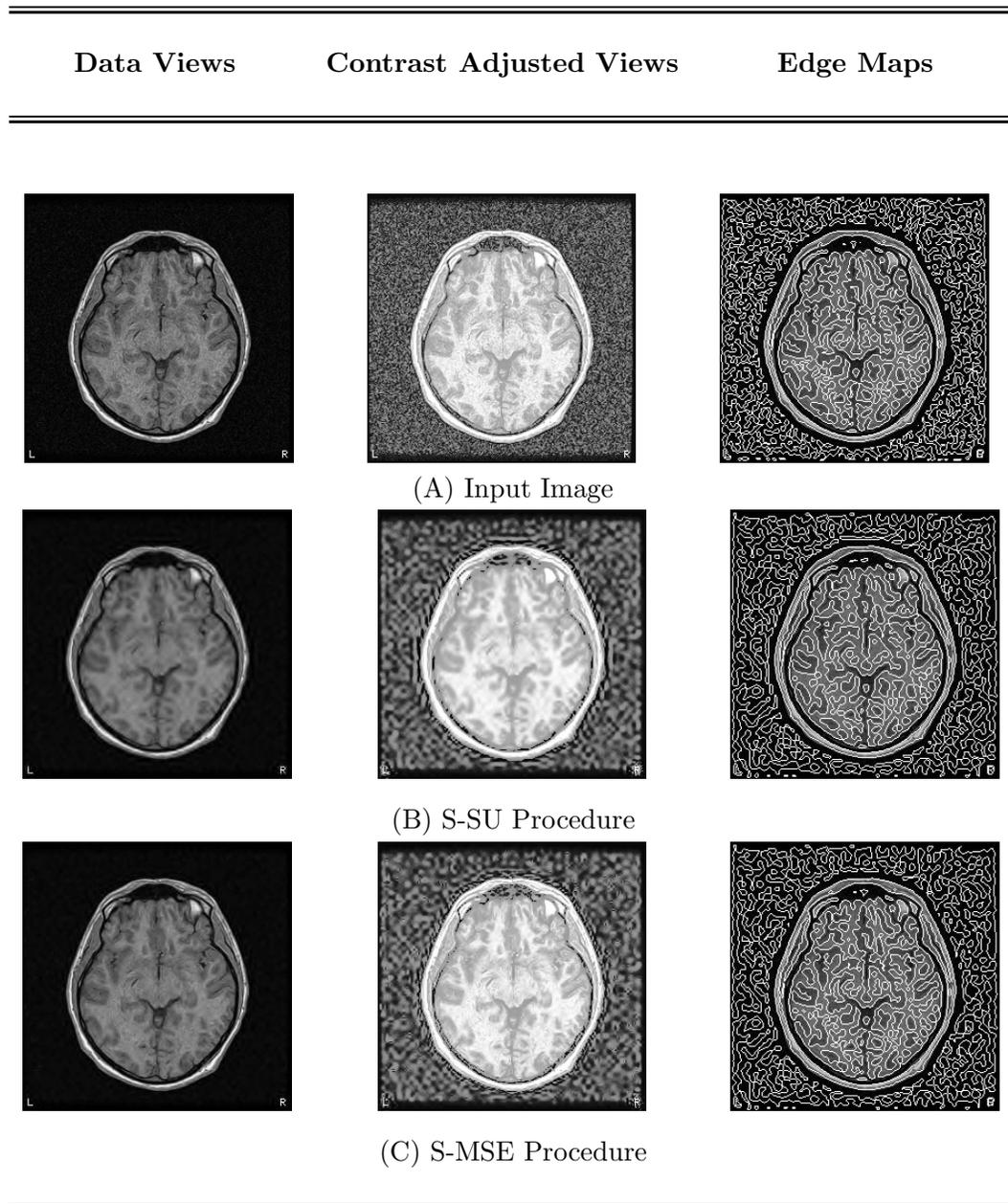


Figure H.17: Brain Image: BG-3. S Group.

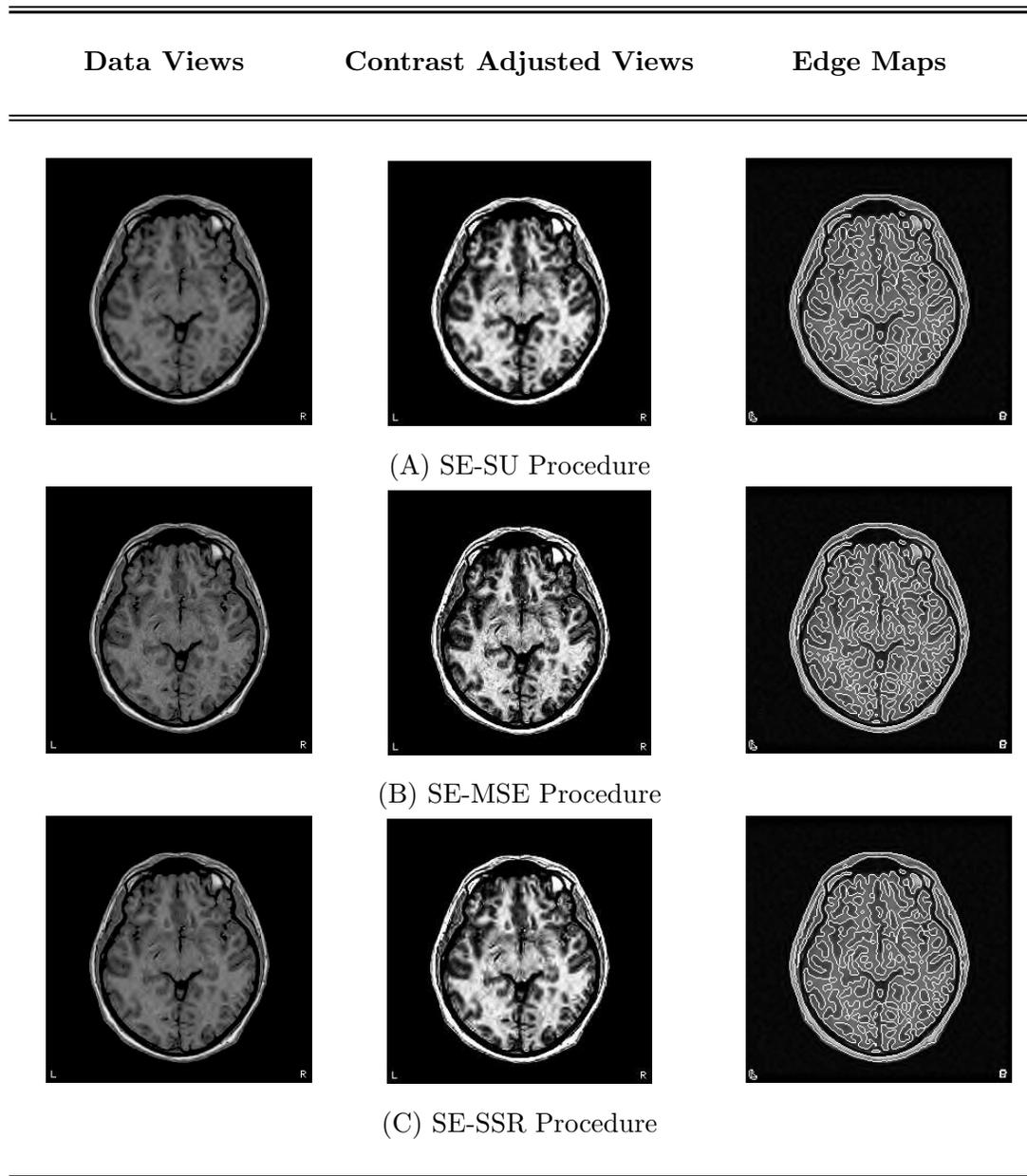


Figure H.18: Brain Image: BG-3. SE Group.

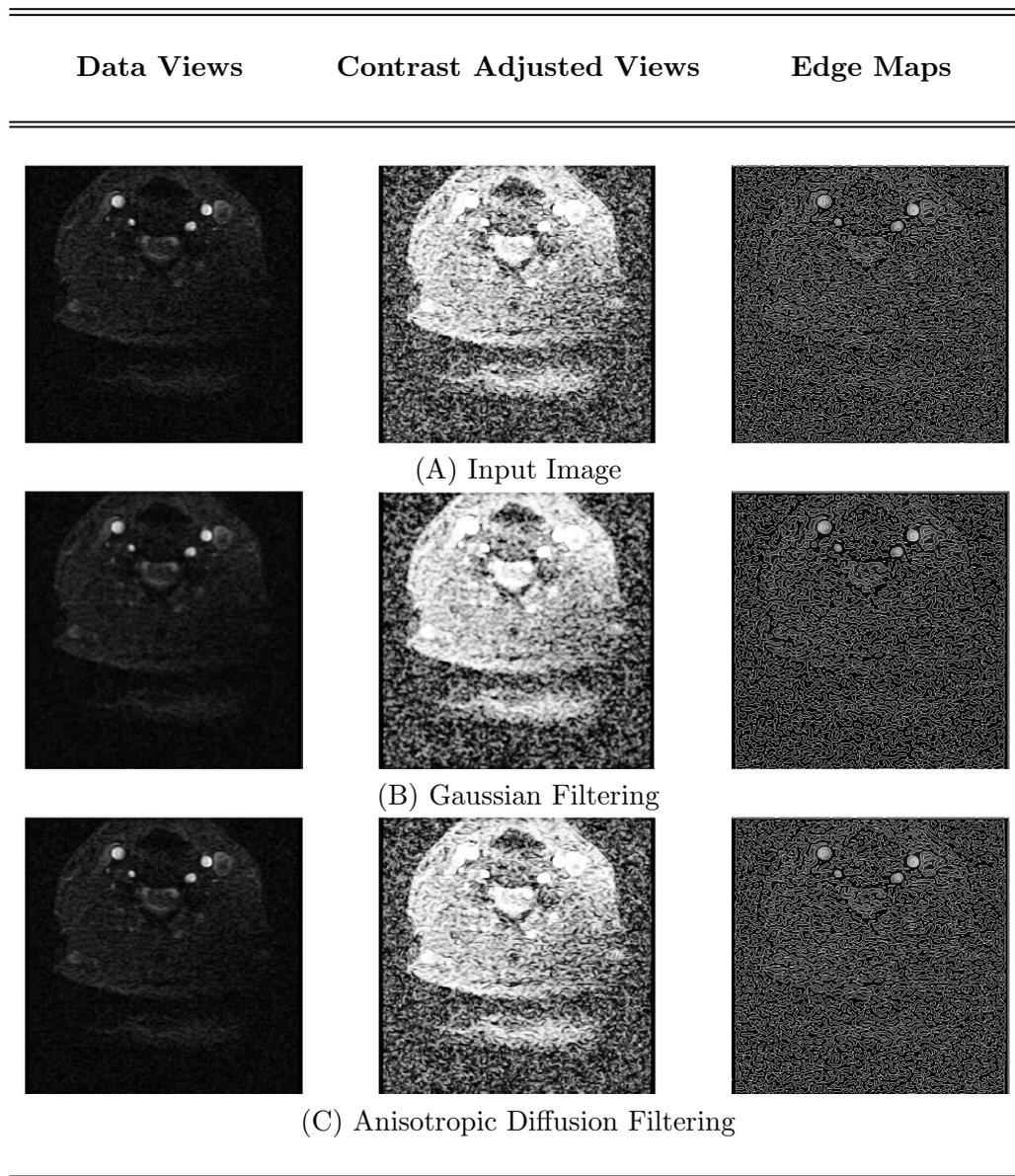


Figure H.19: Cardiovascular Image: Carotid Arteries 2. Diffusion Group.

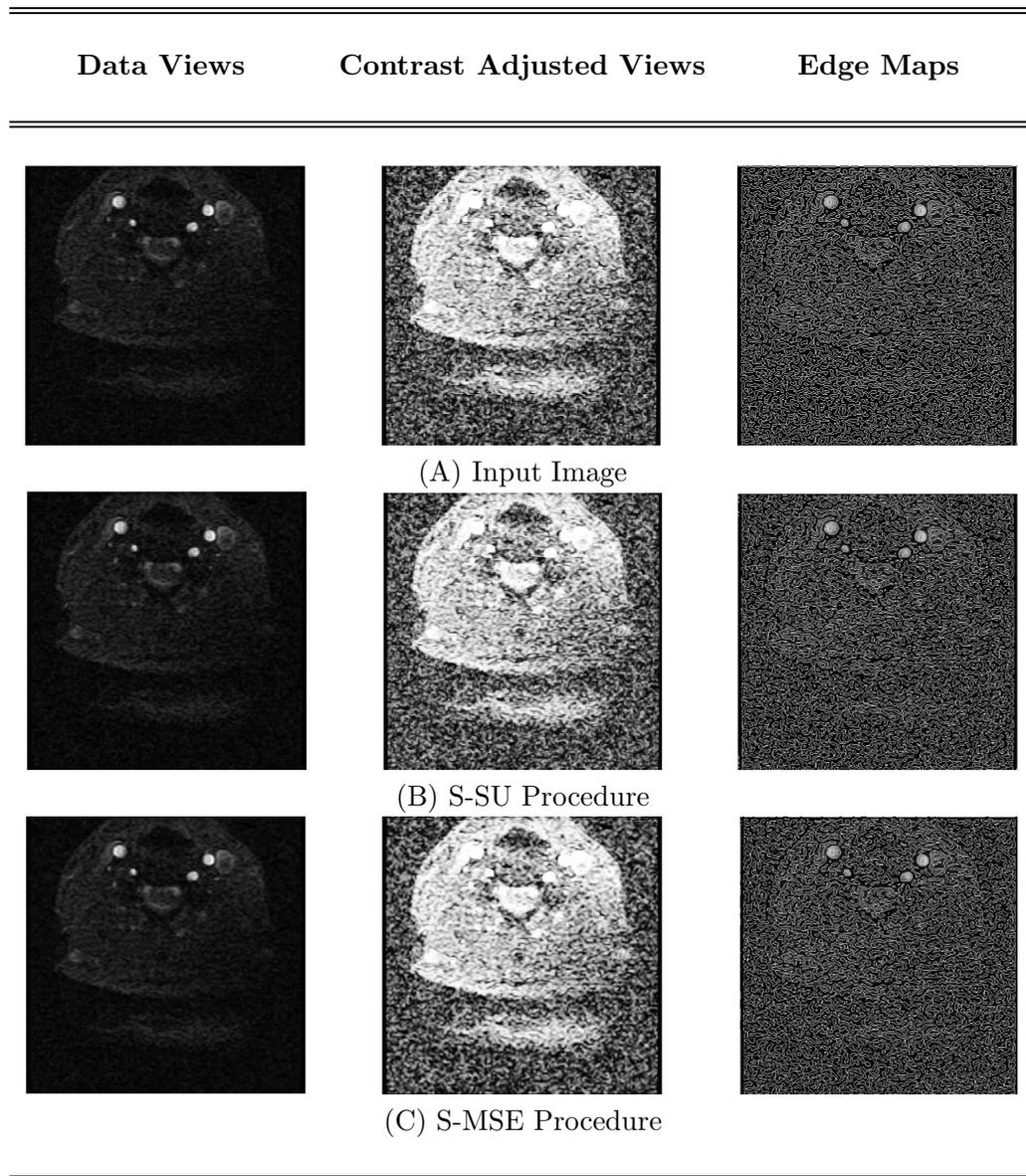
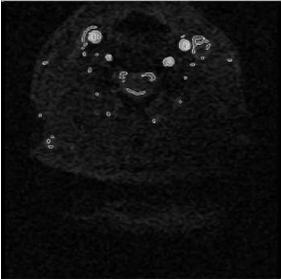
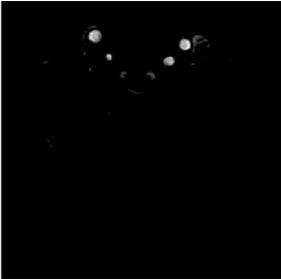
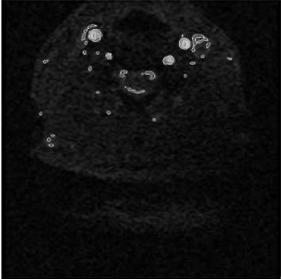
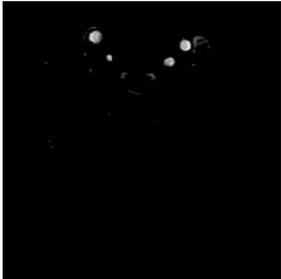
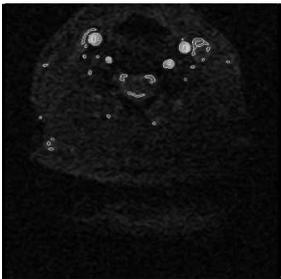


Figure H.20: Cardiovascular Image: Carotid Arteries 2. S Group.

Given Images	Edge Maps
	
(A) SE-SU Procedure	
	
(B) SE-MSE Procedure	
	
(C) SE-SSR Procedure	

‡ Contrast adjustment is not applicable to these images since most of the image features are removed by the SE schemes.

Figure H.21: Cardiovascular Image: Carotid Arteries 2. SE Group.

# Bibliography

- [1] F. Abramovich and Y. Benjamini. Thresholding of wavelet coefficients as multiple hypotheses testing procedure. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, Lecture Notes in Statistics 103, pages 5–14. Springer-Verlag, 1995.
- [2] M. Abramowitz and I. A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. United States. National Bureau of Standards. Applied mathematics series, 55. U.S. Govt. Print. Off., Washington, 1964.
- [3] P. S. Addison. *The illustrated wavelet transform handbook : introductory theory and applications in science, engineering, medicine and finance*. Institute of Physics Publishing, Bristol, 2002.
- [4] M. Alexander, R. Baumgartner, A. Summers, C. Windischberger, M. Klarhoefer, E. Moser, and R. Somorjai. A wavelet-based method for improving signal-to-noise ratio and contrast in MR images. *Magnetic Resonance Imaging*, 18(2):169–180, 2000.
- [5] American Heart Association. *Heart Disease and Stroke Statistics – 2004 Update*. American Heart Association, Dallas, Tex., 2003.
- [6] A. H. Andersen. On the Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 36:331–333, 1996.
- [7] L. C. Andrews and R. L. Phillips. *Mathematical techniques for engineers and scientists*. SPIE Press, Bellingham, Wash., 2003.
- [8] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies. Image coding using wavelet transform. *IEEE Transactions on Image Processing*, 1(2):205 – 220, 1992.
- [9] G. B. Arfken and H.-J. Weber. *Mathematical methods for physicists*. Academic Press, San Diego, fourth edition, 1995.
- [10] R. Averkamp. *Wavelet Thresholding for Non (Necessarily) Gaussian Noise*. PhD thesis, Institut fr Mathematische Stochastik, Universitt Freiburg, Dec. 1999.

- [11] P. Bao and L. Zhang. Noise reduction for magnetic resonance images via adaptive multiscale products thresholding. *IEEE Transactions on Medical Imaging*, 22(9):1089 – 1099, Sept 2003.
- [12] V. Barnett. *Comparative statistical inference*. John Wiley & Sons, New York, second edition, 1982.
- [13] E. J. Borowski and J. M. Borwein, editors. *HarperCollins Dictionary of Mathematics*. HarperCollins Publisher, New York, 1991.
- [14] M. J. Borran and R. D. Nowak. Wavelet-based denoising using hidden Markov models. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 6, Salt Lake City, Utah, May 2001.
- [15] C. S. Burrus, R. A. Gopinath, , and H. Guo. *Introduction to wavelets and wavelet transforms : a primer*. Prentice Hall, Upper Saddle River, N.J., 1998.
- [16] C. R. Carlson, E. H. Adelson, and C. H. Anderson. System for coring an image-representing signal. United States Patent 4,523,230, 1985.
- [17] G. Casella and R. L. Berger. *Statistical inference*. Duxbury Press, Belmont, California, 1990.
- [18] S. G. Chang, B. Yu, and M. Vetterli. Spatially adaptive wavelet thresholding with context modeling for image denoising. In *Proceedings of International Conference on Image Processing (ICIP 98)*, volume 1, pages 535 – 539, Oct. 1998.
- [19] S. G. Chang, B. Yu, and M. Vetterli. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing*, 9(9):1532 – 1546, Sep 2000.
- [20] S. G. Chang, B. Yu, and M. Vetterli. Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Transactions on Image Processing*, 9(9):1522–1531, Sep 2000.
- [21] A. Charmbolle, R. A. D. Vore, N.-Y. Lee, and B. J. Lucier. Non-linear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage. *IEEE Transactions on Image Processing*, 7(3):319–335, March 1998.
- [22] F. Chau, Y. Liang, J. Gao, and X. Shao. *Chemometrics : from basics to wavelet transform*. John Wiley & Sons, Hoboken, N.J., 2004.
- [23] H. A. Chipman, E. D. Kolaczyk, and R. E. McCulloch. Adaptive Bayesian wavelet shrinkage. *Journal of the American Statistical Association*, 92:1413–1421, 1997.

- [24] H. Choi and R. G. Baraniuk. Wavelet statistical models and besov spaces. In *Proceedings of SPIE on Wavelet Applications in Signal Processing VII*, volume 3813, pages 489–501, Oct. 1999.
- [25] C. K. Chu. *An introduction to wavelets*. Academic Press, Boston, 1992.
- [26] M. Clyde and E. I. George. Empirical Bayes estimation in wavelet nonparametric regression. In Muller and Vidakovic [80], pages 309–322.
- [27] M. Clyde and E. I. George. Flexible empirical bayes estimation for wavelets. *Journal of the Royal Statistical Society, Series B*, 62:681–698, 2000.
- [28] R. R. Coifman, Y. Meyer, and V. Wickerhauser. Wavelet analysis and signal processing. In L. Auslander, T. Kailath, and S. K. Mitter, editors, *Signal Processing, Part I: Signal Processing Theory*, pages 59–68. Springer-Verlag, New York, NY, 1990.
- [29] M. S. Crouse, R. G. Baraniuk, and R. D. Nowak. Signal estimation using wavelet-markov models. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97*, volume 5, pages 3429 – 3432, April 1997.
- [30] M. S. Crouse, R. D. Nowak, and R. G. Baraniuk. Wavelet-based statistical signal processing using hidden markov models. *IEEE Transactions on Signal Processing*, 46(4):886 – 902, April 1998.
- [31] I. Daubechies. *Ten lectures on wavelets*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1992.
- [32] K. R. Davidson and A. P. Donsig. *Real analysis with real applications*. Prentice Hall, Upper Saddle River, NJ, 2002.
- [33] A. Davison. *Statistical Models*. Cambridge University Press, New York, 2003.
- [34] R. A. DeVore, B. Jawerth, and B. J. Lucier. Image compression through wavelet transform coding. *IEEE transactions on information theory*, 38(2):719–746, March 1992.
- [35] R. A. DeVore and B. J. Lucier. Fast wavelet techniques for near-optimal image processing. In *IEEE Military Communications Conference Record*, pages 1129–1135, San Diego, 1992.
- [36] D. L. Donoho. Wavelet shrinkage and W. V. D.: A 10-minute tour. *International Conference on Wavelets and Applications*, 1992.

- [37] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455, 1994.
- [38] D. L. Donoho and I. M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90(432):1200–1224, Dec. 1995.
- [39] D. L. Donoho, I. M. Johnstone, G. Kerkyacharian, and D. Picard. Wavelet shrinkage: Asymptopia? *Journal of the Royal Statistics Society Series B.*, 57(2):301–337, 1995.
- [40] S. Efromovich. *Nonparametric curve estimation : methods, theory and applications*. Springer, New York, 1999.
- [41] M. Evans, N. Hastings, and B. Peacock. *Statistical distributions*. J. Wiley, New York, 2nd. edition, 1993.
- [42] G. Fan. *Wavelet-Domain Statistical Image Modeling and Processing*. Ph.d. thesis, University of Delaware, 2001.
- [43] G. Fan and X.-G. Xia. Image denoising using local contextual hidden Markov model in the wavelet domain. *IEEE Signal Processing Letter*, 8(5):125–128, May 2001.
- [44] M. A. T. Figueiredo and R. D. Nowak. Wavelet-based image estimation: an empirical Bayes approach using Jeffrey’s noninformative prior. *IEEE Transactions on Image Processing*, 10(9):1322 –1331, Sept. 2001.
- [45] H.-Y. Gao. Wavelet shrinkage denoising using the non-negative garrote. *Journal of Computational and Graphical Statistics*, 7:469–488, 1998.
- [46] H.-Y. Gao and A. G. Bruce. Waveshrink with firm shrinkage. *Statistica Sinica*, 7:855–874, 1997.
- [47] J. Ge and G. Mirchandani. A simple and efficient wavelet-based denoising algorithm using inter- and intrascale statistics adaptively. In *Seventh International Symposium on Signal Processing and its Applications. ISSPA 2003 1-4 July 2003, Paris, FRANCE*, 2003.
- [48] J. Ge and G. Mirchandani. Softening the multiscale product method for adaptive noise reduction. In *Thirty-Seventh Annual Asilomar Conference on Signals, Systems and Computers*, November 2003.
- [49] J. W. Goodman. *Statistical optics*. Wiley interscience, New York, 1985.

- [50] A. Graps. An introduction to wavelets. *IEEE Computational Sciences and Engineering*, 2(2):50–61, 1995.
- [51] R. L. Gregg and R. D. Nowak. Noise removal methods for high resolution MRI. In *IEEE Nuclear Science Symposium*, volume 2, pages 1117 – 1121, 1997.
- [52] H. Gudbjartsson and S. Patz. The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine*, 34:910–914, 1995.
- [53] N. B. Haaser and J. A. Sullivan. *real analysis*. Dove Publications, New York, 1991.
- [54] W. Hardle, G. Kerkycharian, D. Picard, and A. Tsybakov. *Wavelets, approximation, and statistical applications*. Lecture notes in statistics ; 129. Springer, New York, 1998.
- [55] Heart and Stroke Foundaton of Canada. *The changing face of heart disease and stroke in Canada 2000*. Statistics Canada, Ottawa, Canada, Oct 1999.
- [56] P. J. Huber. *Robust statistics*. John Wiley & Sons, New York, 1981.
- [57] W.-L. Hwang and S. Mallat. Singularities and noise discrimination with wavelets. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1992. ICASSP-92*, volume 4, pages 377–380, 1992.
- [58] M. Jansen. *Noise reduction by wavelet thresholding*. Lecture notes in statistics volume 161. Springer, New York, 2001.
- [59] M. Jansen, M. Malfait, and A. Bultheel. Generalized cross validation for wavelet thresholding. *Signal Processing*, 56(1):33 – 44, 1997.
- [60] A. Jeffrey. *Handbook of mathematical formulas and integrals*. Academic Press, San Diego, 2000.
- [61] H. Jeffreys and B. S. Jeffreys. *Methods of mathematical physics*. Cambridge University Press, New York, 3rd edition, 1999.
- [62] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions*, volume 2. Wiley & Sons, New York, 2 edition, 1994.
- [63] I. Jorhnstone. Wavelets and the theory of nonparametric function estimation. *Phil. Trans. Roy. Soc. Lond. A.*, 357:2475 – 2494, 1999.
- [64] J. Kalbfleisch. *Probability and statistical inference*. Springer-Verlag, New York, 2nd ed. edition, 1985.

- [65] O. T. Karlsen and W. M. B. R. Verhagen. Parameter estimation from Rician-distributed data sets using a maximum likelihood estimator: Application to t1 and perfusion measurements. *Magnetic Resonance in Medicine*, 41(3):614 – 623, 1999.
- [66] S. M. Kay. *Fundamentals of statistical signal processing*, volume II: Detection Theory. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1998.
- [67] M. G. Kendall, A. Stuart, and J. K. Ord. *Kendall's Advanced Theory of Statistics*, volume 1: Distribution Theory. Halsted Press, New York, sixth edition, 1994.
- [68] S. J. Kisner, T. M. Talavage, and J. Ulmer. Testing a model for MR imager noise. In *The 2nd Joint Meeting of the IEEE Engineering in Medicine & Biology Society and the Biomedical Engineering Society*, pages 23–26, Houston, Texas, 2002.
- [69] A. Korostelev and A. Tsybakov. *Minimax Theory of Image Reconstruction*. Springer-Verlag, New York, 1993.
- [70] H. Krim, D. Tucker, S. Mallat, and D. Donoho. On denoising and best signal representation. *IEEE Transactions on Information Theory*, 45(7):2225–2238, Nov 1999.
- [71] D. Leporini and J.-C. Pesquet. Bayesian wavelet denoising using Besov priors. In *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, 1999*, pages 144–147, 1999.
- [72] D. Leporini and J.-C. Pesquet. Bayesian wavelet denoising: Besov priors and non-Gaussian noises. *Signal Processing*, 81:55–67, 2001.
- [73] A. Macovski. Noise in MRI. *Magnetic Resonance in Medicine*, 36:1996, 494-497.
- [74] S. G. Mallat. *A wavelet tour of signal processing*. Academic Press, San Diego, Calif., 2nd edition, 1999.
- [75] S. G. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 32(2):617 – 643, 1992.
- [76] S. G. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(7):710–732, 1992.
- [77] Y. Meyer. *Wavelets : algorithms and applications*. Society for Industrial & Applied Mathematics, Philadelphia, 1993.

- [78] J. Minkoff. *Signals, noise, and active sensors : radar, sonar, laser radar*. Wiley-Interscience, New York, 1992.
- [79] J. Minkoff. *Signal processing fundamentals and applications for communications and sensing systems*. Artech House, Boston, 2002.
- [80] P. Muller and B. Vidakovic, editors. *Bayesian Inference in Wavelet Based Models*. Lecture Notes in Statistics vol 141. Springer-Verlag, New York, 1999.
- [81] G. P. Nason. Wavelet shrinkage using cross-validation. *Journal of Royal Statistical Society Series B*, 58:463–479, 1996.
- [82] A. F. Nikiforov and V. B. Uvarov. *Special functions of mathematical physics : a unified introduction with applications*. Birkhuser, Boston, 1988.
- [83] R. D. Nowak. Wavelet-based rician noise removal for magnetic resonance imaging. *IEEE Transactions on Image Processing*, 8(10):1408 – 1419, Oct. 1999.
- [84] R. T. Ogden. *Essential wavelets for statistical applications and data analysis*. Birkhuser, Boston, 1997.
- [85] R. T. Ogden and E. Parzen. Change-point approach to data analytic wavelet thresholding. *Statistics and Computing*, 6, 1995.
- [86] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, third edition, 1991.
- [87] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. PAMI*, 12(7):629 – 639, 1990.
- [88] A. Pizurica. *Image Denoising Using Wavelets and Spatial Context Modeling*. Ph.d thesis, Ghent University, Belgium, June 2002.
- [89] A. Pizurica and W. Philips. Multiscale statistical image models and Bayesian methods. In F. Truchetet, editor, *Proceedings of SPIE Volume 5266: Wavelet Applications in Industrial Processing*, pages 60–74, February 2004.
- [90] A. Pizurica, W. Philips, and M. Acheroy. A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising. *IEEE Transactions on Image Processing*, 11(5):545–557, May 2002.
- [91] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy. The application of Markov random field models to wavelet-based image denoising. In J. Blanc-Talon and D. C. Popescu, editors, *Imaging and Vision*

*Systems: Theory, Assessment and Applications*,, pages 43–70. NOVA Science Books, Huntington, USA, 2001.

- [92] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy. A versatile wavelet domain noise filtration technique for medical imaging. *IEEE Transactions on Medical Imaging*, 22(3):323–331, March 2003.
- [93] J. Portilla, V. Strela, M. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12(11):1338–1351, November 2003.
- [94] W. K. Pratt. *Digital image processing : PIKS inside*. Wiley, New York, 3rd edition, 2001.
- [95] S. Qian. *Introduction to time-frequency and wavelet transforms*. Prentice Hall PTR, Upper Saddle River, N.J., 2002.
- [96] S. Ray, A. Chan, and B. Mallick. Bayesian wavelet shrinkage in transformation based normal models. In *IEEE International Conference on Image Processing*, volume 1, pages 876 – 879, 2002.
- [97] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE Signal Processing Magazine*, 8(4):14–38, Oct. 1991.
- [98] J. Romberg, H. Choi, and R. G. Baraniuk. Bayesian wavelet domain image modeling using hidden markov models. *IEEE Transactions on Image Processing*, 10(7):1056–1068, July 2001.
- [99] J. Sijbers. *Signal and noise estimation from Magnetic Resonance Images*. PhD thesis, University of Antwerp, Antwerp, Belgium, 1998.
- [100] J. Sijbers, A. den Dekker, M. Verhoye, A. V. der Linden, and D. V. Dyck. Adaptive anisotropic noise filtering for magnitude MR data. In *Proceedings of SPIE Medical Imaging*, pages 1418–1425, San Diego, February 1999.
- [101] J. Sijbers, A. J. den Dekker, P. Scheunders, and D. V. Dyck. Maximum-likelihood estimation of Rician distribution parameters. *IEEE Transactions on Medical Imaging*, 17(3):357 – 361, 1998.
- [102] E. P. Simoncelli. Bayesian denoising of visual images in the wavelet domain. In Muller and Vidakovic [80], chapter 18, pages 292–308.
- [103] E. P. Simoncelli and E. H. Adelson. Noise removal via Bayesian wavelet coring. In *Proceedings of the 3rd International Conference on Image Processing Lausanne, Switzerland*, September 1996.

- [104] J.-L. Starck, F. Murtagh, and A. Bijaoui. *Image processing and data analysis : the multiscale approach*. Cambridge University Press, New York, 1998.
- [105] G. Strange and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, Wellesley, M.A., 2nd edition, 1997.
- [106] A. Stuart, J. K. Ord, and S. Arnold, editors. *Kendall's advanced theory of statistics*, volume 2A: Classical inference and the linear model. Oxford University Press, Inc., New York, sixth edition, 1994.
- [107] C. Torrence and G. P. Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, 79(1):61–78, 1998.
- [108] E. Trucco and A. Verri. *Introductory techniques for 3-D computer vision*. Prentice Hall, Upper Saddle River, N.J, 1998.
- [109] M. Vetterli and J. Kovacevic. *Wavelets and subband coding*. Prentice Hall PTR, Englewood Cliffs, N.J., 1995.
- [110] B. Vidakovic. Nonlinear wavelet shrinkage with Bayes rules and Bayes factors. *Journal of the American Statistical Association*, 93(441):173–179, 1998.
- [111] B. Vidakovic. *Statistical modeling by wavelets*. J. Wiley, New York, 1999.
- [112] D. F. Walnut. *An introduction to wavelet analysis*. Birkhäuser, Boston, 2002.
- [113] G. G. Walter and X. Shen. *Wavelets and other orthogonal systems*. Chapman & Hall/CRC, Boca Raton, 2nd edition, 2001.
- [114] Y. Wang. *MR imaging statistics and model based MR image analysis*. PhD thesis, University of Maryland, Baltimore, 1995.
- [115] Y. Wang and T. Lei. Statistical analysis of MR imaging and its applications in image modeling. In *Proceedings of the IEEE International Conference on Image Processing and Neural Networks*, volume 1, pages 866–870, 1994.
- [116] A. Watt and F. Policarpo. *The Computer Image*. Addison Wesley, Harlow, Essex, 1998.
- [117] J. B. Weaver, Y. Xu, D. H. Jr., and L. Cromwell. Filtering noise from images with wavelet transforms. *Magnetic Resonance in Medicine*, 21, 1991.

- [118] J. Weickert. *Anisotropic Diffusion in Image Processing*. Teubner-Verlag, Stuttgart, 1998.
- [119] J. Weickert. Coherence-enhancing diffusion filtering. *International Journal of Computer Vision*, 31(2-3):111 – 127, April 1999.
- [120] J. Weickert. Nonlinear diffusion filtering. In B. Jahne, H. Haubecker, and P. Geinler, editors, *Handbook on Computer Vision and Applications, Vol. 2: Signal Processing and Pattern Recognition*, pages 423–450. Academic Press, San Diego, 1999.
- [121] D. Weishaupt, V. D. Koechli, and B. Marincek. *How does MRI work? : an introduction to the physics and function of magnetic resonance imaging*. Springer, New York, 2003.
- [122] N. Weyrich and G. T. Warhola. Wavelet shrinkage and generalized cross validation for image denoising. *IEEE Transactions on Image Processing*, 7(1):82 – 90, Jan. 1998.
- [123] J. Whitaker and B. K. Benson, editors. *Standard Handbook of Video and Television Engineering*. McGraw-Hill, New York, fourth edition, 2003.  
(With updated chapters available at ”[http://www.tvhandbook.com//support//TVHB\\_updates.htm](http://www.tvhandbook.com//support//TVHB_updates.htm)”).
- [124] B. Woo and E. Yang. A Bayesian hierarchical model for noise reduction in low-field magnetic resonance imaging. In *Proceedings of Medical Image Understanding and Analysis, MIUA '01*, July 2001.
- [125] J. Wood and K. Johnson. Wavelet-denoising of complex magnetic resonance images. In *Proceedings of the IEEE 24th Annual Northeast Bioengineering Conference*, pages 32 – 34, 1998.
- [126] J. C. Wood and K. M. Johnson. Wavelet packet denoising of magnetic resonance images: importance of Rician noise at low SNR. *Magnetic Resonance in Medicine*, 41(3):631–5, March 1999.
- [127] Z. Q. Wu, J. A. Ware, and J. Jiang. Wavelet-based Rayleigh background removal in MRI. *Electronics Letters*, 39:603 – 604, Apr. 2003.
- [128] J.-H. Xue, A. Pizurica, W. Philips, E. Kerre, R. V. D. Walle, and I. Lemahieu. An integrated method of adaptive enhancement for unsupervised segmentation of MRI brain images. *Pattern Recognition Letters*, 24(15):2549–2560, Nov. 2003.
- [129] Y. Yang. Minimax nonparametric classification i. rates of convergence. *IEEE Transactions on Information Theory*, 45(7):2271–2284, Nov 1999.

- [130] R. K. Young. *Wavelet theory and its applications*. Kluwer Academic Publisher, Boston, 1993.