

DIFFUSION WEIGHTED IMAGE RECONSTRUCTION:

A BAYESIAN APPROACH

by

Mustafa Farrah

A Dissertation submitted in

Partial fulfillment of the

Requirements for Degree of

Doctor of Philosophy

in Engineering

The University of Wisconsin-Milwaukee

May 2015

ABSTRACT
DIFFUSION WEIGHTED IMAGE RECONSTRUCTION:
A BAYESIAN APPROACH

by

Mustafa Farrah

The University of Wisconsin-Milwaukee, 2015
Under the Supervision of Professor Chiu Law

Diffusion weighted imaging (DWI) is a form of Magnetic Resonance Imaging (MRI) based upon measuring the random Brownian motion of water molecules within a voxel of tissue. Recently, DWI has become an important modality in the diagnostic work-up of early identification of ischemic stroke, differentiation of acute stroke from other conditions that mimic it, differentiation of epidermoid cyst from arachnoid cyst, assessment of cortical lesions in Creutzfeldt Jakob disease(CJD), assessment of active demyelination and many others.

Diffusion MRI suffers from the a low Signal to Noise Ratio (SNR) due to the presence of noise from the measurement process. Many techniques have been proposed to increase the SNR of the images. Some of these are averaging several acquisitions in order to reduce the noise variance, but these techniques require more time. Other denoising techniques that work on the DWI images are Principal Component Analysis (PCA), Non local-means algorithm, and Discrete Cosine Transform (DCT).

In this thesis we propose a Bayesian approach to denoise the DWI images. DWI images are acquired at certain b-value, which is a factor of diffusion weighted sequences. A higher b-value leads to a stronger diffusion weighting. When acquiring DWI images, we always get an image with no diffusion, known as the b0 image. This image has less noise than the other DWI images and hence can be used to improve the quality of those images. The Bayesian approach has been applied successfully in many medical applications. In this thesis, we will use the entropy between the DWI and the b0 image as a measure of similarity between the two images, i.e. lower joint entropy for closely matching images.

Contents

Abstract	ii
List of Figures	ix
Acknowledgment	x
Declaration	xi
1 Introduction:	1
1.1 Introduction to Diffusion Weighted Imaging:	1
1.1.1 What is Diffusion	2
1.1.2 Using MRI to Measure Diffusion	3
1.1.3 Diffusion Gradients and the b-value:	5
1.1.4 Diffusion weighted Images Vs Apparent Diffusion Coefficients:	7
1.1.5 Optimum b-value and Its Relation to the Signal Intensity:	9
1.2 Noise in MRI	10
1.3 Contemporary Denoising Methods:	12
1.3.1 Discrete Cosine Transform:	12
1.3.1.1 Introduction	12
1.3.1.2 Description of the Method:	12
1.3.2 Non-Local Means Filter:	14
1.3.2.1 Introduction	14
1.3.2.2 NL-Means Algorithm	16

1.3.3	Overcomplete Local Principal Component Analysis:	17
2	Diffusion Weighted Imaging Reconstruction Using Bayesian Approach:	20
2.1	Image Entropy	21
2.2	Bayes' Theorem	22
3	Experimental Comparison between the Entropy Bayesian Approach and Contemporary Methods:	26
3.1	EBA Comparison with DCT:	27
3.2	EBA Comparison with NLM:	30
3.3	EBA Comparison with PCA:	33
3.4	Comparison Summary:	36
4	Simulation and Experimental Results	37
4.1	Simulation Results	38
4.2	Experimental Results	42
4.2.1	Noise Study	42
4.2.2	Reconstruction Images	48
4.2.2.1	Sphere Phantom	48
4.2.2.2	Sphere Phantom for b=500	49
4.2.2.3	Sphere Phantom for b=1500	52
4.2.2.4	Sphere Phantom for b=2500	55
4.2.2.5	Sphere Phantom for b=5000	59
4.2.2.6	Sphere Phantom for b=10000	62
4.2.3	Logo Phantom	66
4.2.3.1	Logo Phantom b=50	67
4.2.3.2	Logo Phantom b=100	70
4.2.3.3	Logo Phantom b=200	73
4.2.3.4	Logo Phantom b=400	76
4.3	Anatomical Results	79
4.3.1	Brain Results	79

4.3.2	Propeller Images With Fixed Beta Value.	86
4.3.2.1	Slice 5	88
4.3.2.2	Slice 14	91
4.3.2.3	Slice 27	93
4.3.2.4	Slice 28	96
4.3.3	Propeller Data with Variable Beta Value.	99
4.3.3.1	Slice 17	99
4.3.3.2	Slices 19 and 21	101
4.3.3.3	Slice 28	103
5	Conclusion and Future Work	104
	References	108
	Curriculum Vitae	113

List of Figures

1.1	Restricted Diffusion Vs Free Diffusion Image	3
1.2	Diffusion MRI Brain Image During Stroke	4
1.3	Diffusion Weighted Images Pulse Sequence	5
1.4	T2 Shine Through	7
1.5	Diffusion Weighted Image vs Apparent Diffusion coefficient.	8
1.6	Effect of b-value on The SNR	9
1.7	The Rician Distribution for Several Signal to Noise Ratios, N/σ .	11
1.8	Image Similarity in Non-Local Means Algorithm	15
1.9	Weight Assignment in Non-Local Means Algorithm	15
1.10	Principal Component Analysis in 2-D	17
1.11	Example of local matrix X formation from a DW image series	18
2.1	Joint Entropy of Images.	22
3.1	EBA ((b) and (c)) and DCT ((e) and (f)) Comparison, Slice 3.	27
3.2	EBA ((b) and (c)) and DCT ((e) and (f)) Comparison, Slice 5.	28
3.3	EBA ((b) and (c)) and DCT ((e) and (f)) Comparison, Slice 20.	28
3.4	EBA ((b) and (c)) and NLM ((e) and (f)) Comparison, Slice 3.	30
3.5	EBA ((b) and (c)) and NLM ((e) and (f)) Comparison, Slice 5.	31
3.6	EBA ((b) and (c)) and NLM ((e) and (f)) Comparison, Slice 20.	31
3.7	EBA ((b) and (c)) and PCA ((e) and (f)) Comparison, Slice 3.	33
3.8	EBA ((b) and (c)) and PCA ((e) and (f)) Comparison, Slice 5.	34
3.9	EBA ((b) and (c)) and PCA ((e) and (f)) Comparison, Slice 20.	34
4.1	Simulation Icon Images.	39
4.2	Icon Simulation Results.	40
4.3	Icon SNR vs Beta.	41
4.4	Logo (Cylinder) Noise Data for DWI Images with Various b-values.	43
4.5	Logo Noise Histograms.	44
4.6	Spherical Phantoms Noise Data for DWI Images with Various b-values.	45
4.7	Spherical Phantom Noise Histograms.	46
4.8	Spherical Phantoms Noise Histograms.	47
4.9	Sphere Logo Images.	48
4.10	Sphere Logo b500.	49
4.11	b=500 Noise Study	50
4.12	SNR vs Beta for b=500.	51

4.13 Spherical Phantom Image Reconstruction Results for b-value=500.	51
4.14 Spherical Phantom for b=1500.	52
4.15 Noise Study for DWI with b=1500.	53
4.16 SNR vs Beta for b=1500.	54
4.17 Spherical Phantom Image Reconstruction Results for b=1500.	54
4.18 Spherical Phantom for b=2500.	55
4.19 Noise Study for DWI with b=2500.	56
4.20 SNR vs Beta for b=2500.	57
4.21 Spherical Phantom Image Reconstruction Results for b=2500.	58
4.22 Spherical Phantom for DWI with b=5000.	59
4.23 Noise Study for DWI with b=5000.	60
4.24 SNR vs Beta for DWI with b=5000.	61
4.25 Spherical Phantom Image Reconstruction Results for DWI with b=5000.	61
4.26 Spherical Phantom b=10000.	62
4.27 b=10000 Noise Study.	63
4.28 Spherical Phantom Image Reconstruction Results for b=10000.	64
4.29 SNR vs Beta for b=10000.	65
4.30 SNR comparison for different b = 500, 1500, 2500.	65
4.31 Sphere-GE Logo Images.	66
4.32 Logo Phantom Images b=50.	67
4.33 Logo Noise Study for b=50.	68
4.34 Logo SNR vs Beta for b=50.	68
4.35 Logo Results for b=50.	69
4.36 Sphere Logo for b=100.	70
4.37 Noise Study b=100.	71
4.38 Logo SNR vs Beta for b=100.	71
4.39 Logo Results for b=100.	72
4.40 Sphere Logo b200.	73
4.41 Logo Noise Study for b=200.	74
4.42 Logo SNR vs Beta for b=200.	74
4.43 Logo Results for b=200.	75
4.44 Sphere Logo b400.	76
4.45 b400 Noise Study	77
4.46 Logo SNR vs Beta for b=400.	77
4.47 Logo Results for b=400.	78
4.48 Brain Images With Pathology.	80
4.49 Pathology Histograms.	80
4.50 Pathology Gaussian Noise Study.	81
4.51 Pathology Rician Noise Study.	82
4.52 Brain Pathology SNR vs Beta	83
4.53 Brain Results With Gaussian Noise	84
4.54 Brain Results with Rician Noise.	85
4.55 Propeller	86
4.56 Propeller Motion Artifact Correction.	87
4.57 Propeller Images Slice 5.	88
4.58 Propeller Images Slice 5 Noise Analysis Data.	89
4.59 Propeller Images Slice5 Results.	90

4.60	Slice 5 SNR vs Beta.	90
4.61	Propeller Images Slice 14.	91
4.62	Propeller Images Slice 14 Results.	92
4.63	Propeller Images Slice27.	93
4.64	Propeller Images Slice 27 Noise Analysis Data.	94
4.65	Propeller Images Slice 27 Results.	95
4.66	Propeller Images Slice 28.	96
4.67	Propeller Images Slice 28 Noise Analysis.	97
4.68	Slice 28 SNR versus Beta for Data from Slice 28.	98
4.69	Propeller Images Slice 28 Results.	98
4.70	Propeller Slice 17 Images.	99
4.71	Propeller Slice17 Results.	100
4.72	Propeller Slices 19 and 21 Images.	101
4.73	Propeller Slice19 and Slice 21 Results.	102
4.74	Propeller Slice 28 Images.	103
4.75	Propeller Slice 28 Variable Beta	103

Acknowledgment

I want to thank my advisers Professor Chiu Law and Professor Daniel Rowe. Thanks to Professor Law who was my academic adviser and tried his best helping me with my research although it was not his research area. I feel I was so lucky to work with Professor Rowe who was my research adviser. Professor Rowe was so helpful through my research without him; this thesis had no way to be accomplished. Professor Rowe was always available for meeting with me whenever I needed him, he spent time with me deriving the algorithm and providing me with the back-ground I needed. I also want to thank Dr. Kevin Kings at GE for providing experimental data and his insight into the project.

Thanks to my friend Emad Abdel Salam who helped me with acquiring data from GE and introduced me to Dr. Kings. Special thank for Professor George Hanson the chair of the EE department for his support through the difficulties that I had. I also would like to thank my thesis committee, Professor Hosseini, Professor Adel Nasiri, and Professor Ramin Pashaie.

Mustafa Farrah, Milwaukee

I hereby declare that this thesis entitled “Diffusion Weighted Image Reconstruction: A Bayesian Approach” is the result of my own research except as cited in the references. This thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature :

Student : Mustafa Farrah

Date : January 2015

Supervisor : Chiu Law

Co-Supervisor: Daniel Rowe

Chapter 1

Introduction:

Diffusion-weighted imaging (DWI) has been used to help detect early stroke and abscess, epidermic, cellular tumors and other neurological diseases [16]. The use of faster magnetic resonance imaging (MRI) technique made it possible to acquire diffusion weighted images of the whole-body. These techniques include multi-channel coils, echo planar imaging, and stronger gradients [4]. Diffusion-weighted imaging finds applications in oncologist imaging of the liver [12], prostate gland [11], and breast, as well as whole-body imaging [13]. Unfortunately, DWI images are corrupted with noise. Many algorithms have been proposed to denoise DWI images. In the first Section of this chapter, we will explain the principles of DWI. In the second Section, we will discuss the noise distribution in MRI. In the third Section, we will explain some denoising techniques that have been applied to DWI images.

1.1 Introduction to Diffusion Weighted Imaging:

In this Section, we explain the principles of DWI by considering first the diffusion phenomenon of water molecules, and then discussing the use of MRI to measure

diffusion in the body. We will examine the difference between diffusion weighted imaging and apparent diffusion coefficient (ADC) mapping. Finally, we will define the b-value and its relation to the image quality.

1.1.1 What is Diffusion

Water molecules are in constant motion. This motion depends on the thermal energy and the change of the concentration. When there is a change in the concentration of molecules, then the molecules will move from the high concentration to the low concentration. This flow depends on the change of concentration in the x direction and can be described by:

$$F = -D \frac{dC}{dx} \quad (1.1)$$

where F is the diffusion flux which is the amount of substance per unit area per unit time, and diffusion coefficient D controls the amount of substance that will flow through a small area during a small time interval, C is the concentration or the amount of substance per unit volume, and x is the position

The ideal diffusion is described by Brownian motion and is given by:

$$C = C_0 \exp\left[-\frac{x^2}{2\sigma^2}\right] \quad (1.2)$$

where C_0 is the concentration at $t=0$ and σ is the standard deviation of a Gaussian function. This concentration takes a Gaussian distribution in space.

Diffusion is a random process of movement of molecules; it can also be defined as the microscopic movement of atoms or molecules in a solution or gas. In living tissues, molecules of water and other chemical flow freely through the various tissue of the body. However in certain pathological conditions such as the

tightly packed cells of tumors, the movement of molecules is restricted and MRI can be used to evaluate diffusion restriction [14].

The movement of molecules in most fluid is homogeneous and moves in all spatial directions, so called isotropic diffusion. Cell membranes, vascular structures, axon cylinders, and fibers, are highly structured and further restrict the diffusion of molecules. Since the movement will not be the same in different directions in these cases, this type of diffusion is described as anisotropic [3].

1.1.2 Using MRI to Measure Diffusion

Stejskal and Tanner [26] described an MR experiment that could be applied to the detection and quantification of water diffusion in vivo [12]. They applied a symmetric pair of diffusion-sensitizing pulses around the 180° refocusing pulse. Figure 1.1 described this experiment. A spin echo pulse is made of a 90° radio frequency (RF) pulse followed by 180° refocusing pulse.

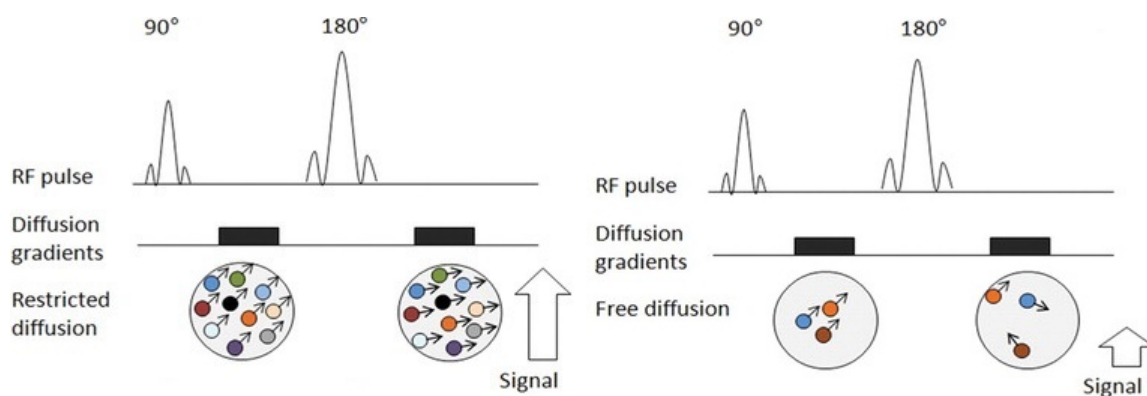


Figure 1.1: Restricted Diffusion Vs Free Diffusion ¹

To acquire a DWI image, two diffusion gradient pulses are applied around the 180° pulse. The first one introduces a phase shift that is dependent on the

¹<http://Principles and Applications of Diffusion-weighted Imaging in Cancer Detection, Staging, and Treatment Follow-up> Ashkan A. Malayeri, Riham H. El Khoul, Atif Zaheer, Michael A. Jacobs, Celia P. Corona-Villalobos, Ihab R. Kamel, and Katarzyna J. Macura RadioGraphics 2011 31:6 , 1773-1791

strength of the gradient at the position of the spin at $t = 0$. Then a 180° RF pulse is applied, this pulse will reverse the phase shift induced by the first gradient pulse. Then a second diffusion gradient pulse will be applied. The diffusion gradient causes the field intensity to vary with position. In the case of restricted diffusion, as shown on the left hand-side of Figure 1.1, the movement of the water molecules will be restricted. So after the second diffusion gradients, all molecules will be at the same location resulting in high signal intensity. However in the case of the free diffusion, as shown on the right hand-side of Figure 1.1, water molecules will not be at the same location when applying the second diffusion gradient pulse. This will cause the molecules to acquire different phase information from the first gradient. The molecules signals will not be completely rephased by the second gradient, which will lead to a loss in the acquired signal. The degree of the molecule diffusion can be inferred by measuring the attenuation of the signal intensity at DWI.

One application of DWI is in the early detection of stroke. Figure 1.2 shows diffusion MRI images of a brain that was taken for a patient who had a stroke. The bright area on the left hand-side of the brain indicates high diffusion signal, which indicates restricted diffusion due to stroke [12].

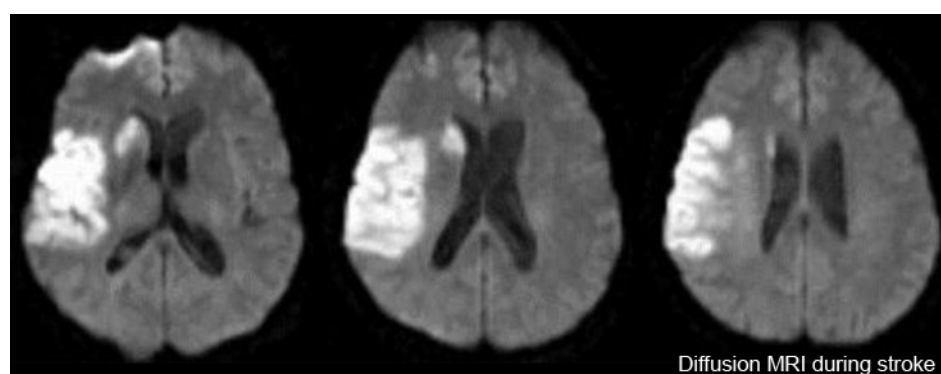


Figure 1.2: Diffusion MRI Brain Image during Stroke: The bright areas of the left side of the brain indicate restricted diffusion due to a stroke. ²

² Schaefer PW, Grant PE, Gonzalez RG. Diffusion-weighted MR imaging of the brain. *Radiology* 2000; 217:331-345.

1.1.3 Diffusion Gradients and the b-value:

In this Section, we explain a pulse sequence that was proposed by Stejskal and Tanner to detect and quantify the diffusion of water [26][1]. The pulse sequence that they proposed is shown in Figure 1.3, where *RF* is the RF pulse, *GS* is the slice selective pulse, *GP* is the phase encoding pulse, and *GM* is the readout direction pulse.

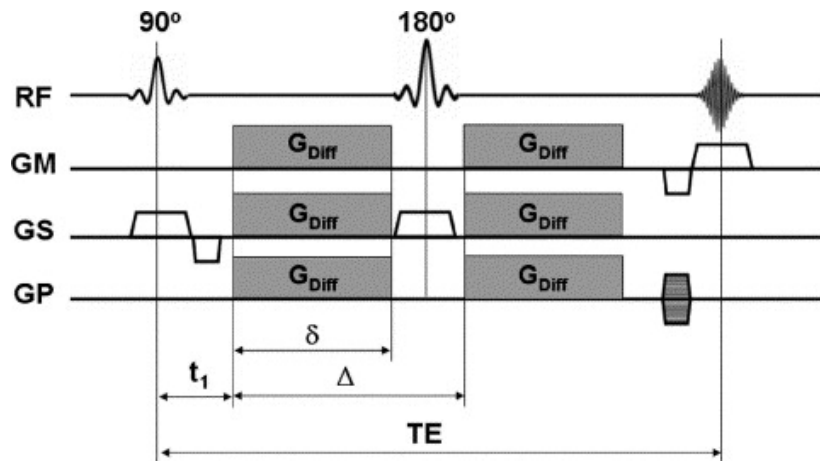


Figure 1.3: Diffusion Weighted Images Pulse Sequence ³

This pulse sequence is a spin echo pulse with two diffusion gradient pulses added around the 180° pulse. The diffusion gradient pulse duration is δ , and the time between the start of the two pulses is Δ .

The phase shift of a single static spin in the presence of a magnetic field gradient is given by:

$$\phi(t) = \gamma B_0 t + \gamma \int_0^t G(\tau) \cdot x(\tau) d\tau \quad (1.3)$$

where γ is the gyromagnetic ratio, B_0 is the strength of the external magnetic field, and $G(\tau)$ is the strength of the diffusion gradient.

³Bammer, Roland, "Basic principles of diffusion-weighted imaging", European journal of radiology (2003), 169–184.

This phase shift depends on the strength of the Gradient that is applied, the time duration that the gradient will be applied and the spatial location of the spin.

At the echo time TE , which represents the time in milliseconds between the application of the 90° pulse and the peak of the echo signal, this phase will be given by:

$$\phi(TE) = \gamma \int_{t_1}^{t_1+\delta} G(\tau).x(\tau)d\tau - \gamma \int_{t_1+\Delta}^{t_1+\Delta+\delta} G(\tau).x(\tau)d\tau \quad (1.4)$$

where δ is the duration of each sensitizing gradients and Δ is time between the start of the two pulses.

The diffusion weighted signal at time TE is given by:

$$DWI(b, TE) = I_0 \exp\left(-\frac{TE}{T_2}\right) \exp(-bD) \quad (1.5)$$

where I_0 is the signal intensity in the b_0 image which is a non-diffusion MRI image, and T_2 is the transverse relaxation time which determines the rate at which excited molecules go out of phase with each other.

This signal depends on the apparent diffusion coefficient (ADC), which is a measure of the magnitude of diffusion within tissue, and a sensitization parameter known as the b-value which is given by:

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3} \right) \quad (1.6)$$

and is measured in s/mm^2 . The b-value can be set by changing the

duration of the pulse, the time interval between the two pulses or the strength of the gradient. In practice, on clinical MR scanners, the diffusion sensitivity is varied by changing the strength of the sensitization gradient.

1.1.4 Diffusion weighted Images Vs Apparent Diffusion Coefficients:

The DWI signal $DWI(b, TE) = I_0 \exp(-\frac{TE}{T_2}) \exp(-bD)$ has two terms, T_2 signal ($\exp(-\frac{TE}{T_2})$), and diffusion ($\exp(-bD)$). In some cases, it is not obvious if a bright area in an image is due to restricted diffusion or due to the T_2 image. Figure 1.4 illustrates this situation. The bright bulbs on the left-hand side of the graph can indicate restricted area. The left-hand side of the graph is a T_2 image. It shows bright bulbs on the left-hand side. The middle image is from DWI. It also has the same bright bulbs which makes it hard to distinguish if these two bulbs are due to restricted diffusion or the T_2 images. The graph on the right is the ADC where the bright bulbs correspond to regions without restricted diffusion.

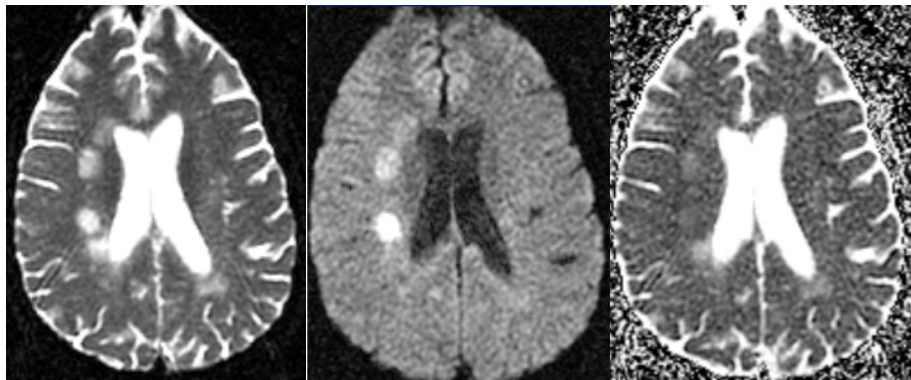


Figure 1.4: T2 shine through. Left image is T2 image, middle image is DWI, right is ADC. ⁴

To eliminate the T_2 factor, two images can be taken; the first one will be T_2 image with no diffusion. This can be done by setting the b-value = 0, so

⁴Schaefer PW, Grant PE, Gonzalez RG. Diffusion-weighted MR imaging of the brain. Radiology 2000; 217:331-345.

Equation 1.5 becomes:

$$DWI(TE) = I_0 \exp\left(-\frac{TE}{T_2}\right) \quad (1.7)$$

A second image at the desired b-value can be taken. If the equation 1.5 is divided by the equation 1.7, we get:

$$\frac{DWI(b, TE)}{DWI(TE)} = \exp(-bD) \quad (1.8)$$

This division is calculated pixel by pixel in the two images. The apparent diffusion coefficient (*ADC*) can be computed by:

$$D = -\frac{1}{b} \log\left(\frac{DWI(b, TE)}{DWI(TE)}\right) \quad (1.9)$$

In Figure 1.5 a diffusion weighted image is shown in the left side and the corresponding ADC image is shown in the right. The DWI image shows a bright bulb on the right side of the brain which can be due to restricted diffusion, ADC image shows the bulb dark, which indicates restricted diffusion is not caused by the T2 effect.

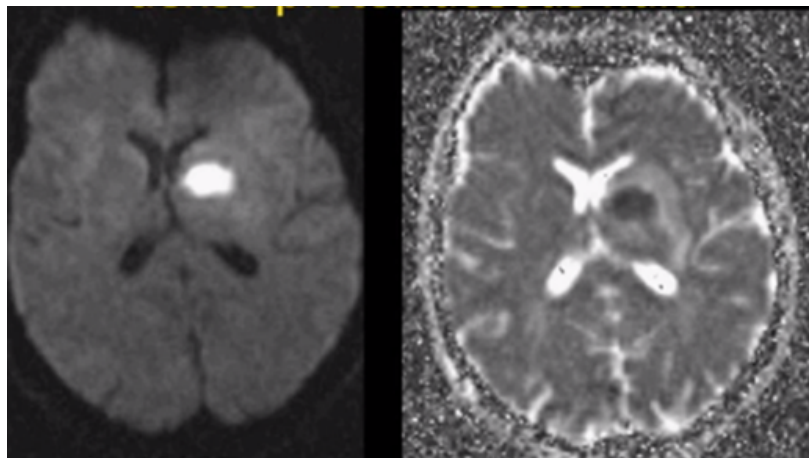


Figure 1.5: Diffusion Weighted Image vs Apparent Diffusion Coefficient. ⁵

1.1.5 Optimum b-value and Its Relation to the Signal Intensity:

The sensitivity of the DWI sequence to water motion can be varied by changing the b-value which is proportional to the three factors mentioned before. If the b-value is increased, the contrast between different tissues will become more apparent; however the image will have more noise. Figure 1.6 illustrates the logarithm of relative signal intensity (SI) versus the b-value for tumor and normal tissue. The slope of the “tumor line” is less than that of the line representing normal tissue, which translates into a lower signal on the ADC map [12]. If we increase the b-value, the contrast between the tumor and the normal cells will also be increased, however, this decreases the SNR value, so there is a trade-off between the SNR and the contrast in the image.

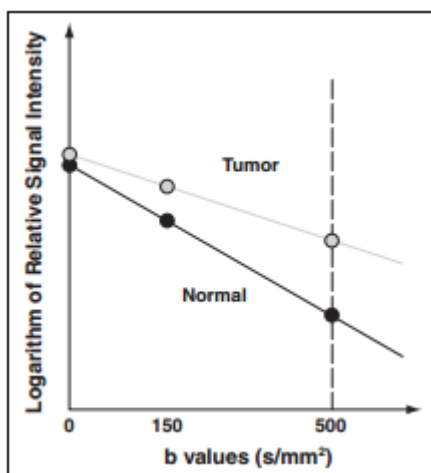


Figure 1.6: Effect of b-value ⁶

If the b-value is more pronounced, then the contrast between the normal tissue and the tumor will be increased, but the strength of the signals will be decreased as shown in Figure 1.1. Since the noise remains the same, increasing the b-value will decrease the signal intensity and hence lower the signal to noise ratio (SNR).

⁵Schaefer PW, Grant PE, Gonzalez RG. Diffusion-weighted MR imaging of the brain. *Radiology* 2000; 217:331-345.

⁶Malayeri, Ashkan A., et al. "Principles and applications of diffusion-weighted imaging in cancer detection, staging, and treatment follow-up." *Radiographics* 31.6 (2011): 1773-1791.

1.2 Noise in MRI

Signals in MRI images are measured using quadrature detectors which measure the MRI signal from two orthogonal directions. These detectors give the real and imaginary part of the signals. The measured signals are contaminated with Gaussian noise. MRI images are formed by taking the Fourier transform of the measured signals. Therefore, the MRI image will be a complex image. The noise distribution of the real part as well as the imaginary part of the complex image, are Gaussian. The magnitude image is then formed by a nonlinear mapping. It has been shown by many studies that noise distribution in the magnitude image is Rician [8].

The Rician distribution has probability density function (pdf) given by:

$$f(x | \nu, \sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2 + \nu^2}{2\sigma^2}\right) I_0\left(\frac{x\nu}{\sigma^2}\right) \quad (1.10)$$

where $I_\alpha(x)$ is the modified Bessel function of the first kind given by:

$$I_\alpha(x) = \frac{1}{\pi} \int_0^\pi \exp(x \cos(\theta)) \cos(\alpha\theta) d\theta - \frac{\sin(\alpha\pi)}{\pi} \int_0^\infty \exp(-x \cosh t - \alpha t) dt, \quad (1.11)$$

and non-centrality parameter $\nu \geq 0$ and scale parameter $\sigma > 0$. The noise distribution in MRI has the modified Bessel function of the first kind and zeroth order, which is given by

$$\begin{aligned}
I_0(x) &= \frac{1}{\pi} \int_0^\pi \exp(x \cos(\theta)) \cos(0\theta) d\theta - \frac{\sin(0\pi)}{\pi} \int_0^\infty \exp(-x \cosh t - 0t) dt, \\
I_0(x) &= \frac{1}{\pi} \int_0^\pi \exp(x \cos(\theta)) d\theta
\end{aligned} \tag{1.12}$$

In Figure 1.7 we show the Rician distribution for different values of $v = \frac{N}{\sigma}$ where N is the mean of the signal and σ is the standard deviation.

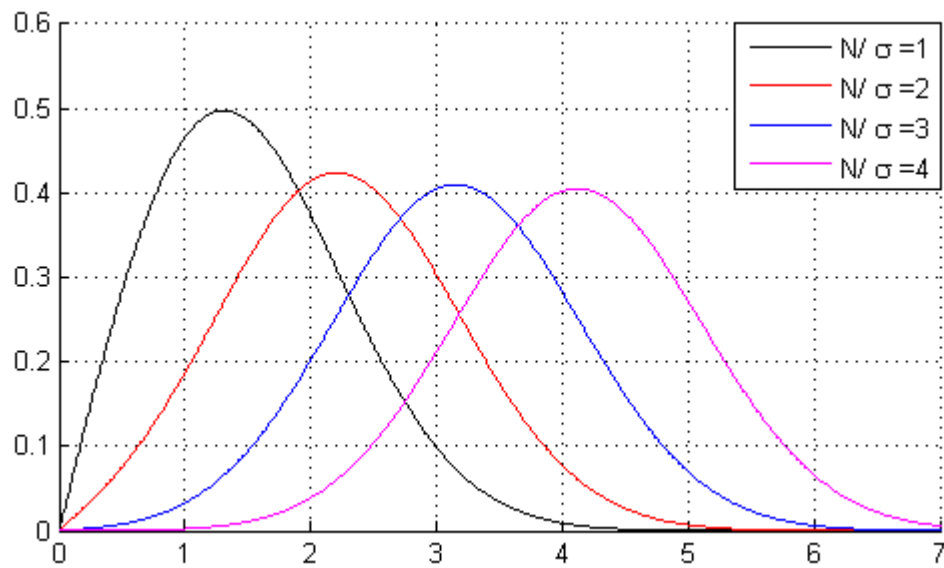


Figure 1.7: The Rician Distribution for Several Signal to Noise Ratios, N/σ .

The graph shows that when the SNR is greater than 3 ($N/\sigma \geq 3$) the Rician distribution can be approximated by the Gaussian distribution.

1.3 Contemporary Denoising Methods:

In this Section, we review 3 contemporary methods that have been applied to denoise DWI. These methods include the Discrete Cosine Transform (DCT) [18], the non local-means algorithm (NLM) [30], and the Overcomplete Local Principal Components Analysis(PCA) [17].

1.3.1 Discrete Cosine Transform:

1.3.1.1 Introduction

The first denoising method is the local 3D Discrete Cosine Transform (DCT) proposed by Guleryuz (2007) [9]. This method makes use of the sparseness of DWI. Sparseness means that an image can be represented by a small number of basis functions. This happens if the representation of the image in another domain or bases, such as the DCT, has many zero coefficients that can be disregarded without affecting the image itself.

1.3.1.2 Description of the Method:

The 3D discrete cosine transform is given by:

$$X_{k_1, k_2, k_3} = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} \sum_{n_3=0}^{N_3-1} x_{n_1, n_2, n_3} \cos \left[\frac{\pi}{N_1} \left(n_1 + \frac{1}{2} \right) k_1 \right] \cos \left[\frac{\pi}{N_2} \left(n_2 + \frac{1}{2} \right) k_2 \right] \cos \left[\frac{\pi}{N_3} \left(n_3 + \frac{1}{2} \right) k_3 \right].$$

where k_1, k_2, k_3 are the new basis of the signal and vary from 0 to $N_1, N_2,$ and N_3 respectively. Guzzler assumes that an image x is corrupted with white noise

n such that the observed image y is given by:

$$y = x + n. \quad (1.13)$$

For 2D as example, the estimated \hat{x} can be obtained by using a hard thresholding method. A set of 8×8 DCT blocks is used. In the case of voxels, a set of $4 \times 4 \times 4$ blocks is used.

A local denoised estimate of the image at block j can be computed by:

$$\hat{x}_j = H^{-1}(\hat{c}_j). \quad (1.14)$$

where H^{-1} is the inverse 3D cosine transform given by:

$$x_{n_1, n_2, n_3} = \sum_{k_1=0}^{N_1-1} \sum_{k_2=0}^{N_2-1} \sum_{k_3=0}^{N_3-1} X_{k_1, k_2, k_3} \cos \left[\frac{\pi}{N_1} \left(n_1 + \frac{1}{2} \right) k_1 \right] \\ \cos \left[\frac{\pi}{N_2} \left(n_2 + \frac{1}{2} \right) k_2 \right] \cos \left[\frac{\pi}{N_3} \left(n_3 + \frac{1}{2} \right) k_3 \right].$$

and \hat{c}_j are the transform coefficients of the j block. These coefficients are given by:

$$\hat{c}_j = T(c_j, \tau) \quad (1.15)$$

where T is a thresholding operator with threshold τ , and c_j is the DCT transform coefficients at block j . The estimation of x at position i , $\hat{x}(i)$ is computed by the weighted average rule given by:

$$\hat{x}(i) = \frac{\sum_{j=1}^V \theta_j \hat{x}_j(i)}{\sum_j^v \theta_j} \quad (1.16)$$

where θ_j relates to \hat{c}_j by:

$$\theta_j = \frac{1}{1 + \|\hat{c}_j\|_0} \quad (1.17)$$

V is the number of overlapping blocks contributing to $\hat{x}(i)$ and $\|\hat{c}_j\|_0$ is the L_0 norm which corresponds to the number of non-zero elements in \hat{c}_j .

1.3.2 Non-Local Means Filter:

1.3.2.1 Introduction

Non-local means filter is proposed by Buades [4]. It differs from the local means filter in the way that it computes the mean of each pixel. In local means filters, the mean value of pixels around a target pixel is computed to denoise the image. However in non-local means algorithm, it is assumed that the image has many similarities as shown in Figure 1.8 with many squares in red, yellow, blue and green. The squares that have the same color are similar to each other.

The estimation of a pixel will be based on the weight of the contribution from similarity blocks. This results in less loss in image fine details. When applying the filter to this image in Figure 1.8, pixels in similar boxes (with the same color) will contribute more than those in different boxes.

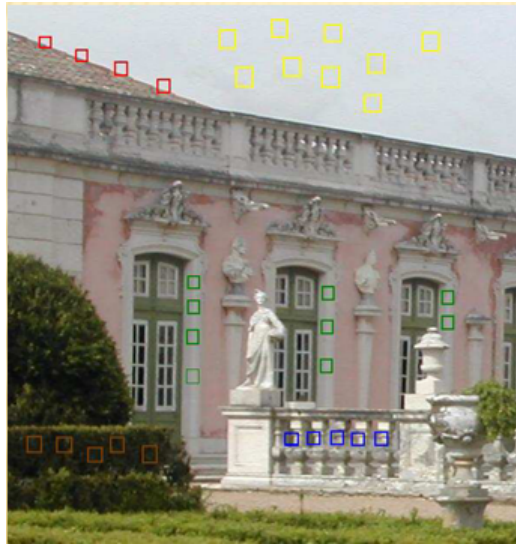


Figure 1.8: Image Similarity in Non-Local means Algorithm. The boxes with the same colors are similar to each others. ⁷

Figure 1.9 shows example of assigning weight of a target pixel p . q_1 will have a large weight $w(p, q_1)$ because its similarity window matches that of p . q_2 will have small weight $w(p, q_2)$ because its similarity window mismatches that of p .

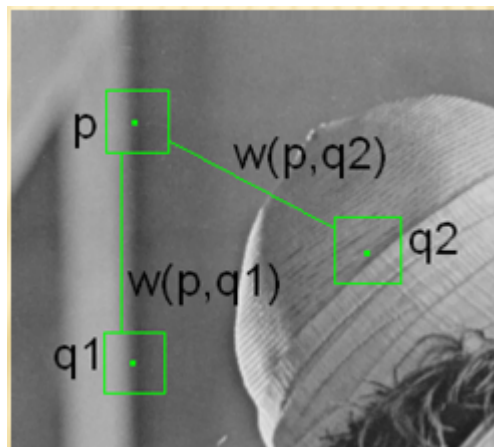


Figure 1.9: Weight Assignment in Non-Local means Algorithm. ⁸

⁷A. Buades, B. Coll, and J. Morel. On image denoising methods. Technical Report 2004-15, CMLA, 2004

⁸A. Buades, B. Coll, and J. Morel. On image denoising methods. Technical Report 2004-15, CMLA, 2004

1.3.2.2 NL-Means Algorithm

The image v is assumed to be distorted with noise. The non-local means (NL-means) can be used to estimate the true value for each pixel i using the formula:

$$NL[v](i) = \sum_{j \in I} w(i, j)v(j) \quad (1.18)$$

where $w(i, j)$ is the weight between pixel i and pixel j and depends on the similarity between the two pixels, and I is the image domain. This similarity depends on the similarity of the two gray vectors $v(N_i)$ and $v(N_j)$ where $N_{i,j}$ is the nearest neighborhood of pixel i and j . The similarity of N_i and N_j can be measured by weighted Euclidean distance $d(N_i, N_j) = \|v(N_i) - v(N_j)\|_{2,a}$ where a is the standard deviation of a Gaussian Kernel.

The Euclidean distance between two vectors x and y with length J is defined as $d_{x,y} = \sqrt{\sum_{j=1}^J (x_j - y_j)^2}$. The weighted Euclidean distance is given by $d_{x,y} = \sqrt{\sum_{j=1}^J (\frac{x_j}{a_j} - \frac{y_j}{a_j})^2}$; where a_j is the standard deviation of the j th variable.

The weights can be computed by:

$$w(i, j) = \frac{1}{Z(i)} e^{-\frac{d(v_i, v_j)^2}{h^2}} \quad (1.19)$$

where $Z(i)$ is normalization constant computed as:

$$Z(i) = \sum_j e^{-\frac{d(v_i, v_j)^2}{h^2}} \quad (1.20)$$

and h is a parameter that controls the decay rate of the exponential function.

1.3.3 Overcomplete Local Principal Component Analysis:

Principal Component Analysis (PCA) has been used for reducing noise in images [28]. In this context, noise can be removed by decomposing the signal into principal components, then disregarding the least significant ones, and finally reconstructing the signal back. PCA is an orthogonal linear transformation that maps the data into a new coordinate system such that the direction with the greatest variance, by any projection of the data, comes to lie on the first axis (called the first principal component), the direction of the second greatest variance on the second coordinate axis, and so on. PCA will ensure the second axis is perpendicular to the first variance and the same enforcement is applied to other axes.

When using PCA to decompose a signal that is distorted with noise into its principal components, the noise will be spread out over all components while the magnitude of the signal itself will be having the most significant component with the largest magnitude along the first direction, and the second largest component in the second direction and so on. Hence disregarding some principal components with small values will reduce noise and use the major principal components of the signal for reconstruction.

Figure 1.10 shows example of a 2D data which can be approximated by a 1D data set.

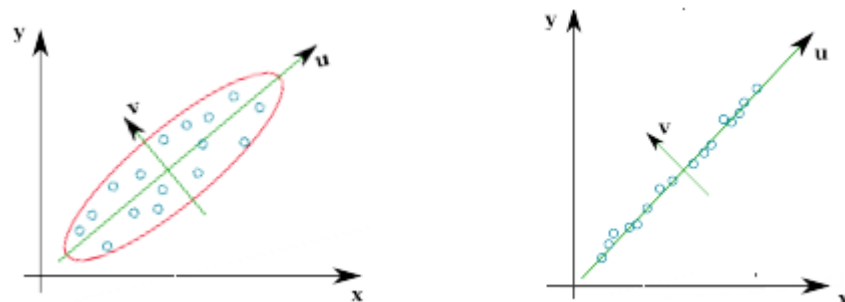


Figure 1.10: Principal Component Analysis in 2-D

The left-hand side of Figure 1.10 shows the 2D data set presented in the

x-y plane. The PCA will find the best coordinate u-v plane which will have the largest variance of the data along the u-axis, and the second largest along the v-axis. The right-hand side of Figure 1.10 shows the approximate representation of the data in 1D along the u-axis.

José V. Manjón [17] used PCA to denoise DWI images. In his method, instead of dividing each image into patches and composing a matrix from these patches, he made use of the idea that corresponding pixels in a series of k DWI images in k direction are very similar and hence these patches can be used to reconstruct a matrix that will be filtered.

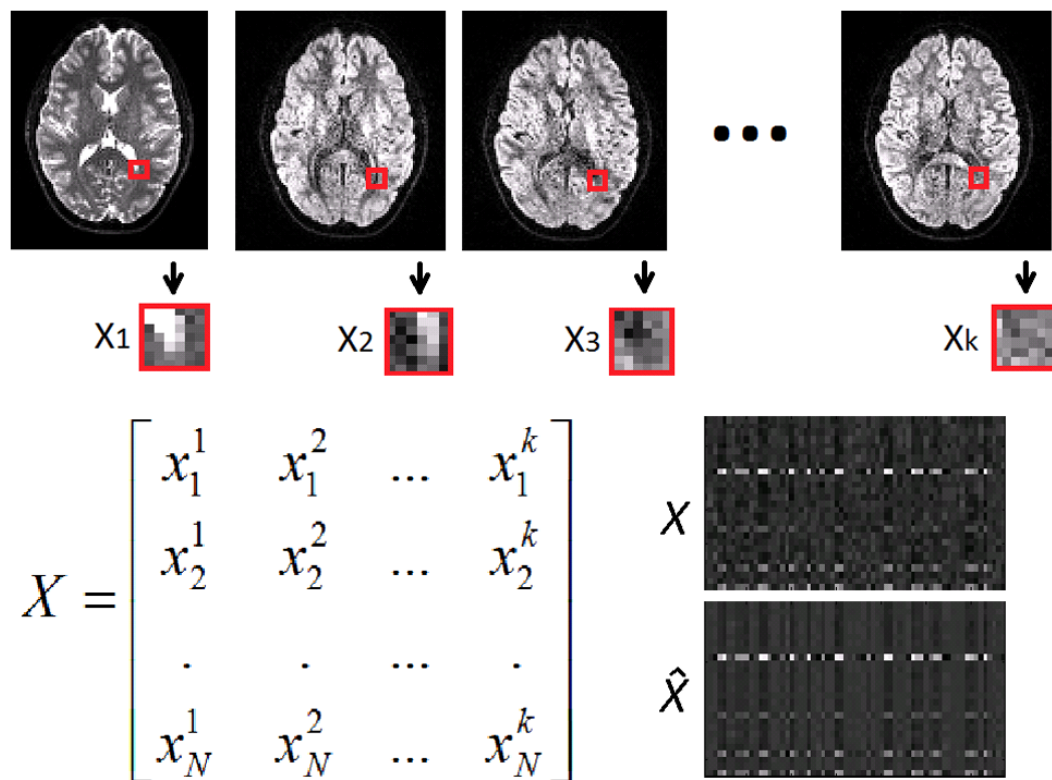


Figure 1.11: Example of local matrix X formation from a DWI image series ⁹

⁹Manjón JV, Coupé P, Concha L, Buades A, Collins DL, et al. (2013) Diffusion Weighted Image Denoising Using Overcomplete Local PCA. PLoS ONE 8(9): e73021. doi:10.1371/journal.pone.0073021

Figure 1.11 explains this idea. For a series of k images, a block of $4 \times 4 \times 4$ voxels is converted into a column of matrix X . Hence, the size of matrix X will be $64 \times k$. Then the PCA algorithm will be applied to matrix X , to compute its approximation \hat{X} from which the patches can be reconstructed. The bottom-right of Figure 1.11 shows matrix X and its filtered version \hat{X} to highlight the high level of the profile signal redundancy in the local matrix X .

In the second chapter of this thesis, we propose a new method for reconstructing DWI based on a Bayesian approach, this method will be called Entropy Bayesian Approach (EBA). Bayesian approach method relies on another image with low noise level to denoise the DWI. Since joint entropy will be used as a measure of similarity, we will explain the Bayesian approach and the joint entropy concept.

In the third chapter, we will compare the EBA with the three contemporary denoising methods discussed here. In the fourth chapter, we will show simulations and experimental results of our work. Finally, we will provide our conclusion and future work.

Chapter 2

Diffusion Weighted Imaging Reconstruction Using Bayesian Approach:

In the previous chapter, we explained the principle of DWI and the limitation that the noise imposed. We also discussed three methods that have been applied to denoise DWI.

In this chapter, we propose a new method to reconstruct DWI. This method uses Bayesian approach to reconstruct the DWI. In this approach [23], the joint entropy between the DWI image and the prior image is used as a measure of similarity. Since the proposed method is based on the Bayesian approach and the joint entropy, we call it Entropy Bayesian Approach (EBA). The EBA algorithm is independent on the noise level in the image, i.e. no matter how much noise the image has, the algorithm can improve it. This will make it possible to use this method with very high b-values where other methods are limited. If complex images are available, EBA can be readily applied and should be more accurate since the approximation on the noise can be removed. The EBA method can work on only one DWI and its prior while other methods, such as the PCA, depend on a sequence of images. We will explain the Bayesian approach in the next Section,

the image entropy in the next Section and then the EBA algorithm.

2.1 Image Entropy

Entropy is a measure of the uncertainty in a distribution of a random variable. Shannon defined the entropy $H(x)$ of a discrete random variable X with possible values x_1, \dots, x_n and probability mass function $P(X)$ as:

$$H(X) = - \sum_i P(x_i) \log_b P(x_i) \quad (2.1)$$

where b is the number of bits used to represent the outcome of the experiment. The joint entropy $H(X, Y)$ of a pair of discrete random variables with a joint distribution $p(x, y)$ is a simple extension that measures the uncertainty in the joint distribution of a pair of random variables [5]. Joint entropy $H(X, Y)$ is defined as:

$$H(X, Y) = - \sum_{i,j} p(x_{ij}, y_{ij}) \ln p(x_{ij}, y_{ij}) \quad (2.2)$$

For digital images, the joint entropy of two images can be computed using the joint histogram of the two images. The joint histogram is computed using the intensities of the corresponding pixels of the two images, and the value in the histogram represents the number of occurrences of intensity value pairs [7][22].

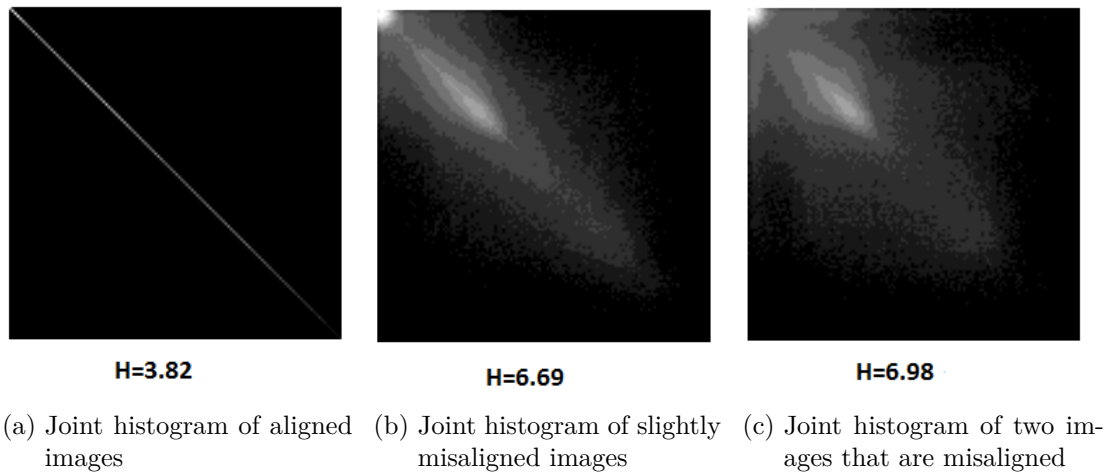


Figure 2.1: Joint Entropy of Images.

Figure 2.1 shows an example of the joint histogram of two images and the corresponding joint entropy. When the two images are the same and aligned with each others, the joint histogram will not be spread out and it has the least entropy. In Figure 2.1 (b) and (c), the joint histograms of misaligned images are shown, the more two images are misaligned or different, the more the joint histogram will be spread out and the joint entropy will be higher.

2.2 Bayes' Theorem

Bayes' theorem [27][8] allows us to write $p(x | y) = p(y | x)p(x)/p(y)$ where the likelihood, $p(y | x)$, describes the data formation process for a particular underlying image while the prior, $p(x)$, encodes any prior beliefs about the properties of such underlying images. In this thesis, we propose a new method to reconstruct DWI using Bayesian approach by incorporating prior information from another image that has high SNR. The prior image can be a T2 weighted image or it can be another DWI with lower b-value that has higher SNR.

While the Bayesian approach has been used in many applications such as image registration [24][15][33], it has not been successfully used in DWI yet.

Unlike other applications with real images, DWI has complex data which require a different approach. Numerous solutions have been proposed to solve functions with complex variables [2] [25]. Here, we treat the complex image as two real images, one corresponds to its real part, and another corresponds to its imaginary part.

Joint entropy has previously been used in many applications such as image registration [32][20][29] [10]. In EBA, it is applied as a measure to quantify similarity. We assume that the DWI is corrupted with Gaussian noise. A typical model of a signal corrupted by noise is

$$y_{ij} = x_{ij} + \epsilon \quad (2.3)$$

where y_{ij} is the observed image pixel, x_{ij} is the true image pixel, and ϵ is a Gaussian noise with 0 mean. A valid assumption about the observed image is that the real and imaginary parts of each pixel obey independent Gaussian distributions with identical standard deviation. This assumption is valid because they are measured by quadrature coils which constitute two pairs of identical coils oriented perpendicular to each other. As these coils are identical and uncoupled, the real signal from one pair of coils will not interfere the imaginary signal from the second pair of coils.

The distribution of the intensity of each image can be described by:

$$p(y_{ij} | x_{ij}) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\frac{1}{2\sigma^2}(y_{ij}-x_{ij})^2} \quad (2.4)$$

Using Bayes theorem, the posterior distribution can be found as:

$$p(x_{ij} | y_{ij})p(y_{ij}) = p(y_{ij} | x_{ij}) \cdot p(x_{ij}) \quad (2.5)$$

where $p(y_{ij}/x_{ij})$ is the likelihood function, and $p(x_{ij})$ is the prior density. We assume the prior density follows a Gibbs distribution of the form:

$$p(x_{ij}) = \frac{1}{w} e^{-\beta h(x_{ij}, g_{ij})} \quad (2.6)$$

where $h(x_{ij}, g_{ij})$ is the joint entropy between x_{ij} , g_{ij} :

$$h(x_{ij}, g_{ij}) = p(x_{ij}, g_{ij}) \cdot \ln p(x_{ij}, g_{ij}) \quad (2.7)$$

and w is a normalization constant. The prior image is the b0 image with pixel g_{ij} . Substituting equations 2.4, 2.6, 2.7, in equation 2.5 we get

$$p(x_{ij} | y_{ij}) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{\frac{-1}{2\sigma^2}(y_{ij}-x_{ij})^2} \frac{1}{w'} e^{-\beta p(x_{ij}, g_{ij}) \cdot \ln p(x_{ij}, g_{ij})} \quad (2.8)$$

where β is a parameter that controls the joint entropy term. If $\beta = 0$, the reconstructed image will be the same as the DWI. As β increases, the reconstructed image will become similar to the prior image or the b0 image.

Since the distribution of each pixel is identically independently distributed, we can write $p(x_{ij} | y_{ij})$ as the product of $\frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{\frac{-1}{2\sigma^2}(y_{ij}-x_{ij})^2} \frac{1}{w} e^{-\beta p(x_{ij}, g_{ij}) \cdot \ln p(x_{ij}, g_{ij})}$ for all possible values of i and j .

If we take the log of equation 2.7 we get

$$\ln(p(x_{ij} | y_{ij})) = \ln\left(\frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}}\right) + \frac{1}{w} - \sum_{ij} \frac{1}{2\sigma^2}(y_{ij} - x_{ij})^2 - \beta p(x_{ij}, g_{ij}) \cdot \ln p(x_{ij}, g_{ij}) \quad (2.9)$$

The solution of this equation $\hat{x} = \underset{x}{argmax} p(x | y)$ where $argmax$ is the set of points x for which $p(x | y)$ attains its largest value, can be found using nonlinear optimization technique like the nonlinear conjugate gradient [6][31]. We used Matlab built-in function `fminunc` to find the solution of equation 2.9. The first two terms in equation 2.9 are constants and the solution of equation does not

depend on them. Hence, they can be dropped without affecting its solution.

The inputs of the program are as follows:

y : the DWI image.

g : the b0 image

σ : the standard deviation of the DWI image which is computed by converting the DWI into a vector and calculating its standard deviation of the vector.

β : is the control parameter that is fixed for each image in the range of .01 to 0.5

x_o : the initial guess of the reconstructed image, which is the average of the DWI image and the b0 image.

The output of the program is x which is the new reconstructed DWI image.

We use the default values from Matlab for the input parameters below:

DiffMinChange: the minimum change in variables for finite-difference gradients with the default of 10^{-8}

MaxIter : the maximum number of iterations with the default value of 400.

Chapter 3

Experimental Comparison between the Entropy Bayesian Approach and Contemporary Methods:

The performance of the EBA algorithm has been compared to the three denoising methods discussed in chapter 1. The data for comparisons are obtained from Laboratoire Bordelais de Recherche en Informatique¹. There are 21 slices with each slice having 21 directions. The comparisons have been performed with laptop computer, 8-core 2GHz i7 processor with 8 Giga Bytes Memory. The performance of the algorithms is evaluated basing on the processing time, the mean squared error of difference between the DWI and the filtered image, and on fidelity in preserving the fine details or the edges in the image. For each algorithm, three slices are used for evaluation. The first denoising method to compare is the DCT algorithm.

¹<http://www.labri.fr>

3.1 EBA Comparison with DCT:

To compare the EBA method to the DCT method, three slices (slices 3,5 and 20) of a DWI are used. Figures 3.1, 3.2, and 3.3 show the results.

In these figures, images (a) and (d) are the noisy DWI, image (b) is the denoised image with EBA algorithm, and image (c) is the difference between (a) and (b). Image (c) has some areas where the anatomy of the image can be seen. They appear when the prior and the DWI image are not totally aligned. The mean squared errors (mse) are 8.58×10^{-4} for the slice 3, 7.39×10^{-4} for slice 5, and 7.325×10^{-4} for slice 20. The difference image in (c) shows that the noise was suppressed from all areas of the image. Image (e) in Figures 3.1, 3.2, and 3.3 are the denoised DWI images with DCT algorithm. Areas in (e) look noisier than those of the EBA.

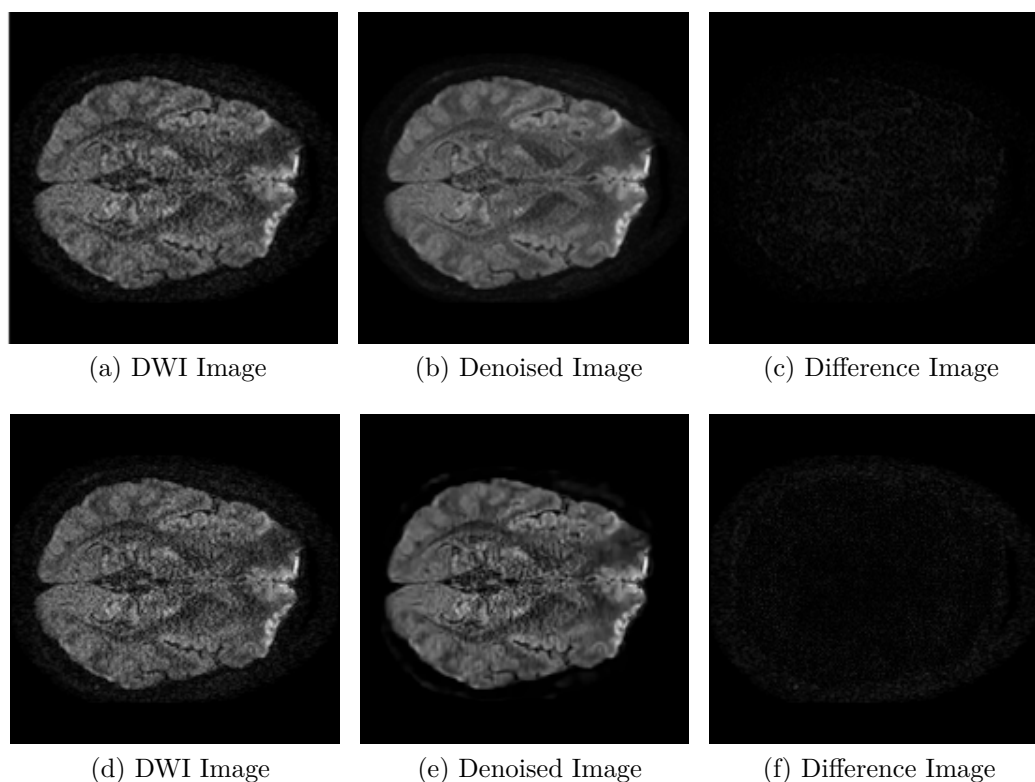


Figure 3.1: EBA ((b) and (c)) and DCT ((e) and (f)) Comparison, Slice 3.

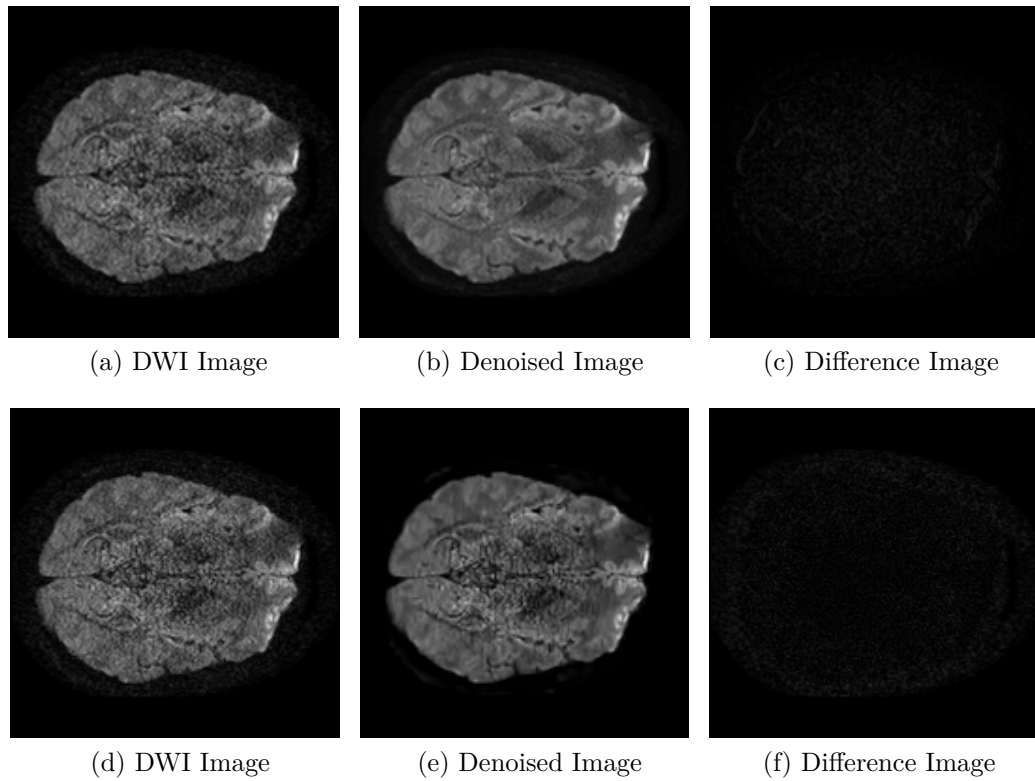


Figure 3.2: EBA ((b) and (c)) and DCT ((e) and (f)) Comparison, Slice 5.

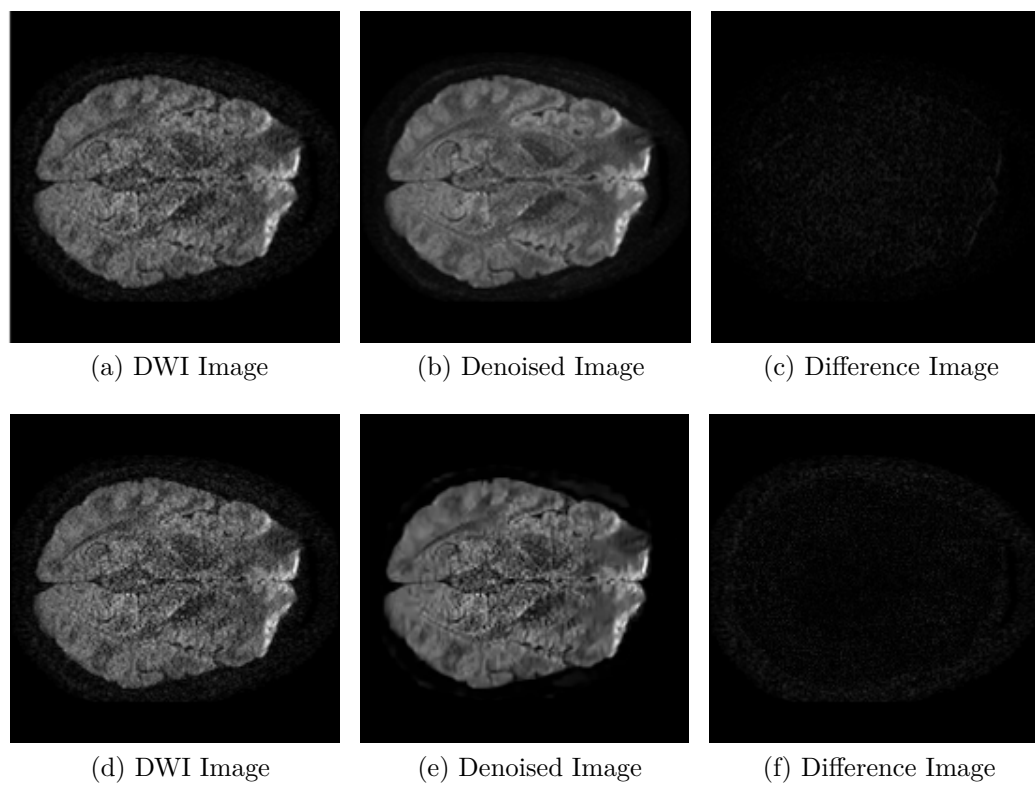


Figure 3.3: EBA ((b) and (c)) and DCT ((e) and (f)) Comparison, Slice 20.

The difference between the DWI image and the denoised image with DCT techniques is shown in image (f). The mses of the DCT algorithm are 9.06×10^{-4} for slice 3, 9.63×10^{-4} for slice 5, and 9.26×10^{-4} for slice 20. Consistently, the mse of the DCT slices is higher than the mse of the EBA slices. The difference images (f) in Figures 3.1, 3.2, and 3.3 show some dark areas in the middle, which indicate that there is no difference between the DWI image and the denoised image, i.e. the denoising algorithm is ineffective in these areas. The computation time of the DCT was 4 times higher than the computation time of the EBA.

3.2 EBA Comparison with NLM:

Next we compare the performance of the EBA method to the NLM algorithm. In NLM, noise is reduced according to the similarity in the image. The NLM algorithm is applied to the same slices in Section 3.1 and the results are compared to the EBA methods. Figures 3.4, 3.5, and 3.6 show the results for comparing EBA to NLM.

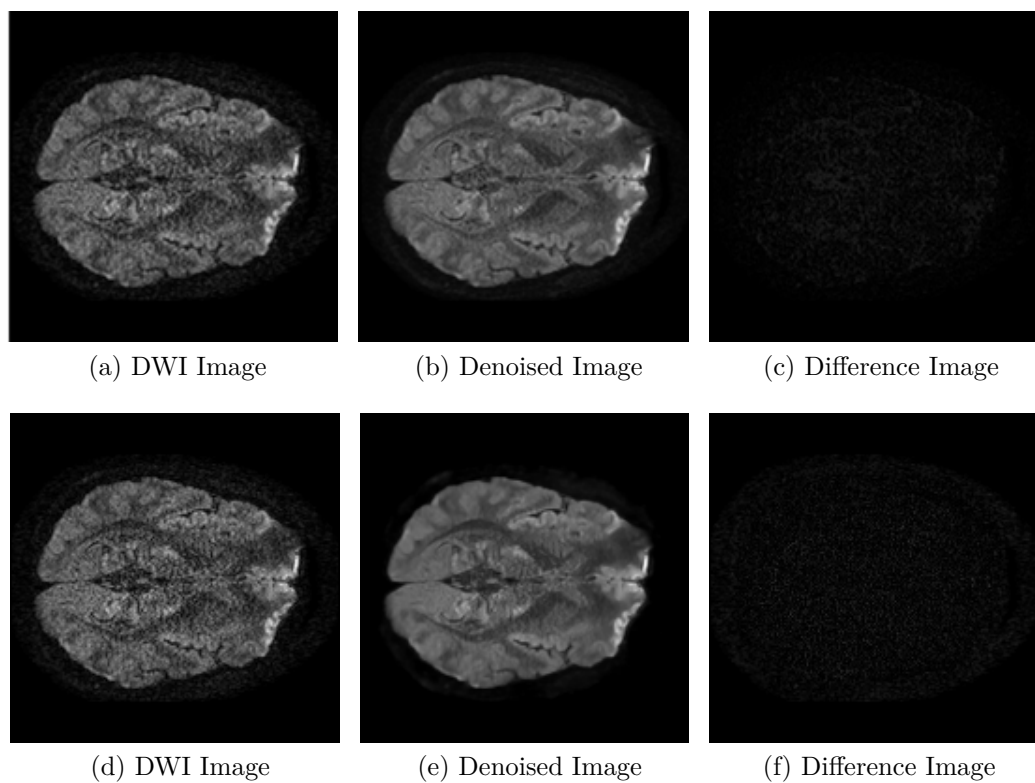


Figure 3.4: EBA ((b) and (c)) and NLM ((e) and (f)) Comparison, Slice 3.

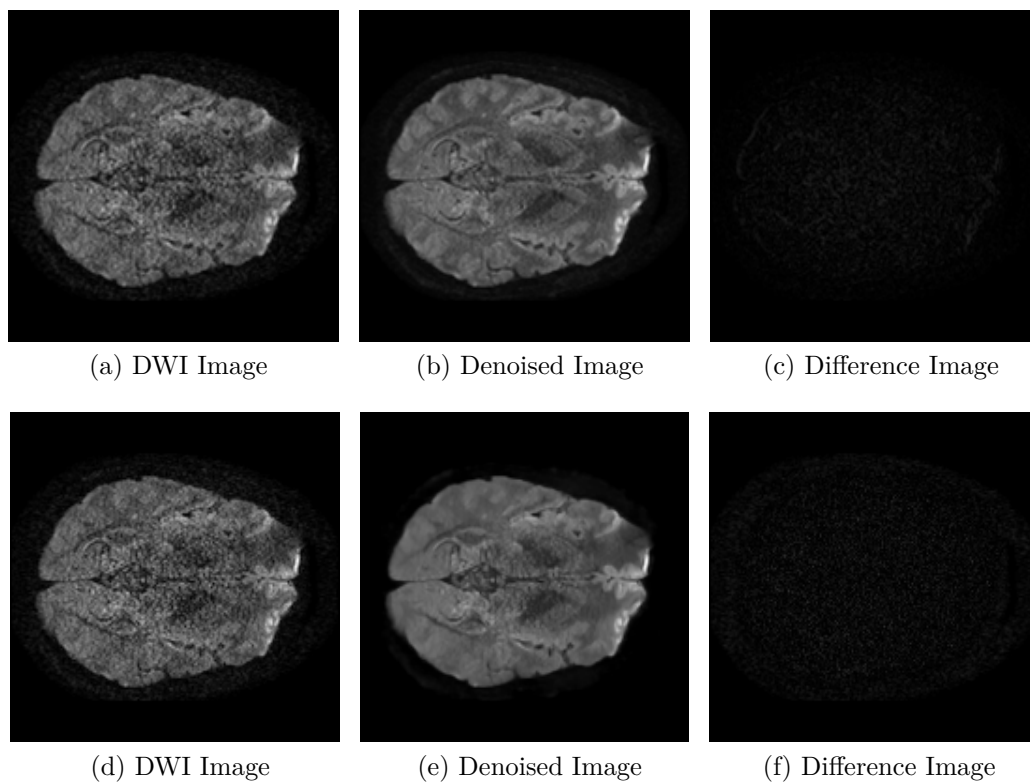


Figure 3.5: EBA ((b) and (c)) and NLM ((e) and (f)) Comparison, Slice 5.

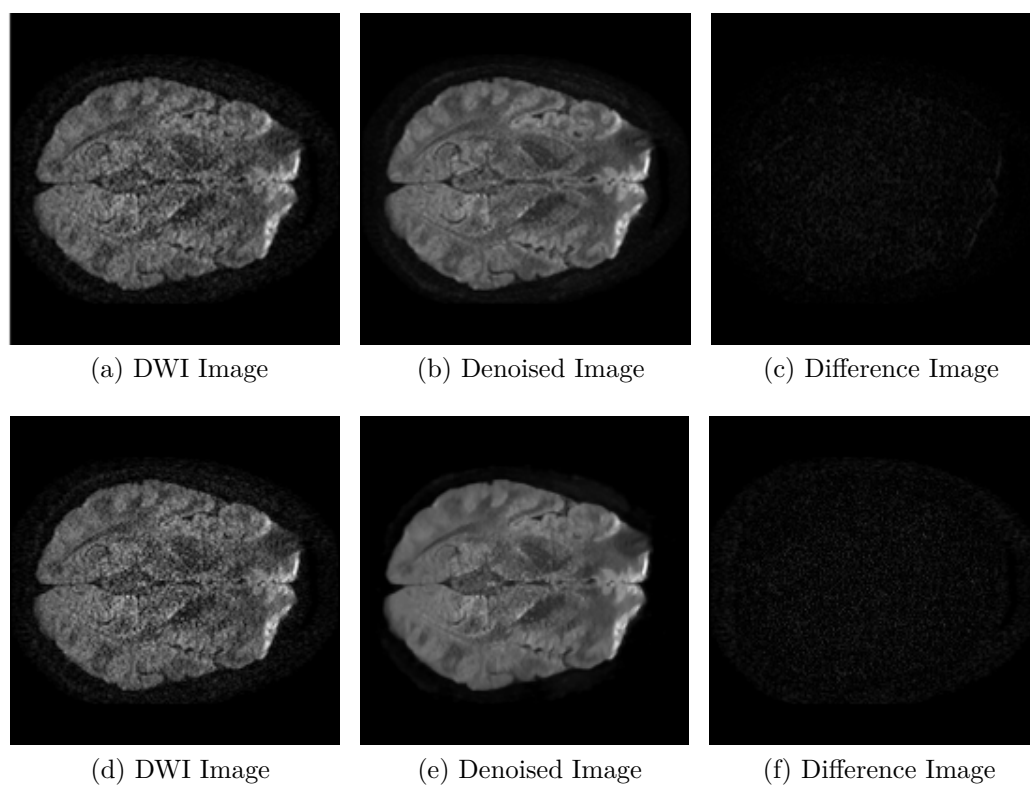


Figure 3.6: EBA ((b) and (c)) and NLM ((e) and (f)) Comparison, Slice 20.

The first three images in these figures, images (b) and (c), are the results of the EBA algorithm discussed in Chapter 2. Images (e) show the filtered images using the NLM algorithm. These images show that the noise is removed and the fine details are preserved in the image except in boundary areas.

Images (f) show the differences between images (d) and (e). They show that the noise is removed from all parts of the image which is not the same case when the NLM algorithm is used. The mses for the NLM algorithm are 13×10^{-4} for slice 3, 12×10^{-4} for slice 5, and 13×10^{-4} for slice 20. These numbers are higher than the EBA numbers.

Finally, the processing time for the NLM algorithm was about 30 s, which is the same as that of the EBA algorithm.

3.3 EBA Comparison with PCA:

Finally, we compared the EBA method to the PCA. Again the same DWI images in Section 3.1 were used. The results are shown in Figures 3.7, 3.8 and 3.9. The images in (a) and (d) show the DWI before denoising. Images (b) and (c) are the same ones shown in the previous Section. The denoised image using PCA is shown in (e). It does not show the details in the middle as that of the EBA algorithm in image (b). PCA acts like a low pass filter in some areas of the image.

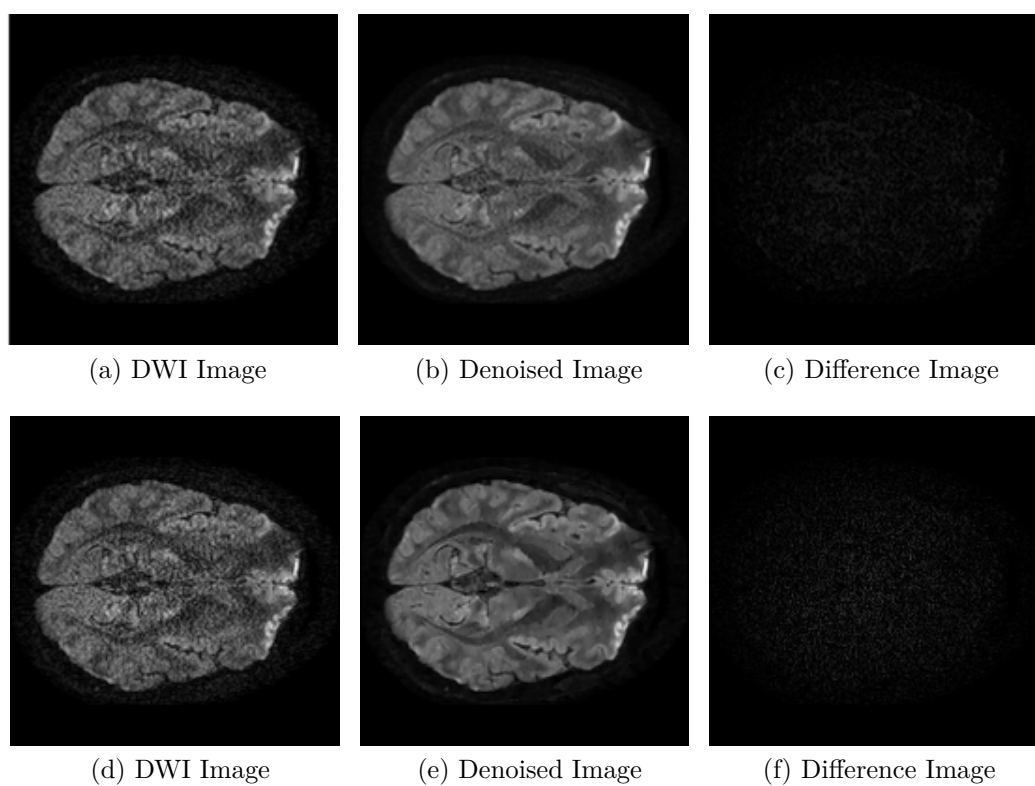


Figure 3.7: EBA ((b) and (c)) and PCA ((e) and (f)) Comparison, Slice 3.

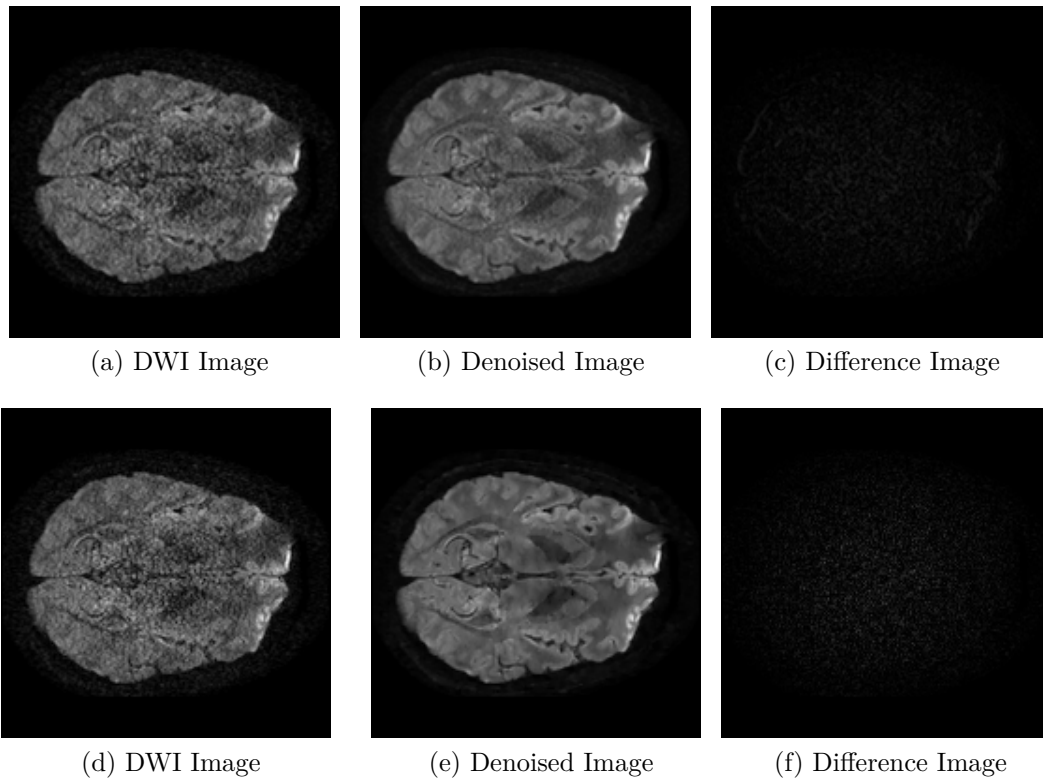


Figure 3.8: EBA ((b) and (c)) and PCA ((e) and (f)) Comparison, Slice 5.

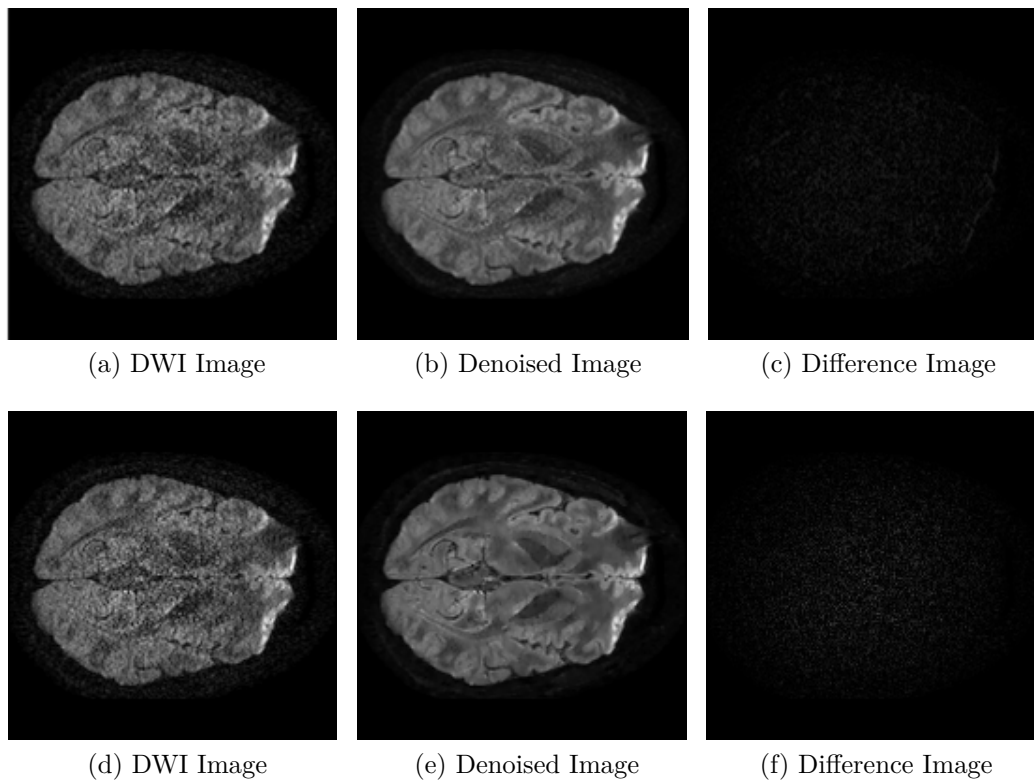


Figure 3.9: EBA ((b) and (c)) and PCA ((e) and (f)) Comparison, Slice 20.

The differences between images (d) and (e) are shown in images (f). It appears that the noise is removed from the image without affecting its details. The mses for all the three slices are 13×10^{-4} , which are higher than that of EBA algorithm. The processing time for this algorithm was about 30 sec which is the same as the processing time for the EBA algorithm.

3.4 Comparison Summary:

Here, we summarize our comparisons of the EBA algorithm to three different denoising methods in Table 3.1.

Method	Mean Squared Error			Processing Time
	Slice 3	Slice 5	Slice 20	
EBA	8.58×10^{-4}	7.39×10^{-4}	7.325×10^{-4}	~30 Sec
DCT	9.06×10^{-4}	9.63×10^{-4}	9.26×10^{-4}	~120 Sec
NLM	13×10^{-4}	12×10^{-4}	13×10^{-4}	~30 Sec
PCA	13×10^{-4}	13×10^{-4}	13×10^{-4}	~30 Sec

Table 3.1: Comparison Results.

The processing time of the EBA is about the same as that of the DCT and PCA. However, the processing time for the DCT algorithm is 4 times more than those of the other methods. The mse of the EBA algorithm is the least among all the denoising methods. The EBA and the PCA perform better in keeping the fine details of the image, the NLM comes next and the DCT is the last. We can conclude that the EBA method is very competitive with the most predominant denoising methods for DWI.

Chapter 4

Simulation and Experimental Results

In this chapter, we present more intensive testing results for EBA. The first part shows the results of simulation experiments in which noise is added to a clean image and then the clean image is used as a prior. Part two shows the results of phantom images were acquired at a local GE facility. The images are acquired using GE standard Head RF coil and GE 3.0T 750W GEM, 32ch MRI system. The scan parameters are as follows: Matrix Size= 256x256, TE=20msec, TR=500msec, NEX=1, BW=15.16KHz, Slice Thickness=10mm, Number of Slices:1, Slice Spacing: 0, Slice Orientation: Axial, sagittal, and coronal. FOV: 240mm, PSD: DWI-Propeller.

Two different phantoms are used to acquire the images; GE large sphere NiCl phantom and GE standard Head coil NiCL cylindrical phantom with RF coil loader. Prior to scanning, the phantoms are placed inside the GE Head coil and slice location is land-marked at the center of the phantoms. DICOM and raw data are saved for post-processing. A proprietary GE software program is used to convert the DICOM and raw data to a readable matrix for processing in Matlab.

Part three shows the results using in-vivo anatomical DWI of the brain

with Propeller acquisition. The images are obtained using GE standard DWI-Propeller protocol. This chapter also addresses the noise distribution of MRI images.

4.1 Simulation Results

This Section shows the results of the simulation work. Gaussian noise with variance of .025 is added to a clean image, and the clean image is used as the prior image. The EBA algorithm is tested using different Beta values. The results are compared visually for cleanliness with the SNR as a measure to quantify the image quality. The SNR is computed by selecting the most homogenous area in the image and is defined as the ratio of the mean and the standard deviation of pixel value in the selected area.

Figure 4.1 (a) shows a 128×128 pixels clean image that is used as a b_0 image. Gaussian noise is added to (a) to create DWI with low $SNR = 9.41$. The difference of the prior image and the DWI is computed and its is shown in (c). The histogram shows that most of the corresponding pixels in the two images have similar brightness. This is a good example in which the EBA algorithm will perform under ideal situation.

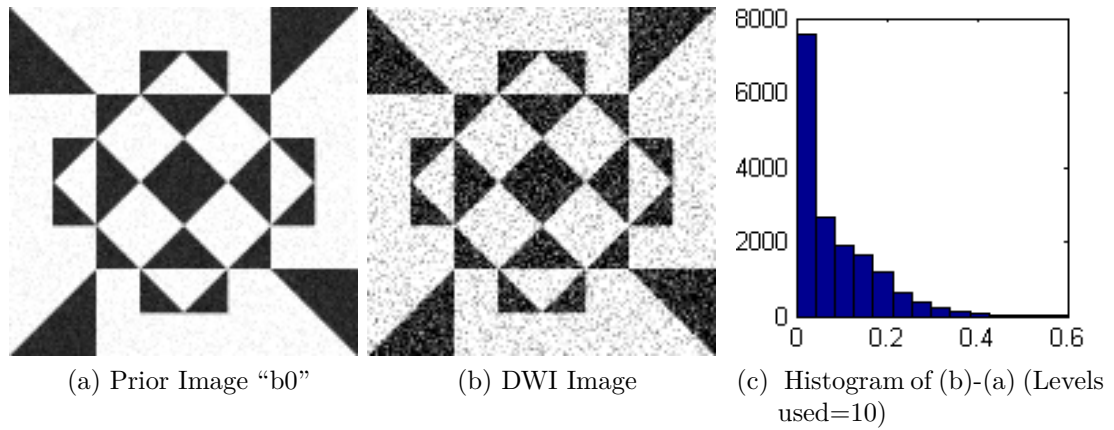


Figure 4.1: Simulation Icon Images.

Figure 4.2 shows the result of the reconstructed images for different beta values. As beta increases, the reconstructed image becomes cleaner since more weight is put on the prior image.

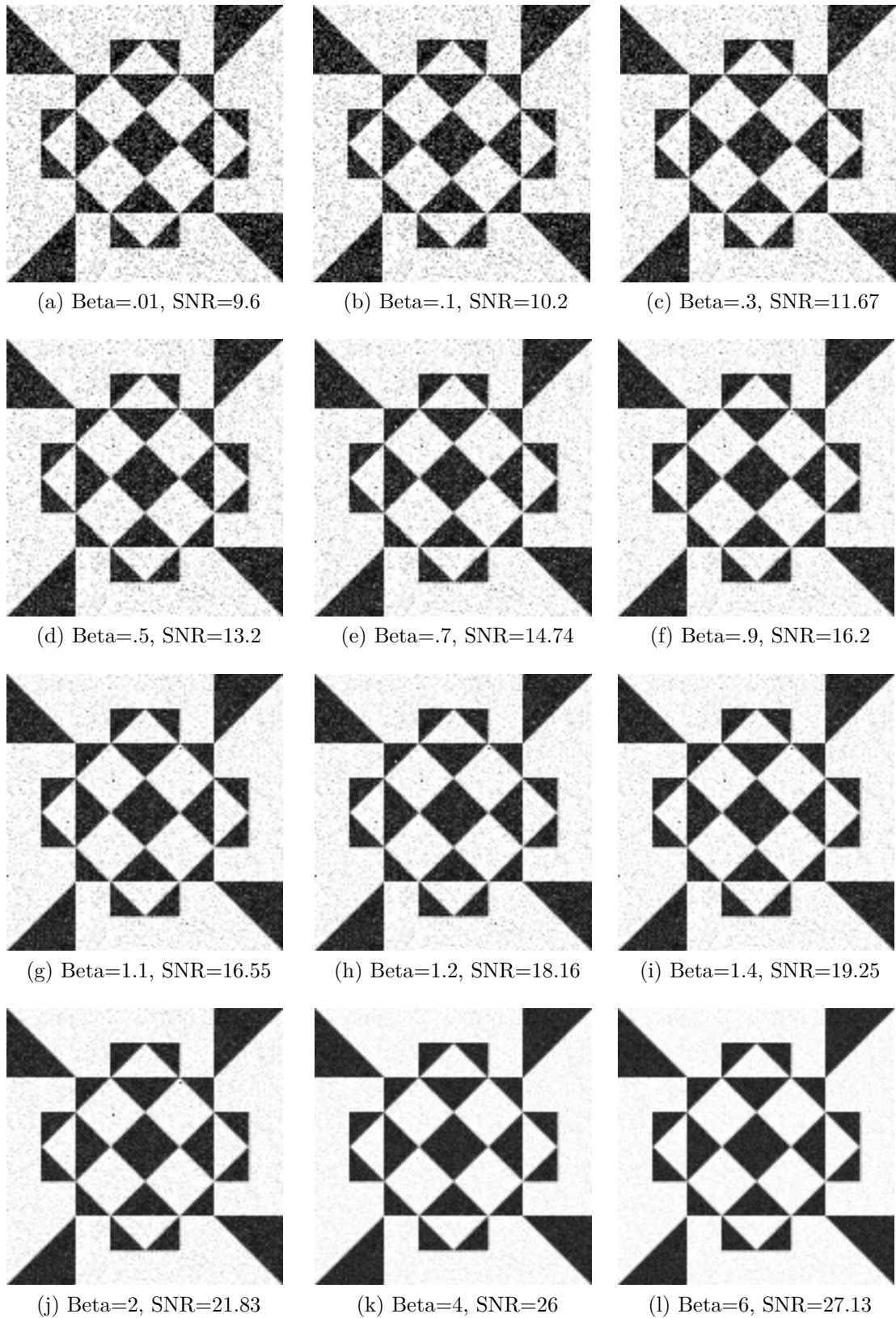


Figure 4.2: Icon Simulation Results.

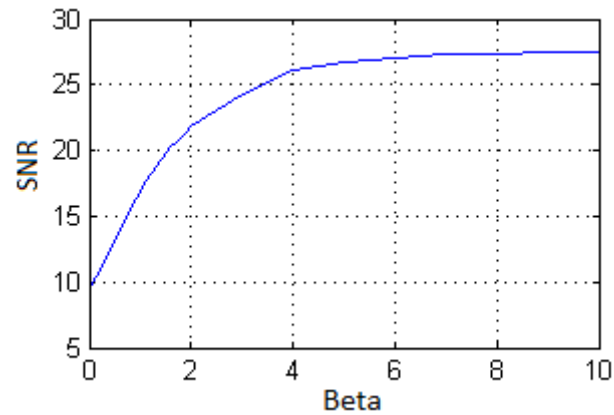


Figure 4.3: Icon SNR vs Beta.

Figure 4.3 shows the SNR versus beta. It shows that the SNR increases as beta increases. When beta=0.3, the SNR increases to 24% above its original value of 9.41.

We have seen in this Section an example in which the proposed method improved the quality of the reconstructed images. The histogram of the difference between the DWI image and the prior is used as a measure of similarity. It is found using Matlab that 83.64% of the pixels are less than 20% difference in the intensity.

4.2 Experimental Results

The second part of the experiment was testing the algorithm on real diffusion weighted images. Two phantoms were used at GE to collect two sets of images. The first data set was collected using the GE standard Head coil NiCL cylindrical phantom “logo”, we used 4 different b values 50, 100, 200, and 400. The second data set was using the large sphere NiCl phantom. We used 5 different b-values 500, 1500, 2500, 5000, and 10000. The purpose of this part was to study the noise distribution on actual MRI images and to test how better the reconstructed images would be.

4.2.1 Noise Study

It is usually assumed the noise in MRI has a Gaussian distribution, however it has been shown by [8] that the noise of the magnitude data has a Rician distribution for small SNR (≤ 3). As the SNR increases then the noise can be approximated by Gaussian distribution. In this part, we tried to study the distribution of the noise for different b-values based on the noise histogram analysis.

According to [19] the noise distribution will be Gaussian distribution if the SNR > 3 . We find that if the b-value is 1000 or less, then the SNR >3 , as it will be shown in the next two subSections.

Figure 4.4 shows the noise data of the logo phantom. The b-values were 50, 100, 200, and 400.

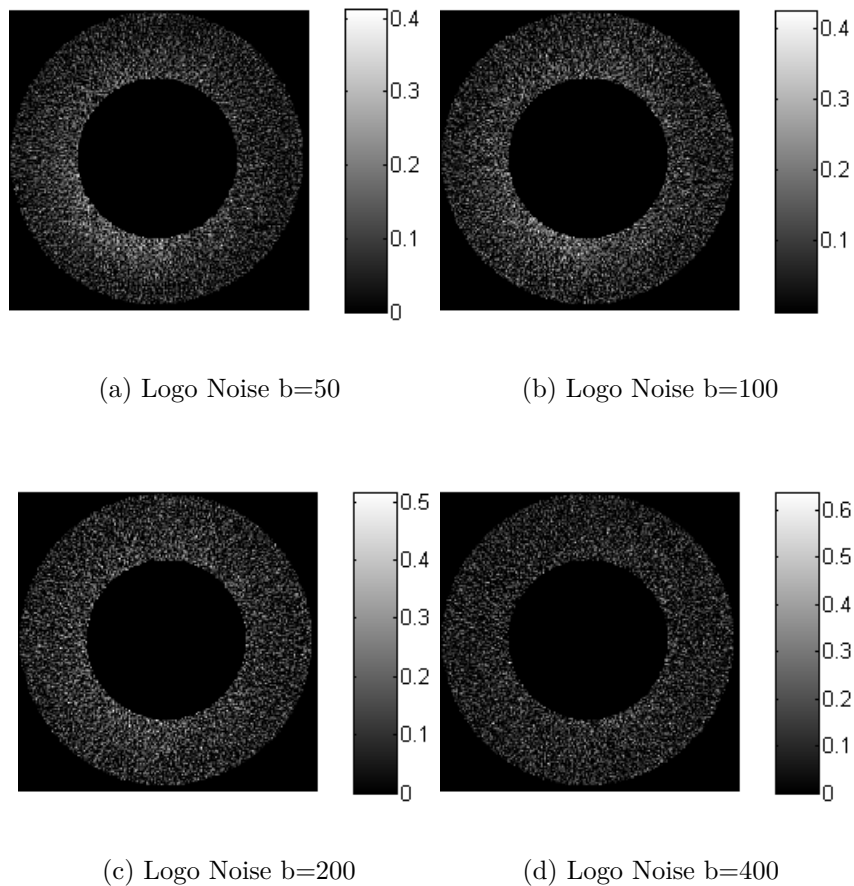


Figure 4.4: Logo (Cylinder) Noise Data for DWI Images with Various b -values.

The noise data of each DWI image is used to form a histogram. We use different quantization levels to create the histograms. The number of quantization levels is chosen manually. Figure 4.5 shows the histogram of each data set. The DWI images are magnitude images which have positive values. Data with negative values are added to the original data in order to form a symmetric Gaussian distribution. Matlab is used to estimate the mean and variance of each data set with the maximum likelihood estimator. The estimated mean and variance, are used to fit a Gaussian probability distribution for each noise histogram. Figure 4.5 shows that in four cases with different b -values, all with $b < 1000$, a Gaussian distribution fits the data correctly as the SNR is larger than 3 in all cases.

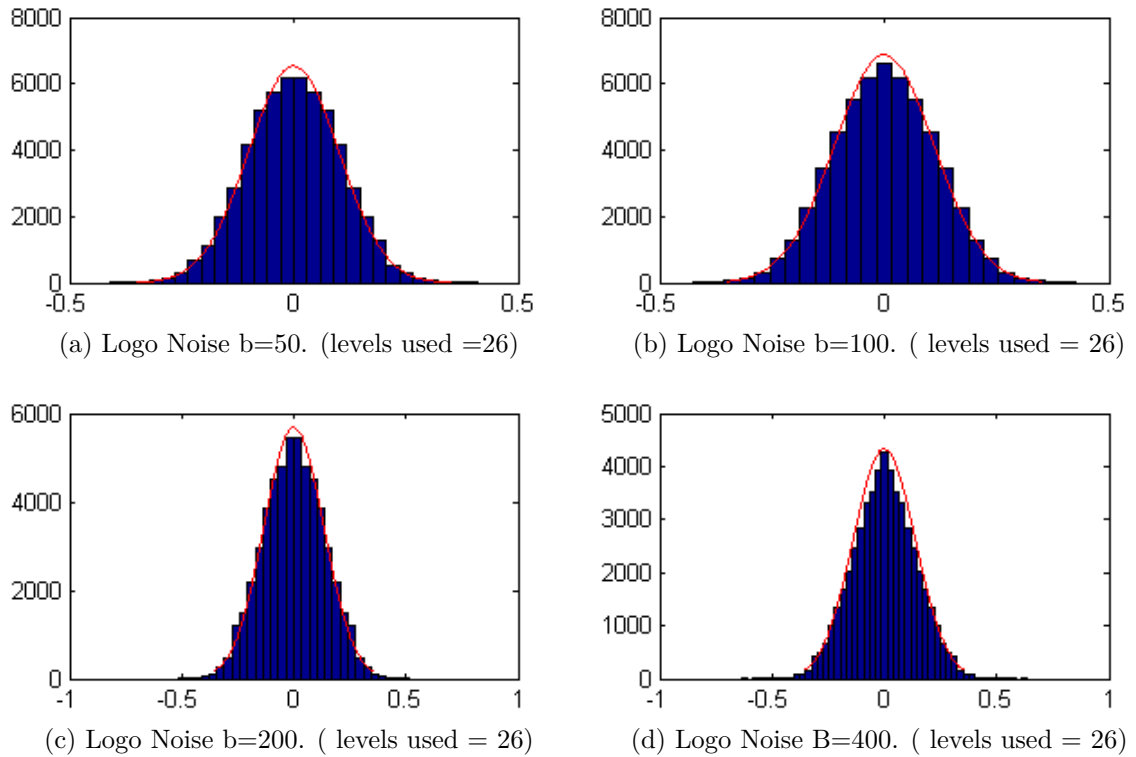


Figure 4.5: Logo Noise Histograms.

The second phantom is a sphere. Five sets of DWI images have been collected with various b -values (500, 1500, 2500, 5000, 10000). Figure 4.6 shows the noise data of each image.

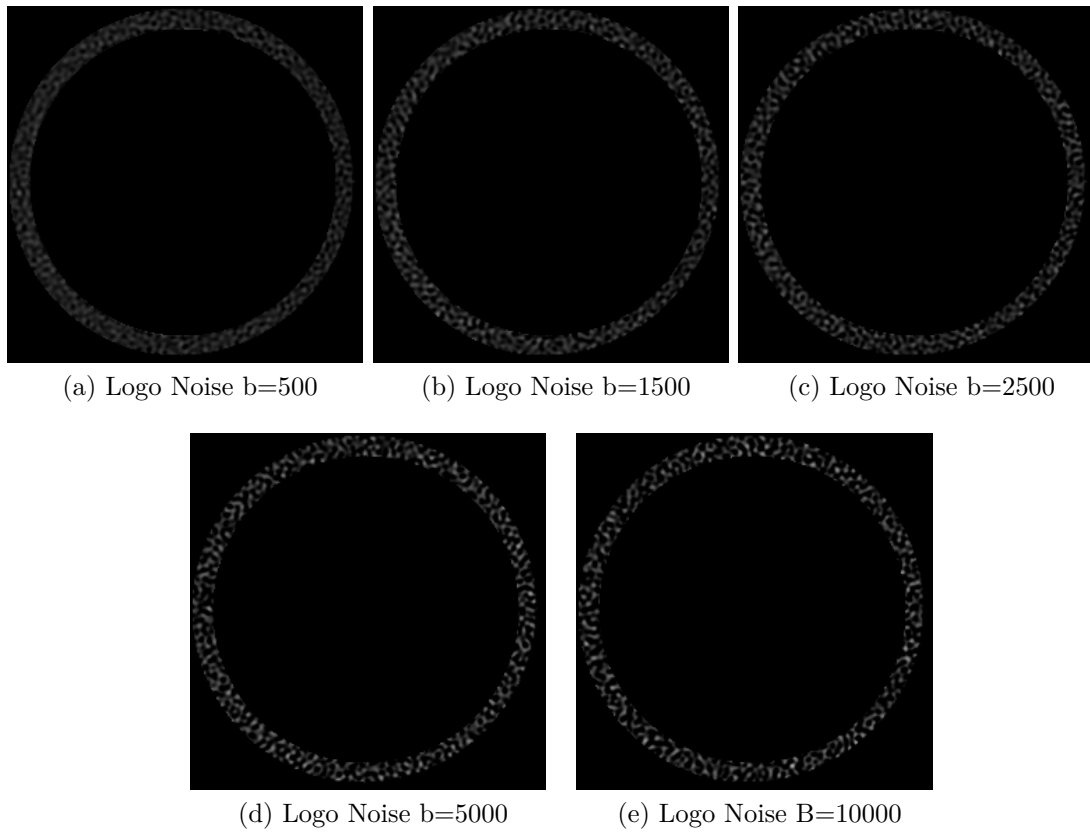


Figure 4.6: Spherical Phantoms Noise Data for DWI Images with Various b -values.

The noise data from each image are used to form a histogram. Figure 4.7. shows the histograms of the data for $b=500$ and $b=1500$ which have Gaussian distributions. In these two cases, the SNR was higher than 3 which made the Gaussian distribution a good approximation for the Rician distribution.

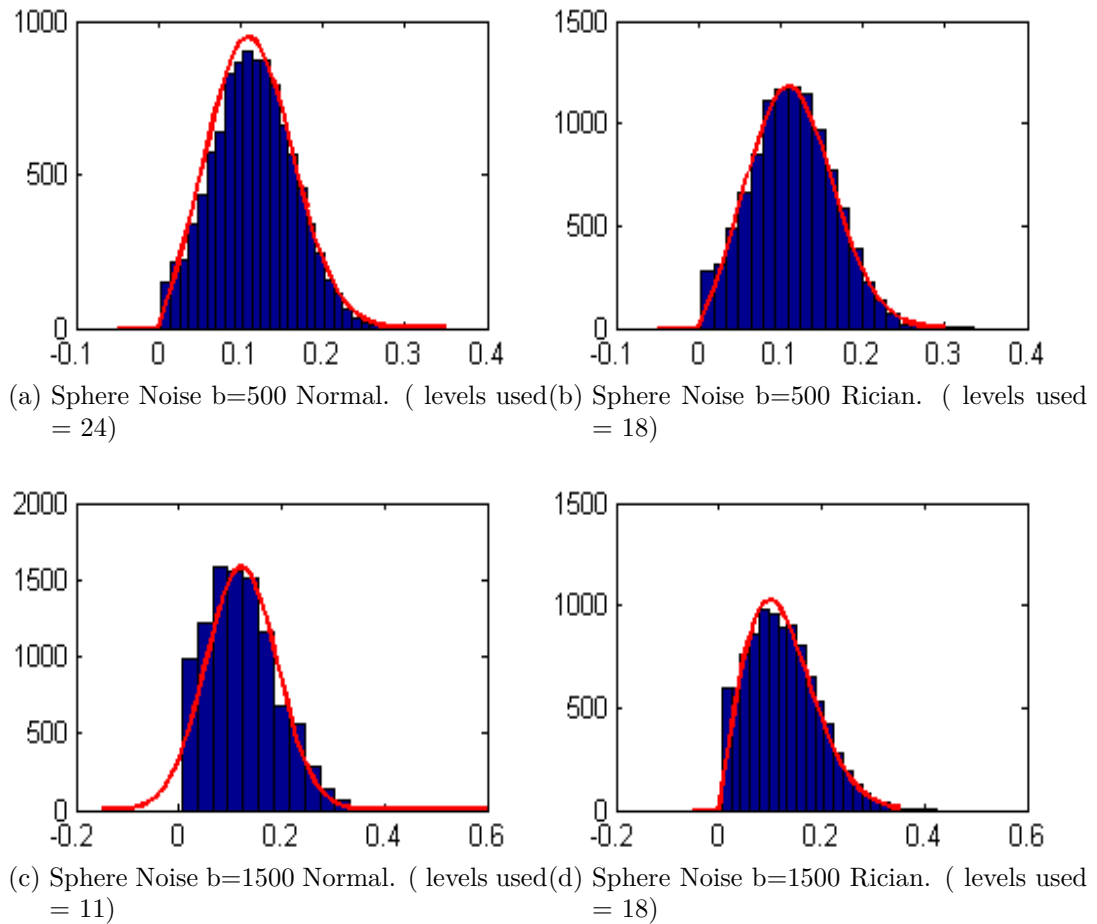


Figure 4.7: Spherical Phantom Noise Histograms.

Figure 4.8 shows the results obtained for $b=2500, 5000, 10000$. Two distributions are fitted to each histogram. On the left side of the figure we fit data with the Gaussian distribution. On the right side of the figure, we fit data with Rician distribution. It is clear from these histograms that the noise distribution follows Rician distribution for $b \geq 2500$ as the SNR is less than 3.

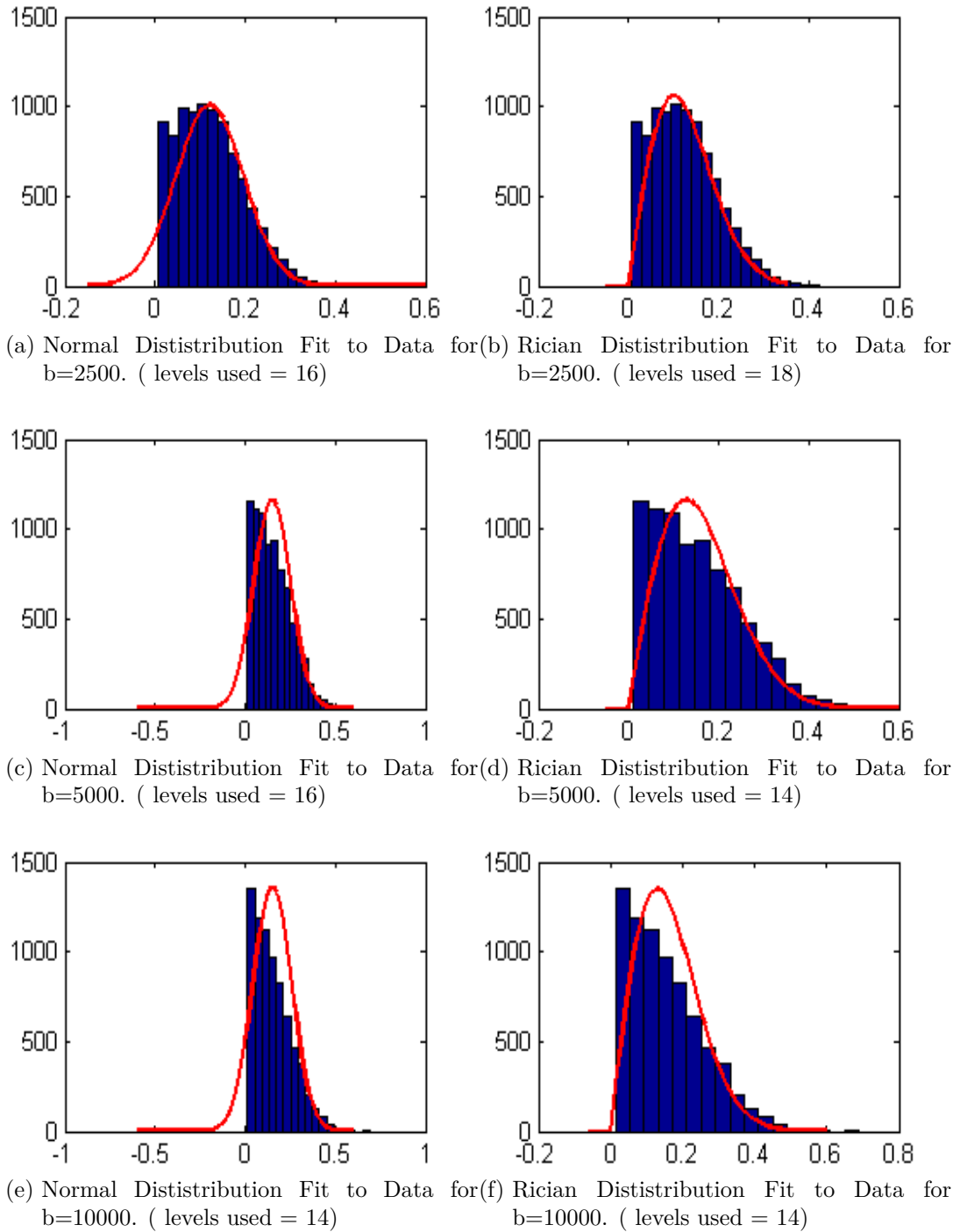


Figure 4.8: Spherical Phantoms Noise Histograms.

4.2.2 Reconstruction Images

4.2.2.1 Sphere Phantom

This Section shows the results of different experiments that have been done using a sphere phantom. The purpose of the study is to validate the algorithm and to study the noise distribution related to the b-value. The similarity between the DWI image and the prior has been also computed using the absolute difference between the two images and forming the histogram of the difference. The spread of the histogram has been used as a threshold for the success of the EBA reconstruction algorithm.

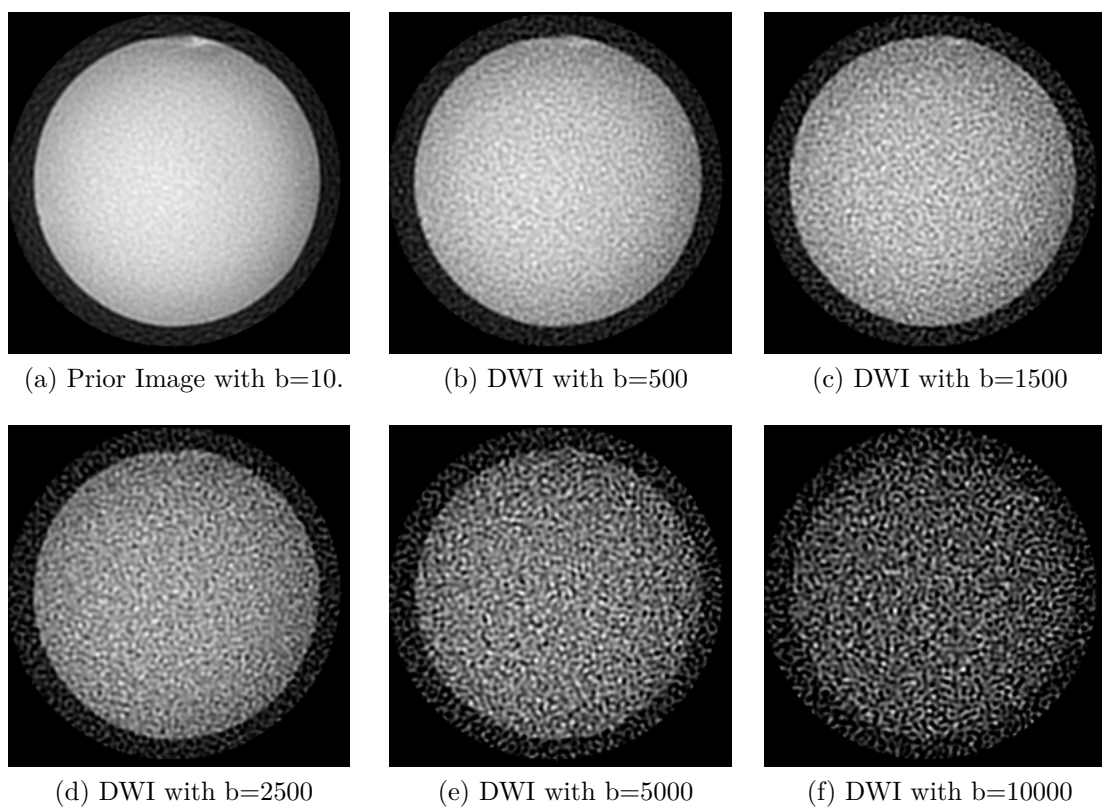


Figure 4.9: Sphere Logo Images.

Figure 4.9 shows the images of the sphere logo collected with different b-values.

4.2.2.2 Sphere Phantom for $b=500$

The first real DWI in this study is for a sphere phantom. We collect different images with different b -values. The first one has $b = 500$. To compare the DWI with its prior, the histogram of the difference between the two images is computed.

Figure 4.10(a),(b) and (c) shows the b_0 or prior image, the DWI image and the histogram of the difference between the b_0 and the DWI respectively.

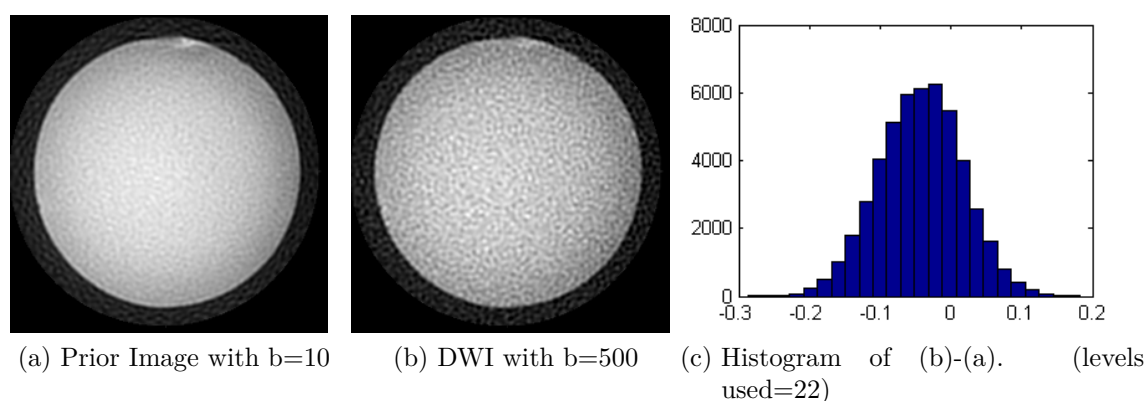


Figure 4.10: Sphere Logo b500.

The histogram shows the difference between the DWI and the prior image is small with most pixels within 20% range of difference. The noise histogram also have the shape of a Gaussian distribution for $b=500$.

The noise of the DWI image is obtained by masking the phantom on the image. The rest of the image will be the back-ground noise. The noise was examined to determine its distribution. In Figure 4.11 we show the noise of the image in (a) and its histogram in (c). The data of the noise was used to estimate the mean and the standard deviation of a Gaussian function. These values were used to fit a Gaussian function to the histogram in (c).

Figure 4.11 (b), (d) shows the image SNR and its histogram. In (b), the image SNR has high values (>10) which imply the Gaussian distribution for

noise. Clearly (d) shows that most of the data pixels have SNR greater than 10 which again infers Gaussian noise, further confirmed by the noise histogram in (c).

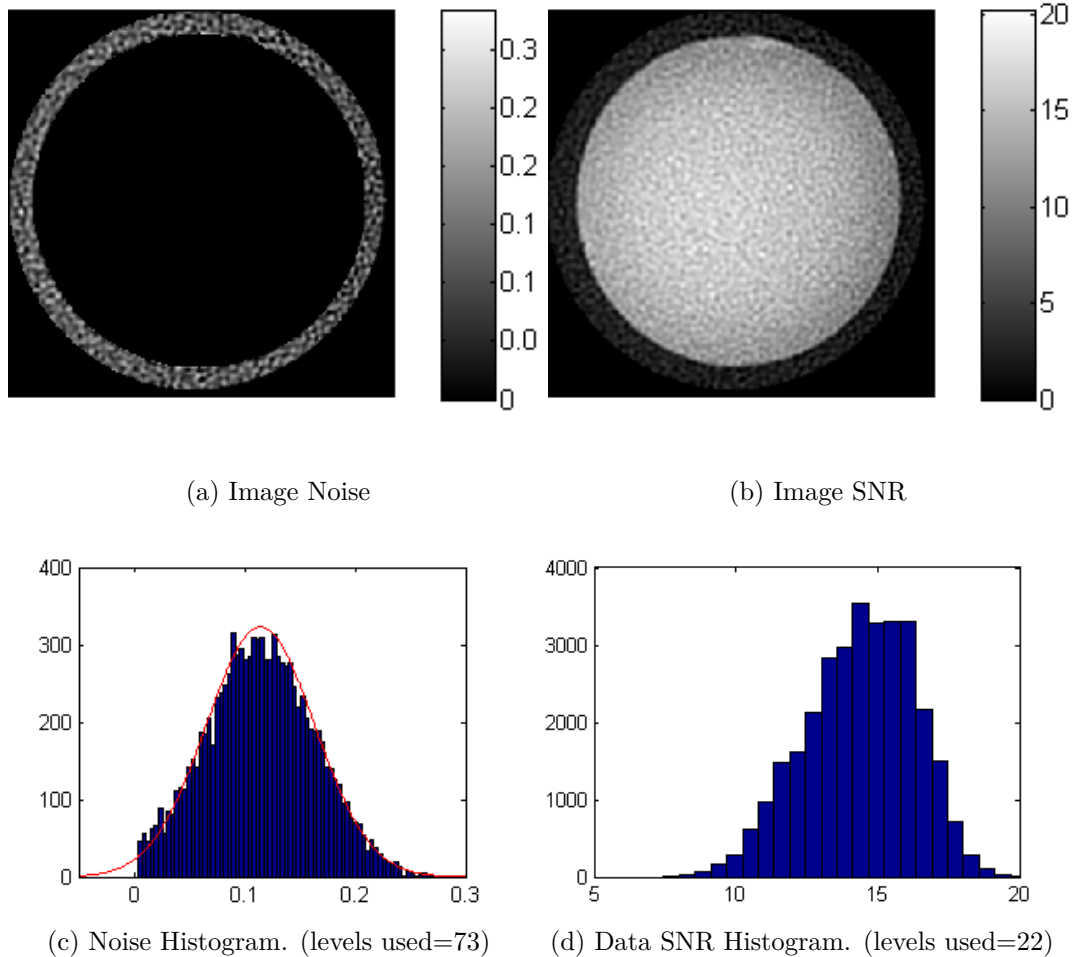


Figure 4.11: $b=500$ Noise Study

The EBA algorithm is then used to reconstruct the DWI and to test the improvement in the quality of the images. Different b -values are chosen, and the SNR is computed for each b -value using the a circle in the center of the image. Figure 4.12 shows the SNR versus beta and demonstrates the improvement in the SNR. A typical beta value is 0.3 at which the SNR improved by 20% with respect to the DWI without applying the EBA algorithm.

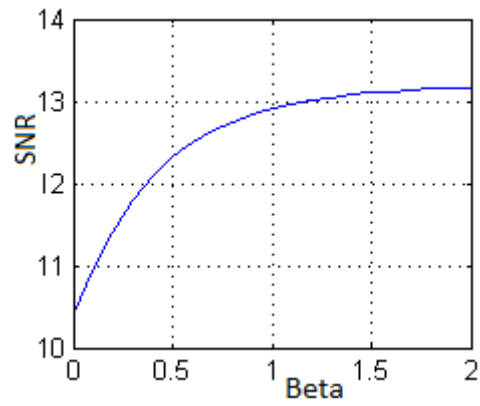


Figure 4.12: SNR vs Beta for $b=500$.

Finally, we show the reconstructed images in Figure 4.13. As beta increases, the images have less noise.

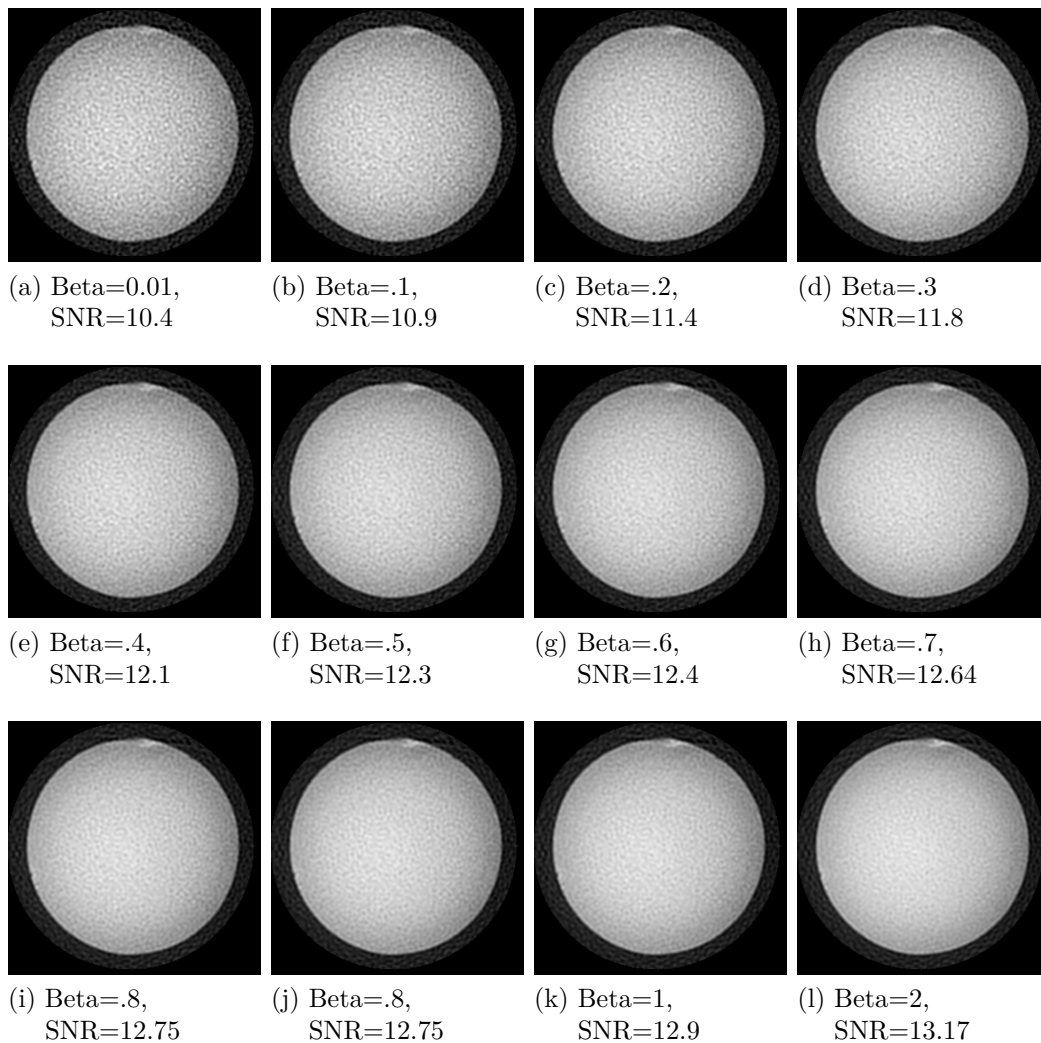


Figure 4.13: Spherical Phantom Image Reconstruction Results for b -value=500.

4.2.2.3 Sphere Phantom for $b=1500$

The second set of images for the spherical phantom is collected when $b=1500$. The DWI image and the prior are shown in Figure 4.14 (a) and (b) respectively. In 4.14 (c) the histogram of the difference between the DWI image and the b_0 image is shown. The histogram has the shape of a Gaussian function but it was more spread out than the previous case when $b=500$.

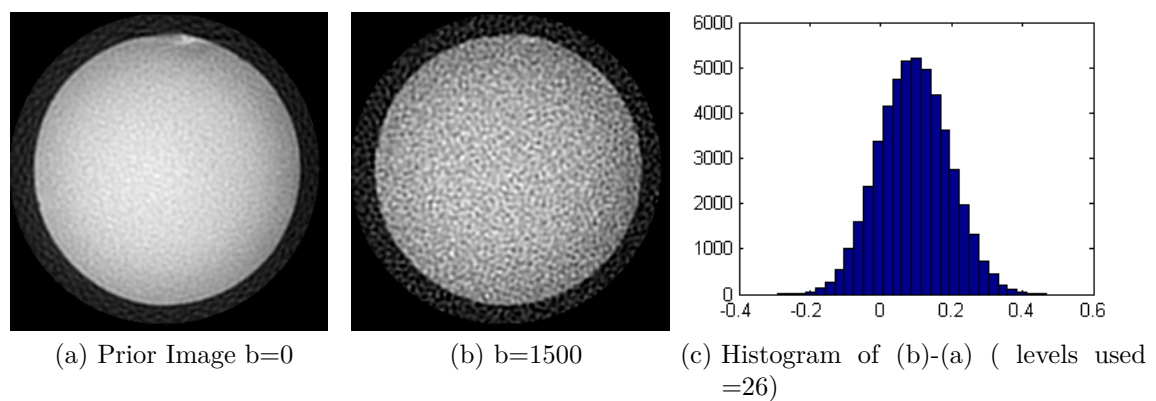


Figure 4.14: Spherical Phantom for $b=1500$.

Figure 4.15 shows the noise data in (a) and its histogram in (c). A Gaussian distribution is fit to the histogram showing that the noise distribution is close to Gaussian when $b=1500$. Part (b) of the graph shows the image SNR and part (d) shows the image SNR histogram. The SNR histogram shows that most of the image pixels in the phantom had SNR greater than 5 which confirms the validity of the Gaussian noise assumption.

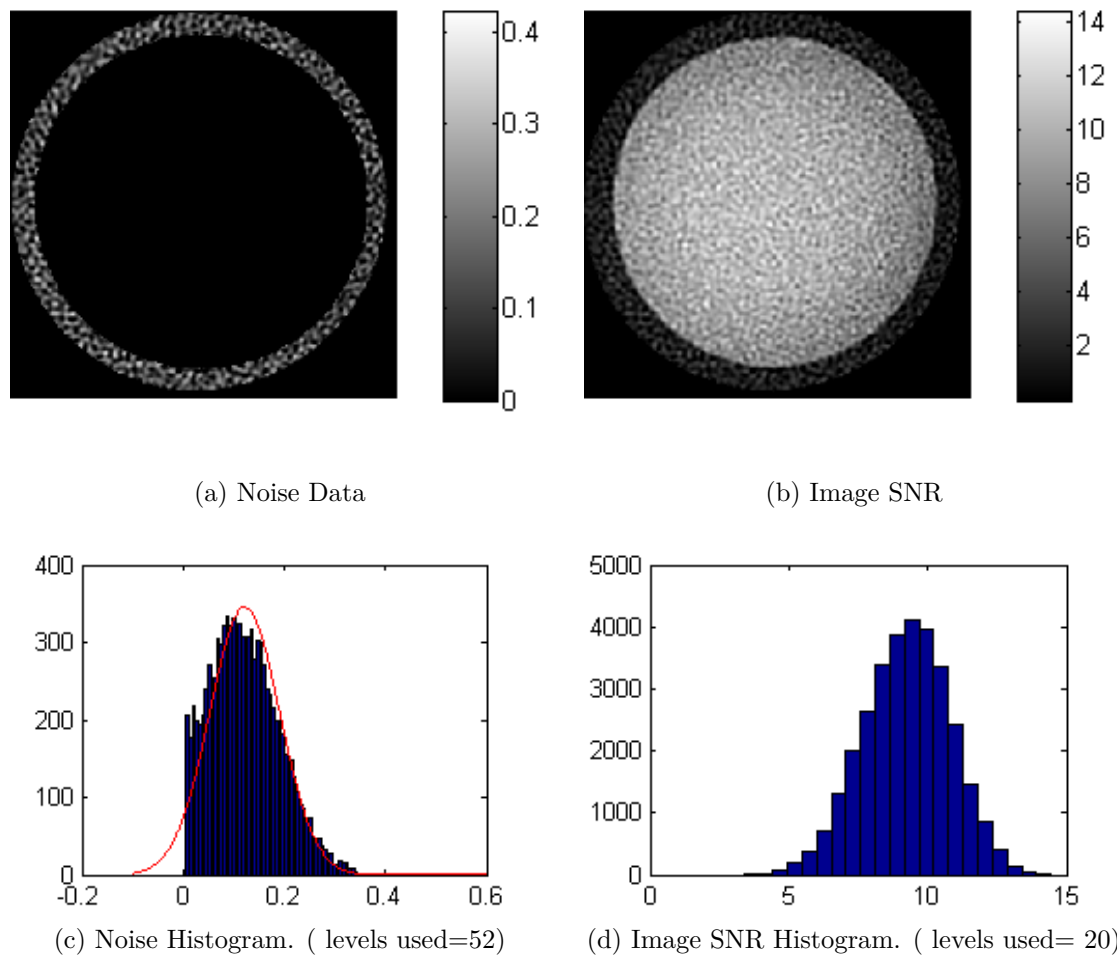


Figure 4.15: Noise Study for DWI with $b=1500$.

To verify the performance of the EBA algorithm, different numbers of beta between 0.01 and 2 are used, and for each beta, the SNR is computed. Figure 4.16 shows the SNR vs Beta . The graph shows improvement in the SNR. A typical value of Beta is 0.3 at which there was a 25% improvement in the SNR.

Figure 4.17 shows the results of the reconstructed images for different beta values. The images in Figure 4.17 become cleaner with less noise as beta is increased.

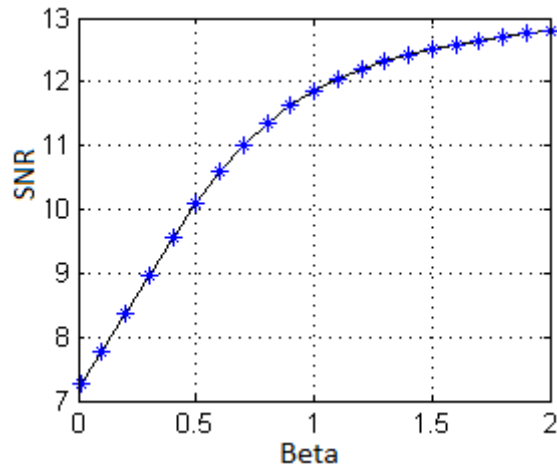
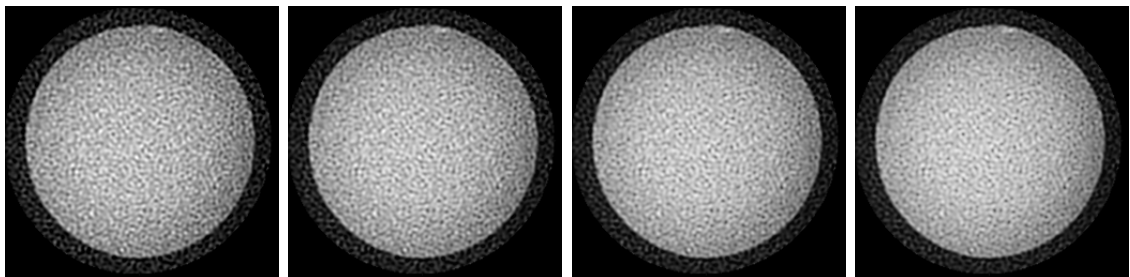
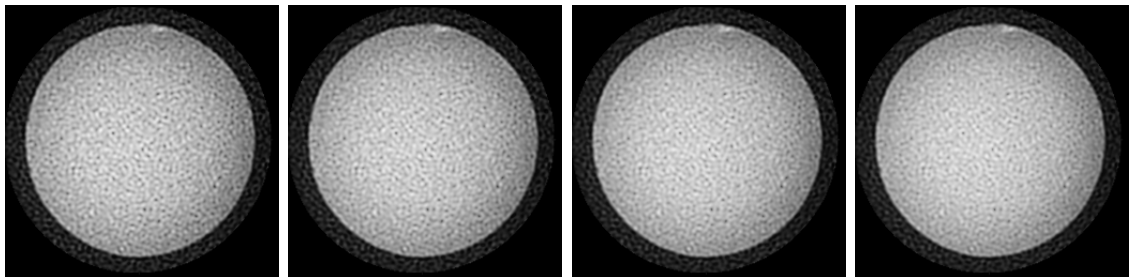


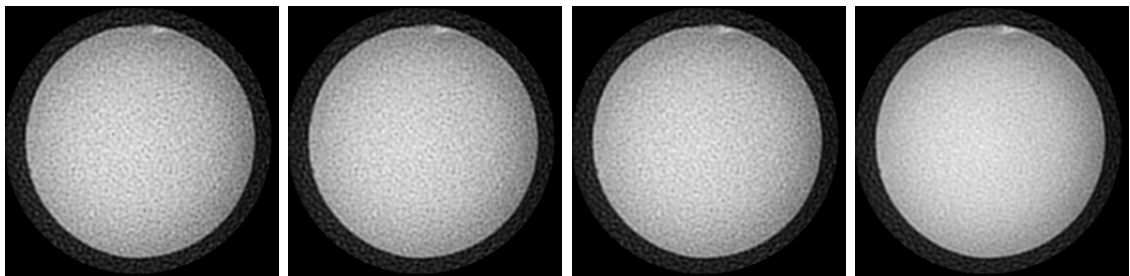
Figure 4.16: SNR vs Beta for $b=1500$.



(a) $\text{Beta}=0.01$, $\text{SNR}=7.25$ (b) $\text{Beta}=0.1$, $\text{SNR}=7.75$ (c) $\text{Beta}=0.2$, $\text{SNR}=8.35$ (d) $\text{Beta}=0.3$, $\text{SNR}=10.2$



(e) $\text{Beta}=0.4$, $\text{SNR}=9.55$ (f) $\text{Beta}=0.5$, $\text{SNR}=10.1$ (g) $\text{Beta}=0.6$, $\text{SNR}=10.6$ (h) $\text{Beta}=0.7$, $\text{SNR}=11$



(i) $\text{Beta}=0.8$, $\text{SNR}=11.35$ (j) $\text{Beta}=0.9$, $\text{SNR}=11.2$ (k) $\text{Beta}=1$, $\text{SNR}=11.85$ (l) $\text{Beta}=2$, $\text{SNR}=12.8$

Figure 4.17: Spherical Phantom Image Reconstruction Results for $b=1500$.

4.2.2.4 Sphere Phantom for $b=2500$

This Section shows the images when $b=2500$. The DWI image and the b_0 image are shown in Figure 4.18 (a) and (b) respectively. In Figure 4.18 (c) the histogram of the difference between the DWI image and b_0 is shown. It can be noticed that the histogram is more spread out than the previous cases and has many bins where the difference between the DWI image and the prior was greater than 20%. This would be an example where the algorithm would perform poorly since the prior image is not similar to the DWI.

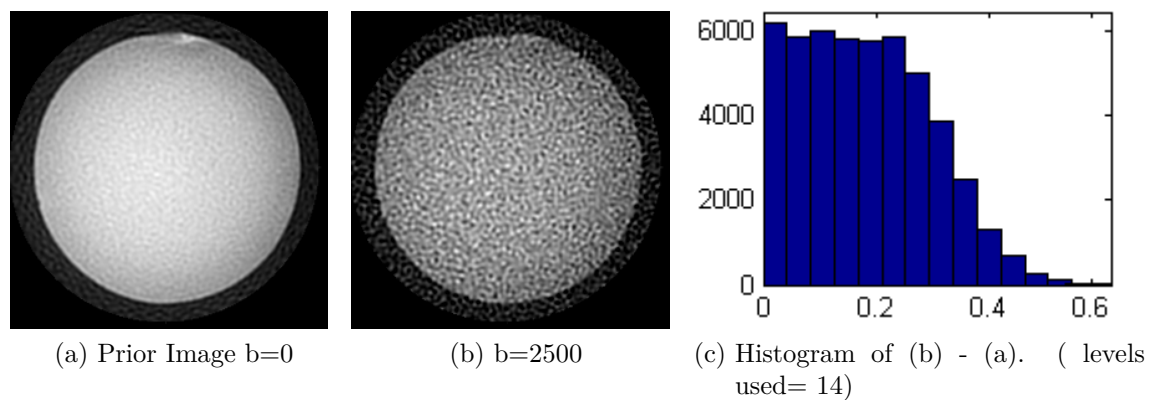


Figure 4.18: Spherical Phantom for $b=2500$.

Figure 4.19 (a), (c) shows the noise of the image and its histogram. In Figure 4.19 (b) the image SNR is shown and its SNR histogram is illustrated in Figure 4.19 (d) .

Figure 4.19 (d) proves that most of the image data has SNR greater than 5 with more pixels falling with SNR below 5 than those in previous cases. This suggests that the noise cannot be approximated by a Gaussian distribution for the pixels that had $\text{SNR} < 5$. If we compare the SNR histogram the ones in the previous cases, we see that the SNR decreases as the b -value increases.

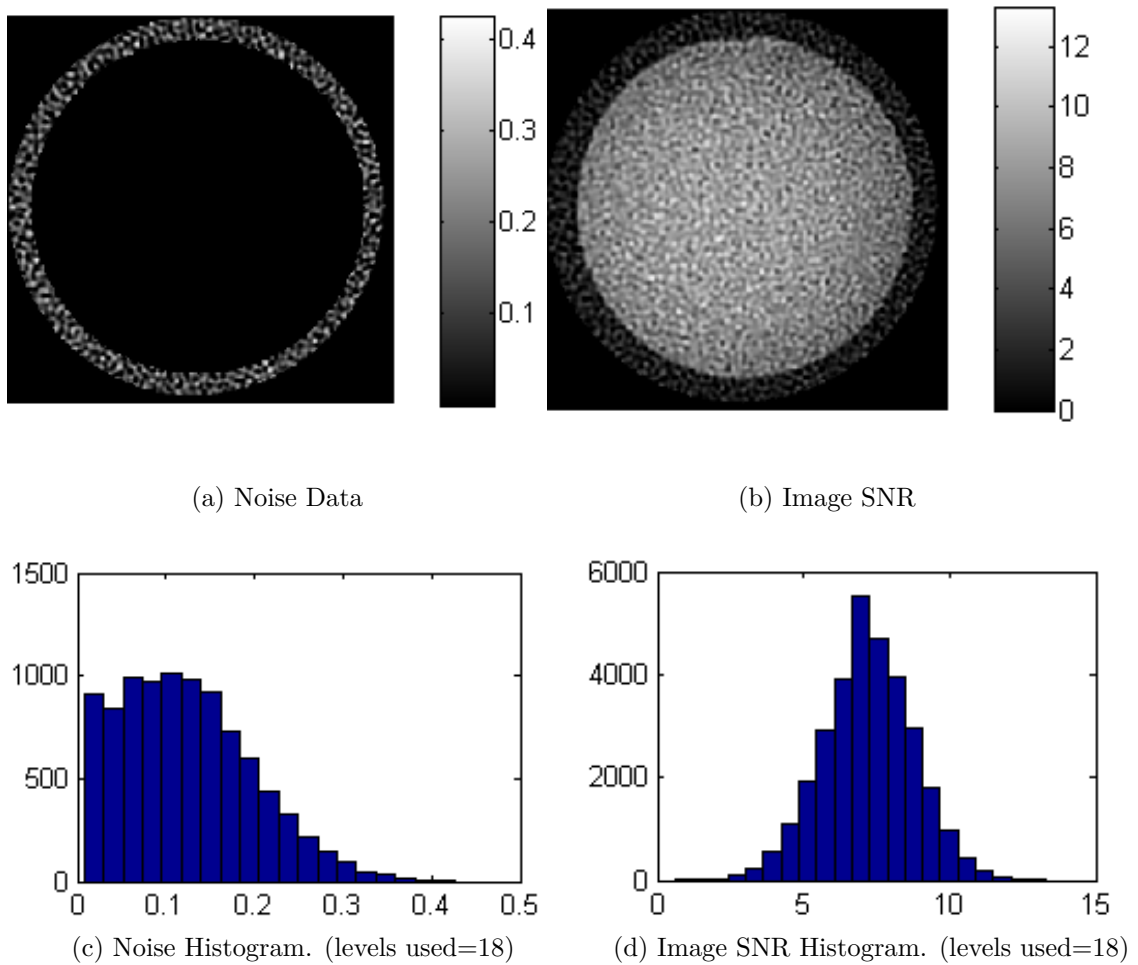


Figure 4.19: Noise Study for DWI with $b=2500$.

The DWI reconstruction for $b=2500$ using different values of β has been examined. Figure 4.20 shows the improvement in the SNR for different values of β . We can see that the SNR improved by a factor of 2 when $\beta=2$. However, $\beta=2$ is not a typical value. A typical value of β is 0.3 at which the SNR is 5.62 which is about 20% improvement.

Figure 4.21 shows the quality of the reconstructed images that has been improved as β is increased. However, there are some pixels on which the noise is dominant and the algorithm fails to improve the image. The algorithm have kept the same values of these pixels as it classifies it as true values of the DWI instead of noise since the difference between the corresponding pixels in the DWI

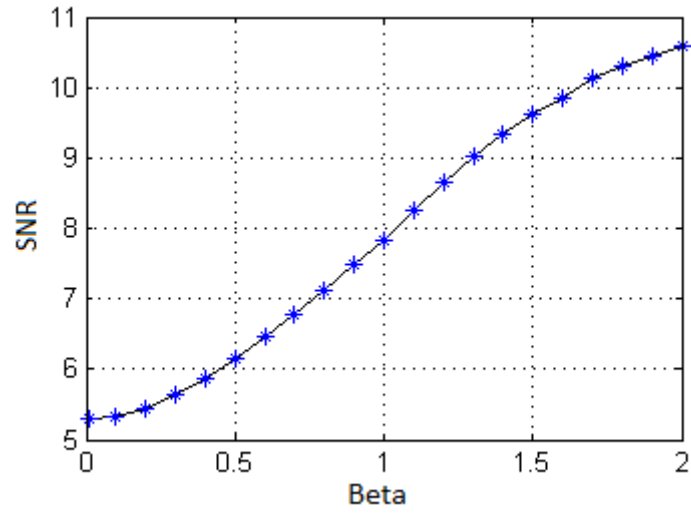


Figure 4.20: SNR vs Beta for $b=2500$.

image and the b_0 image are high.

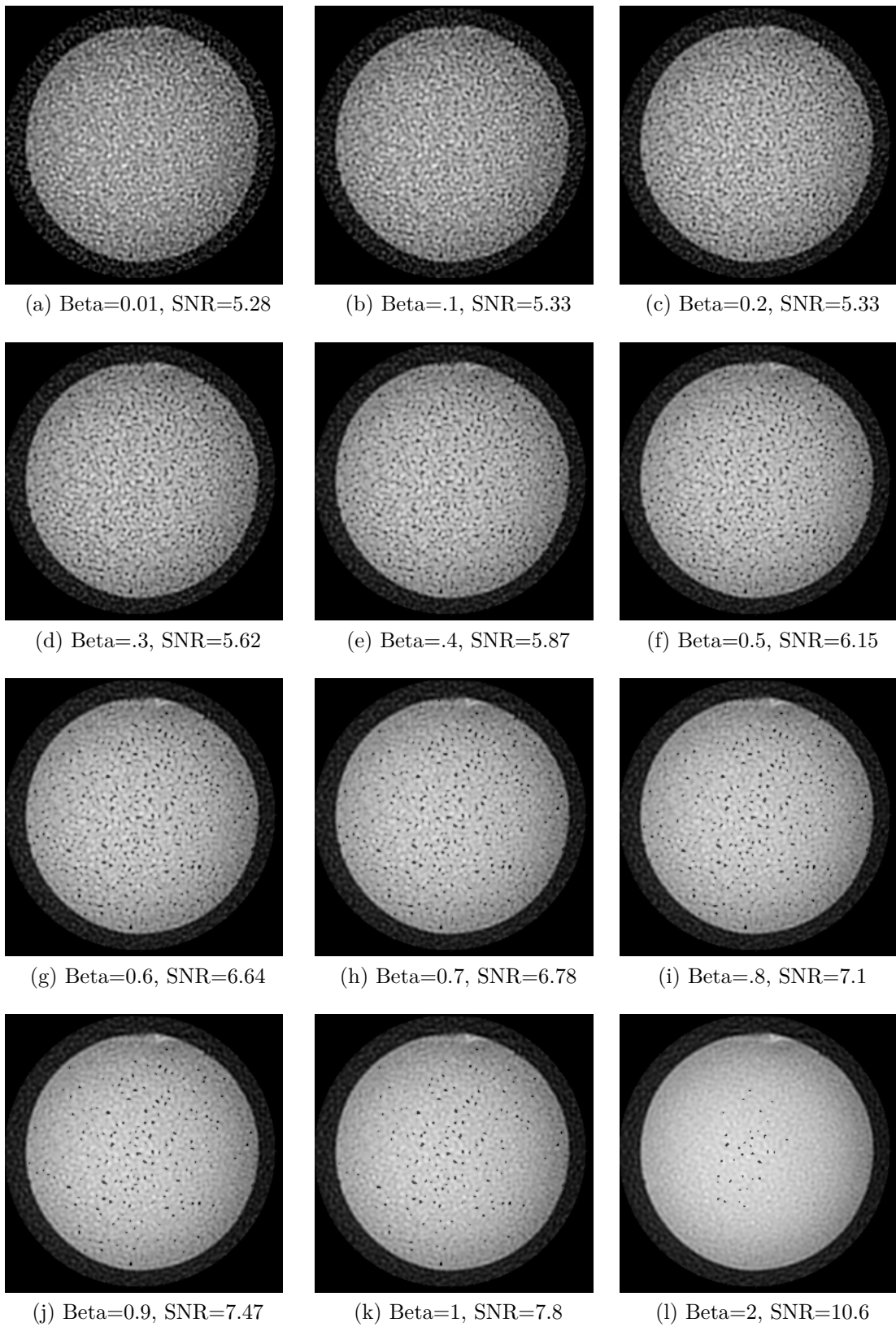


Figure 4.21: Spherical Phantom Image Reconstruction Results for $b=2500$.

4.2.2.5 Sphere Phantom for $b=5000$

In this Section, we show the results of the reconstructed images when $b = 5000$. Figure 4.22 (a) shows the prior image and 4.22 (b) shows the DWI. Figure 4.22 (c) shows the histogram of the difference between the DWI and the prior. The histogram is spread out which means a lot of the corresponding pixels in the two images are very different. This makes the b_0 image not the right prior and hence the algorithm will not improve the DWI image quality significantly.

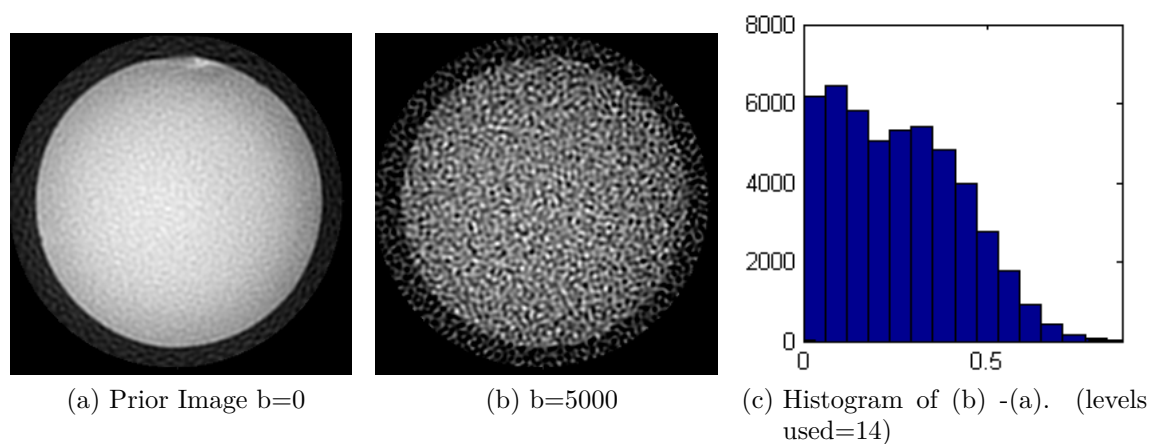


Figure 4.22: Spherical Phantom for DWI with $b=5000$.

Figure 4.23 (a) and (c) show the noise of the DWI image and its histogram. The histogram is spread with many pixels having high-intensity values that are larger than the true value of the pixel, i.e. the noise value is much higher than the pixel value which makes it hard to properly estimate the image in this case. In Figure 4.23 (b) and (d) the image SNR and its histogram are shown. A large number of pixels have SNR value less than 5 which indicates invalidity of the Gaussian noise approximation.

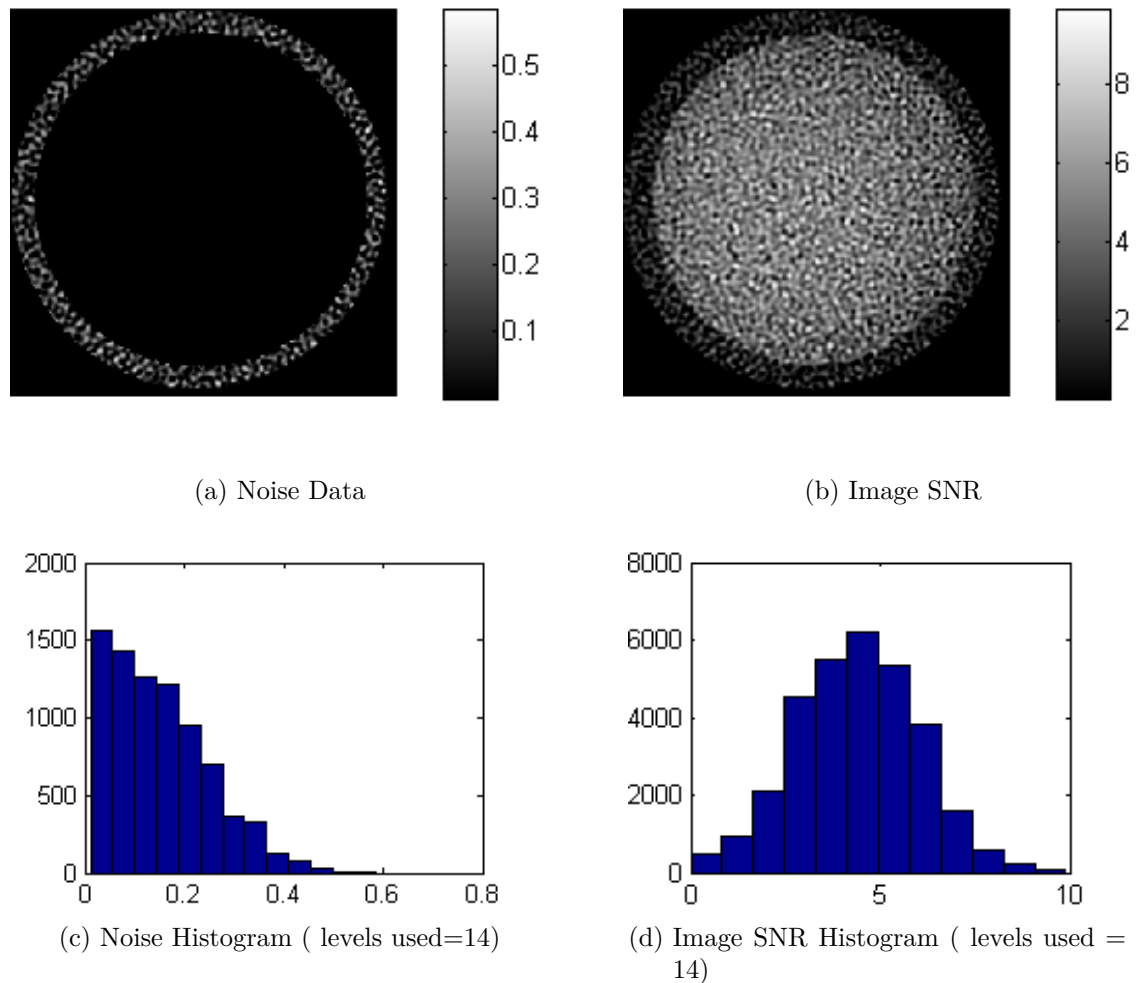


Figure 4.23: Noise Study for DWI with $b=5000$.

Figure 4.24 shows a plot of the SNR vs Beta. It shows that the SNR swings down and then up again owing to the improper use of the prior. For $.01 \leq \text{beta} \leq .8$, the SNR is below the reconstructed DWI without prior. For $.9 \leq \text{beta} \leq 1.5$, the SNR increases but it is still below the original DWI SNR=3. As beta increases, the prior image will have dominant effect on the reconstructed image. For large beta, when the difference between the corresponding pixel of the DWI image and the prior is large, the algorithm will choose the value from the prior image. This can be seen in Figure 4.25 which shows the image quality has not improved until beta becomes high such that the prior becomes the DWI image itself.

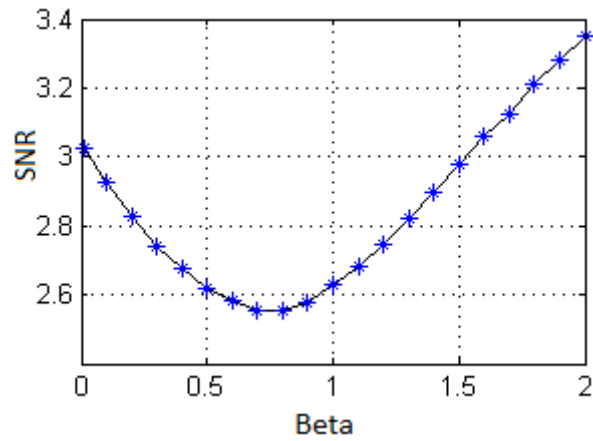


Figure 4.24: SNR vs Beta for DWI with $b=5000$.

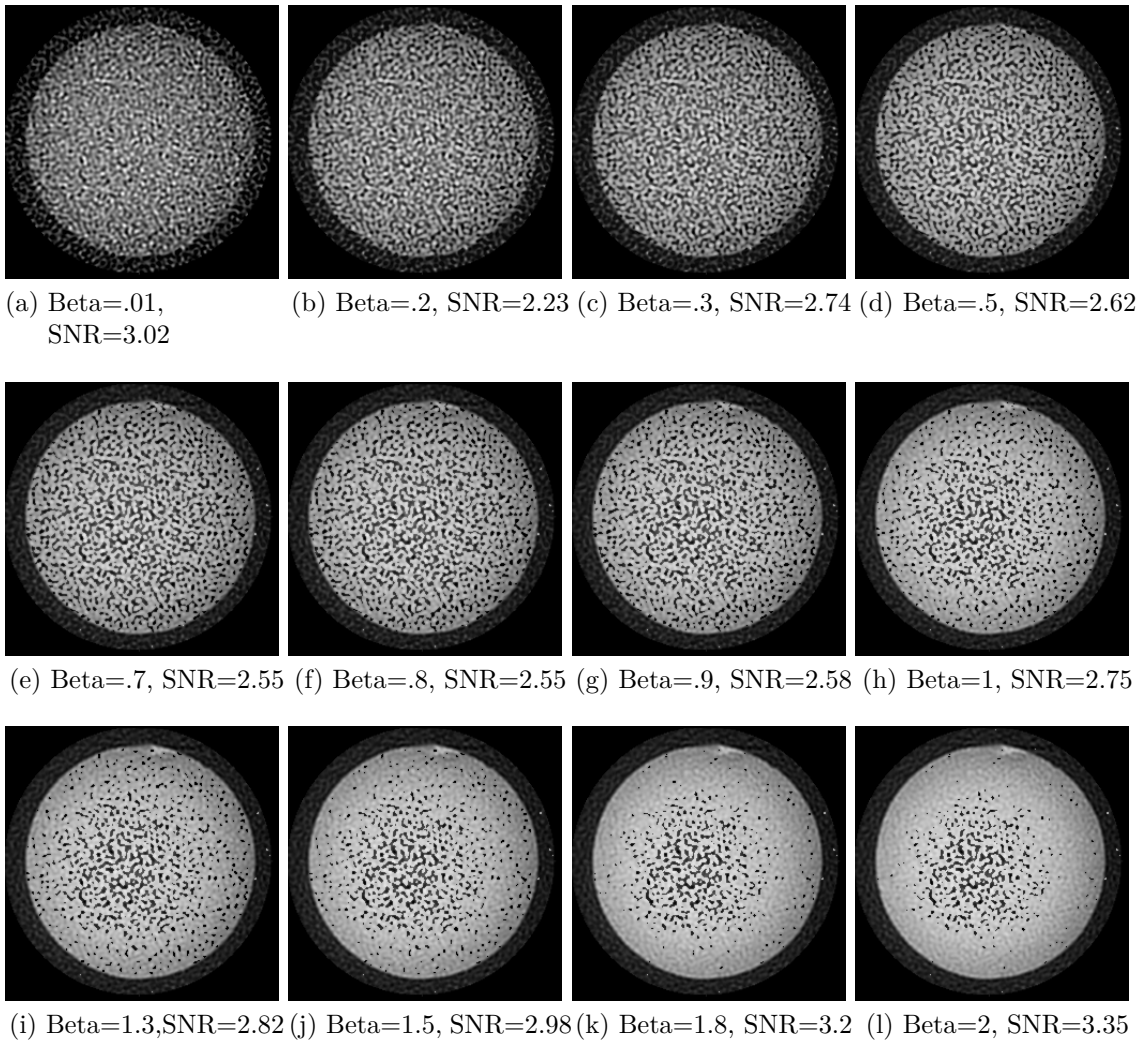


Figure 4.25: Spherical Phantom Image Reconstruction Results for DWI with $b=5000$.

4.2.2.6 Sphere Phantom for $b=10000$

The last set of the sphere phantom images was when $b = 10000$. Figure 4.26 (a) shows the prior image and 4.26 (b) shows the DWI. Figure 4.26 c shows the histogram of the difference between the DWI and the prior image. The histogram shows the high variation in the intensity of the pixels of the DWI and the corresponding pixels of the prior. Here, the prior has no definite relationship with the image and for this reason the EBA algorithm will not work. We can also notice that the histogram has two populations, the first one in the 0 to .3 range, which is coming from the noise in the two images, and the second population takes a Gaussian shape and is coming from the sphere data itself.

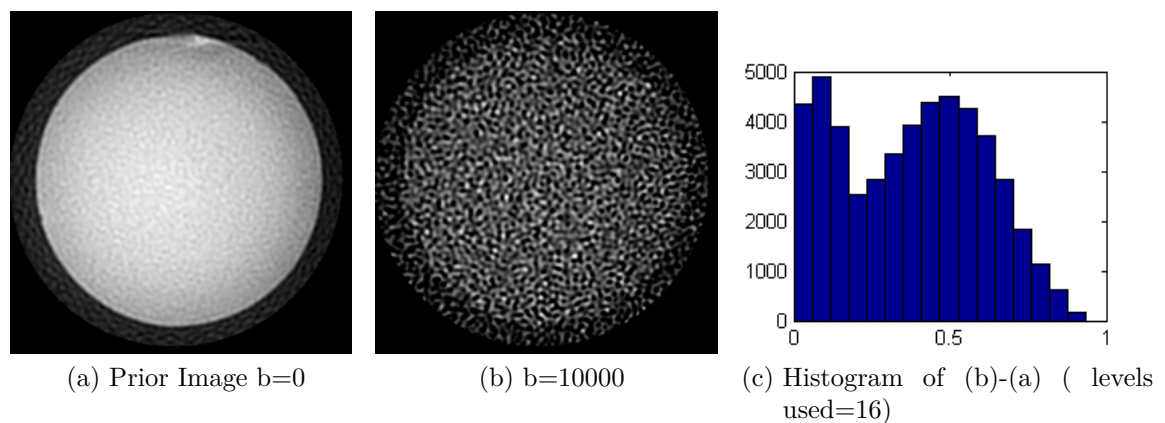


Figure 4.26: Spherical Phantom $b=10000$.

Figure 4.27 (a,c) show the noise in the image and its histogram. The noise histogram is widely spread out. In Figure 4.27 (b,d) the image SNR and the data histogram are shown.

The low image SNR reflects the extremely low intensity values in the DWI. The data histogram in (d) shows that the histogram has two populations, the first one can be Rician distribution for the pixels that had low SNR, and the second one was Gaussian distribution which was coming from pixels that had SNR value greater than 3, which indicate the noise distribution cannot be approximated by

Gaussian distribution.

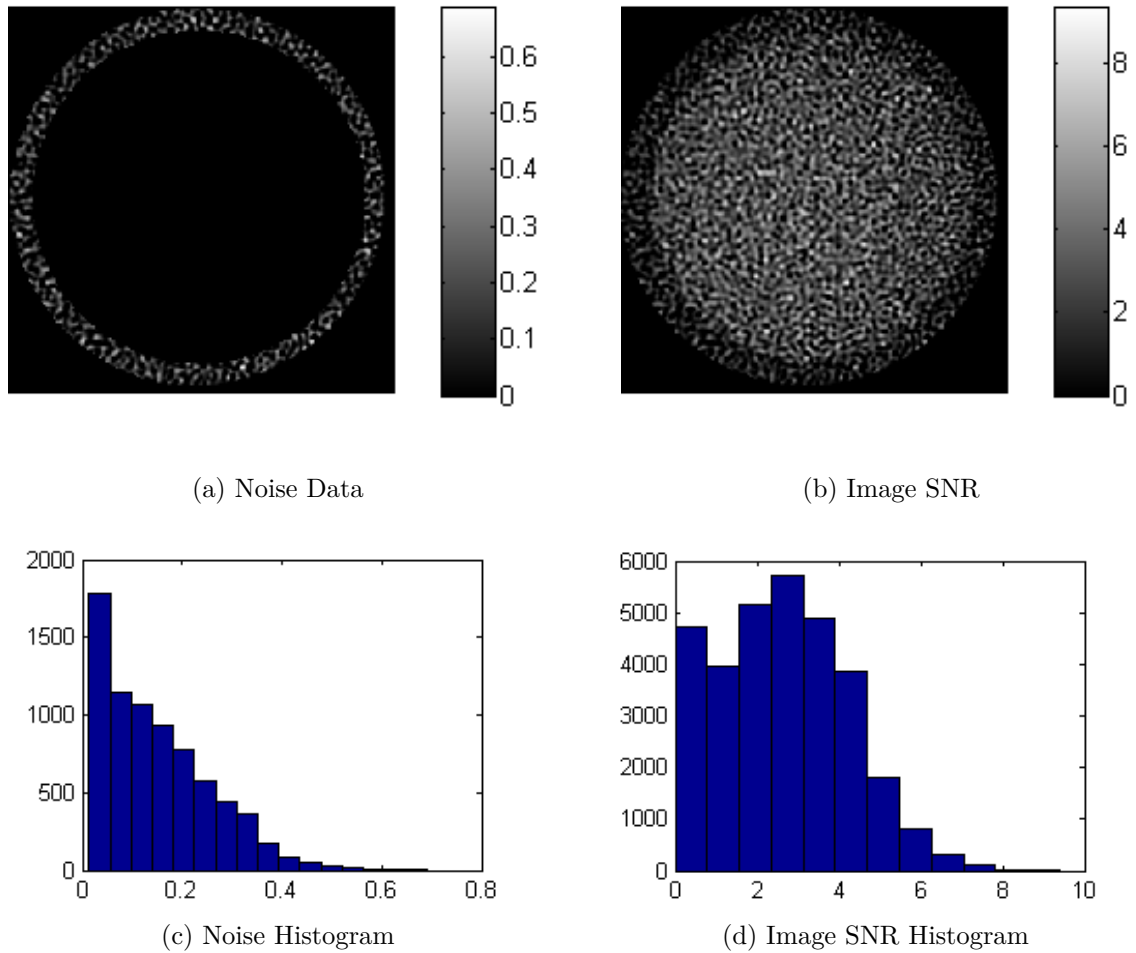


Figure 4.27: $b=10000$ Noise Study.

Next we show the EBA algorithm image reconstruction results. Figure 4.28 shows the reconstructed images with Beta values between 0.01 and 2. For low values of Beta, the images looked the same as the DWI. As Beta became large the images looked like the DWI in the center of the image and looked like the b_0 image on the side of the image.

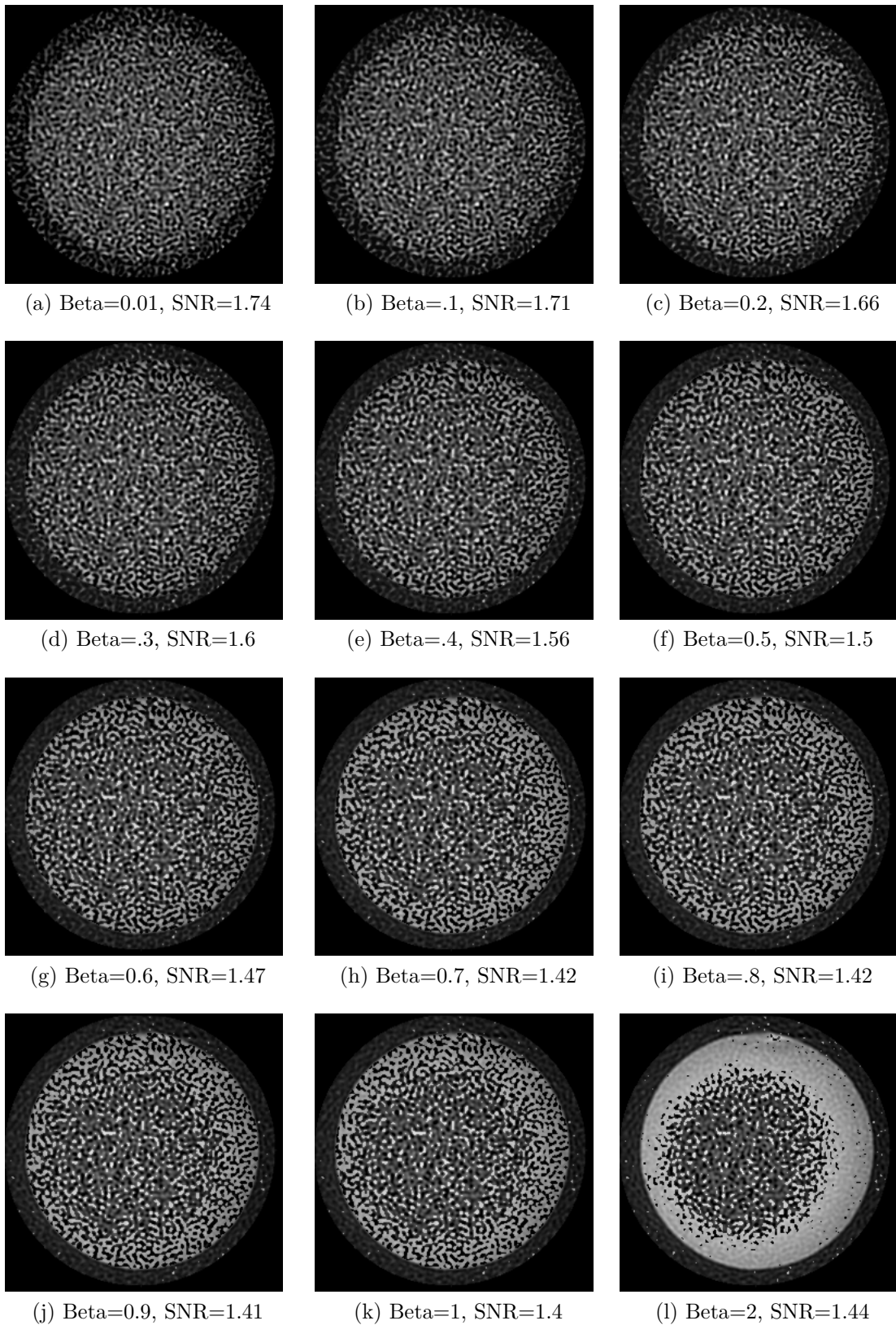


Figure 4.28: Spherical Phantom Image Reconstruction Results for $b=10000$.

The image SNR vs Beta is shown in Figure 4.29. The SNR went down for all values of Beta as the EBA algorithm increased the variance of the image.

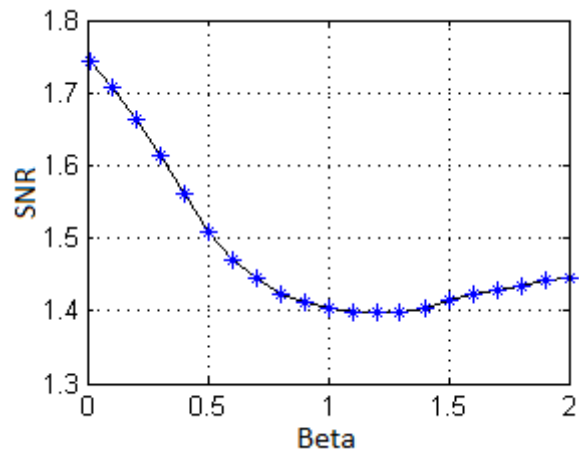


Figure 4.29: SNR vs Beta for $b=10000$.

Finally, we show the SNRs vs Beta for the three cases that illustrate the enhancement in image reconstruction for $b=500$, $b=1500$ and $b=2500$.

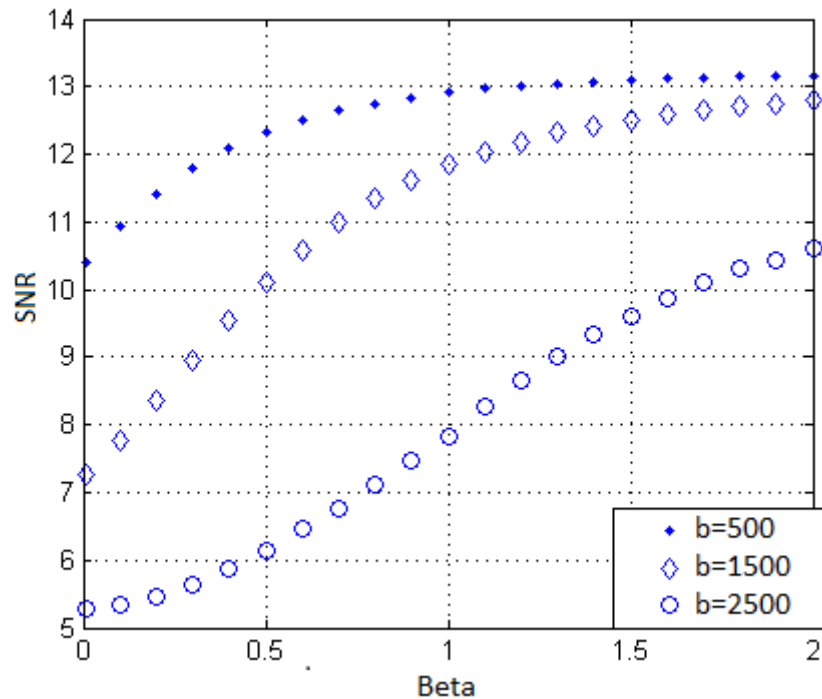


Figure 4.30: SNR comparison for different $b = 500, 1500, 2500$.

4.2.3 Logo Phantom

The second experimental data were generated with the GE logo phantom. Four sets of DWI images were collected. The b 's were 50, 100, 200, and 400. Figure 4.31 shows these images.

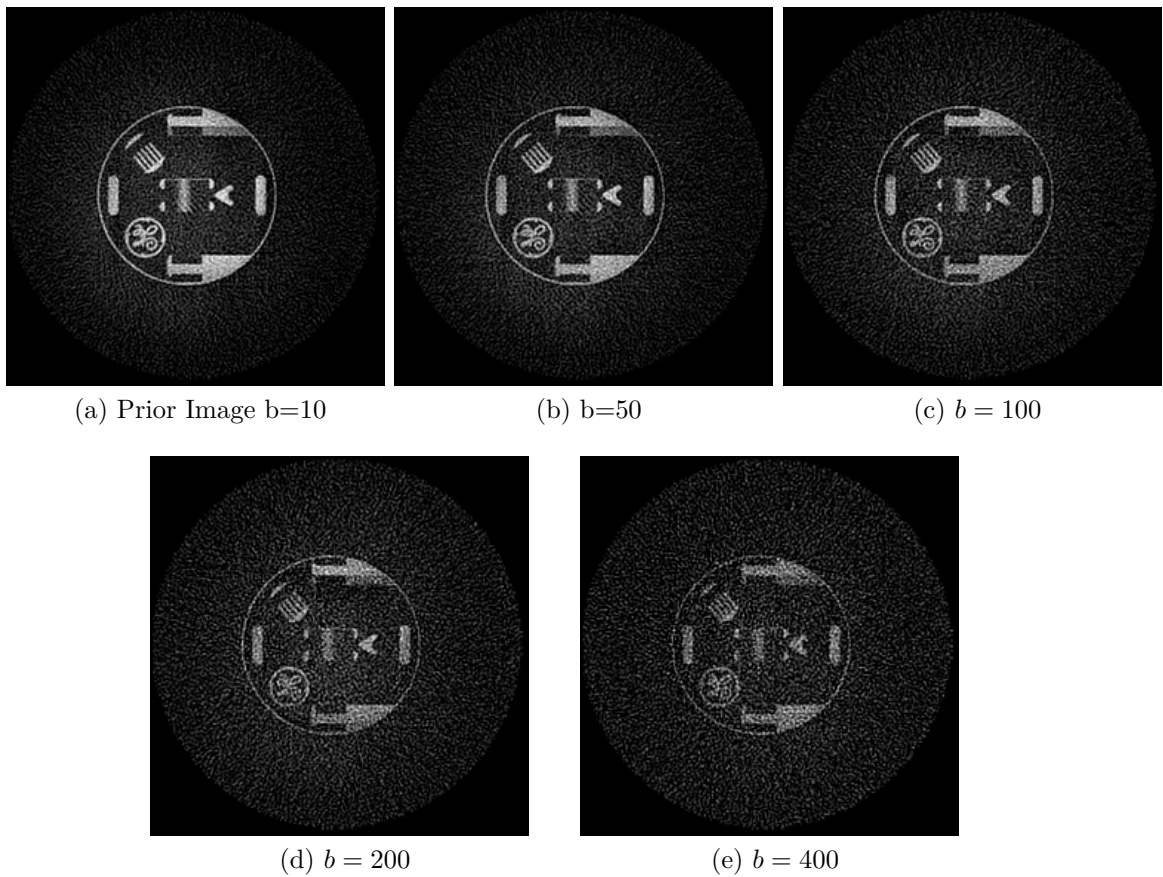


Figure 4.31: Sphere-GE Logo Images.

In the next subSections, we will show the comparison between the contemporary DWI reconstruction methods and the EBA algorithm.

4.2.3.1 Logo Phantom $b=50$

This Section shows the results when $b=50$. Figure 4.32 (a), (b) show the prior image and the DWI. In 4.32 (c) the histogram of the difference between the DWI and the prior is shown, the histogram is not much spread out, with most of the pixels had intensity difference less than 20%, which indicated the high similarity between the DWI and the prior image.

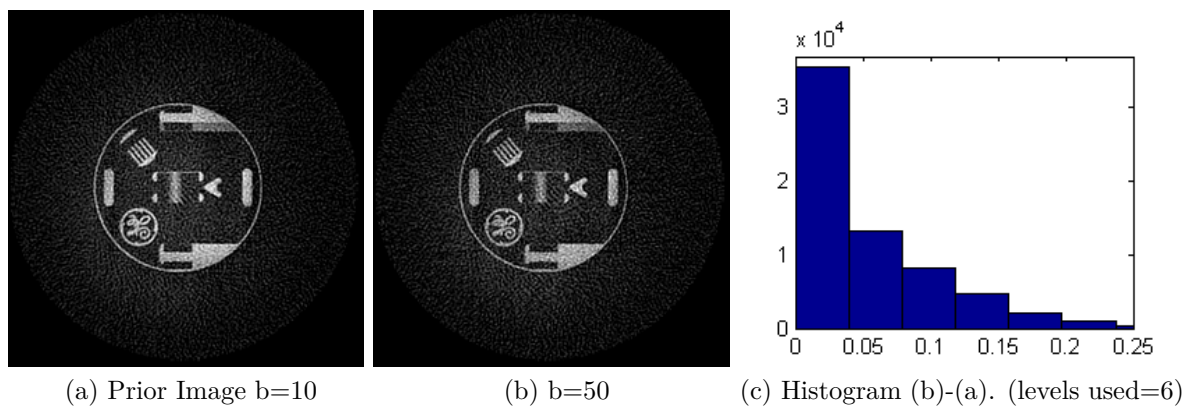
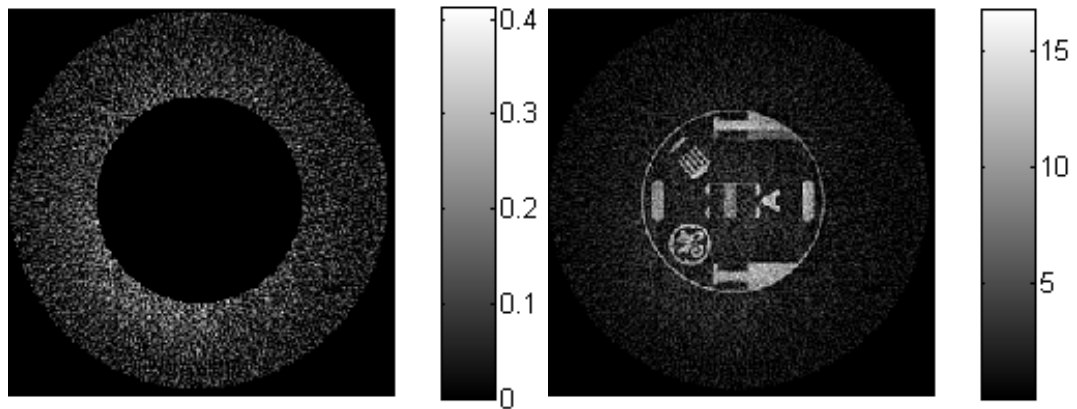


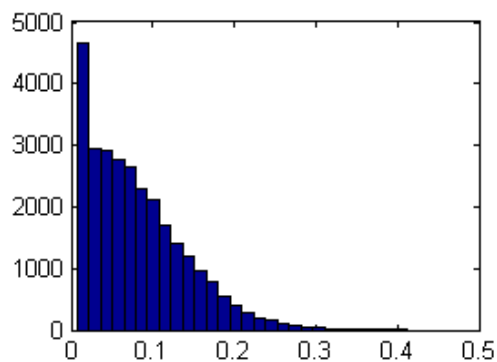
Figure 4.32: Logo Phantom Images $b=50$.

Figure 4.33 (a) shows the noise in the image. In Figure 4.33 (c) the histogram of the noise is shown, the histogram shows that the noise distribution can be approximated by a Gaussian function. In Figure 4.33 (b) the image SNR is shown. Most pixels had $\text{SNR} > 3$ which made it possible to approximate the noise by a Gaussian distribution. Figure 4.33 (d) shows the data SNR histogram. Again all of the pixels had SNR greater than 3 confirming the conclusion that was made before about the noise in this Section.

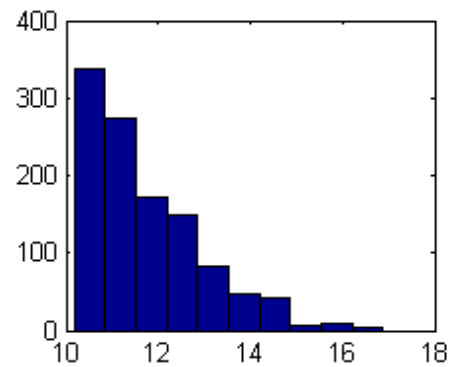


(a) Noise Data

(b) Image SNR



(c) Noise Histogram. (levels used=20)



(d) Image SNR Histogram. (levels used=10)

Figure 4.33: Logo Noise Study for $b=50$.

Figure 4.34 show the improvement in the SNR vs Beta. A typical value of Beta is 0.3 at which the improvement of the SNR was 15%.

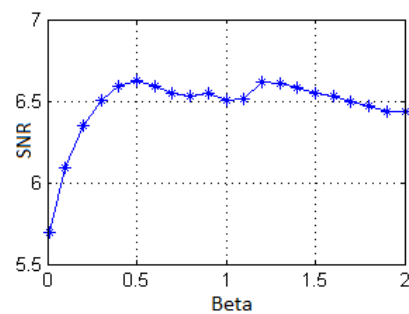
Figure 4.34: Logo SNR vs Beta for $b=50$.

Figure 4.35 shows the reconstructed images. The images had less noise as Beta was increased.

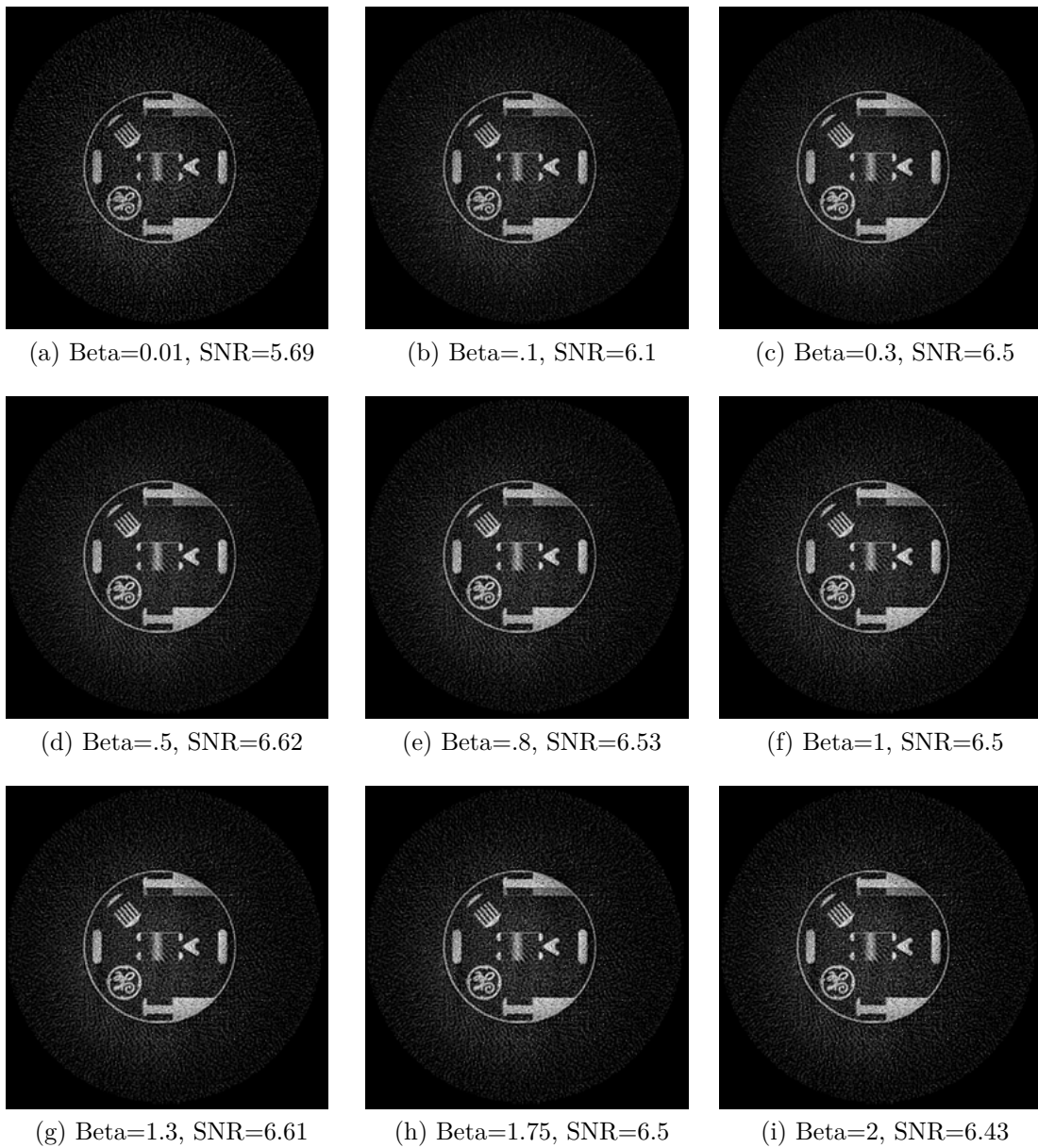


Figure 4.35: Logo Results for $b=50$.

4.2.3.2 Logo Phantom $b=100$

This Section shows the results for $b=100$. Figure 4.36 (a), (b) show the prior image and the DWI, in 4.36 (c) the histogram of the difference between the DWI and the prior is shown, the histogram is not much spread out, indicating the similarity is high between the DWI and the prior image. This histogram is very similar to the previous case 4.32 (c), where Beta=50.

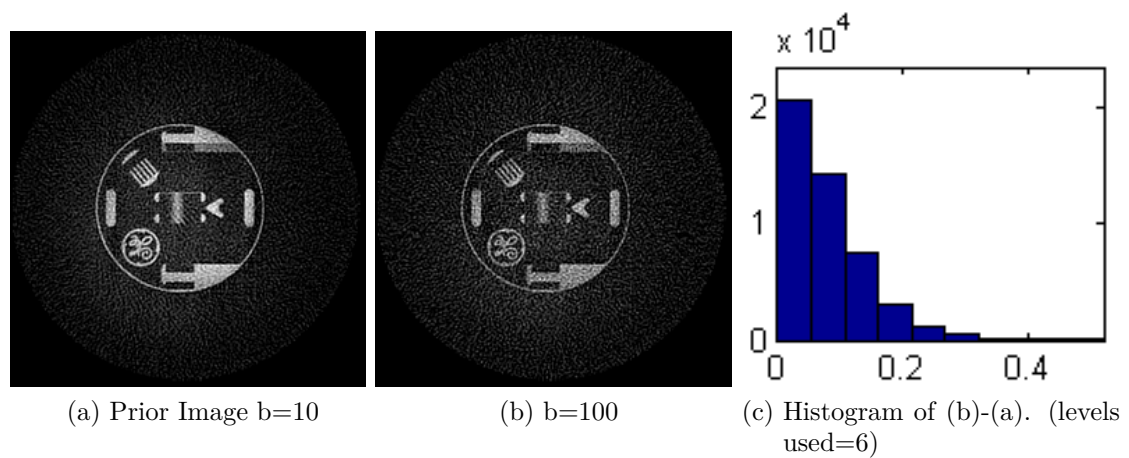
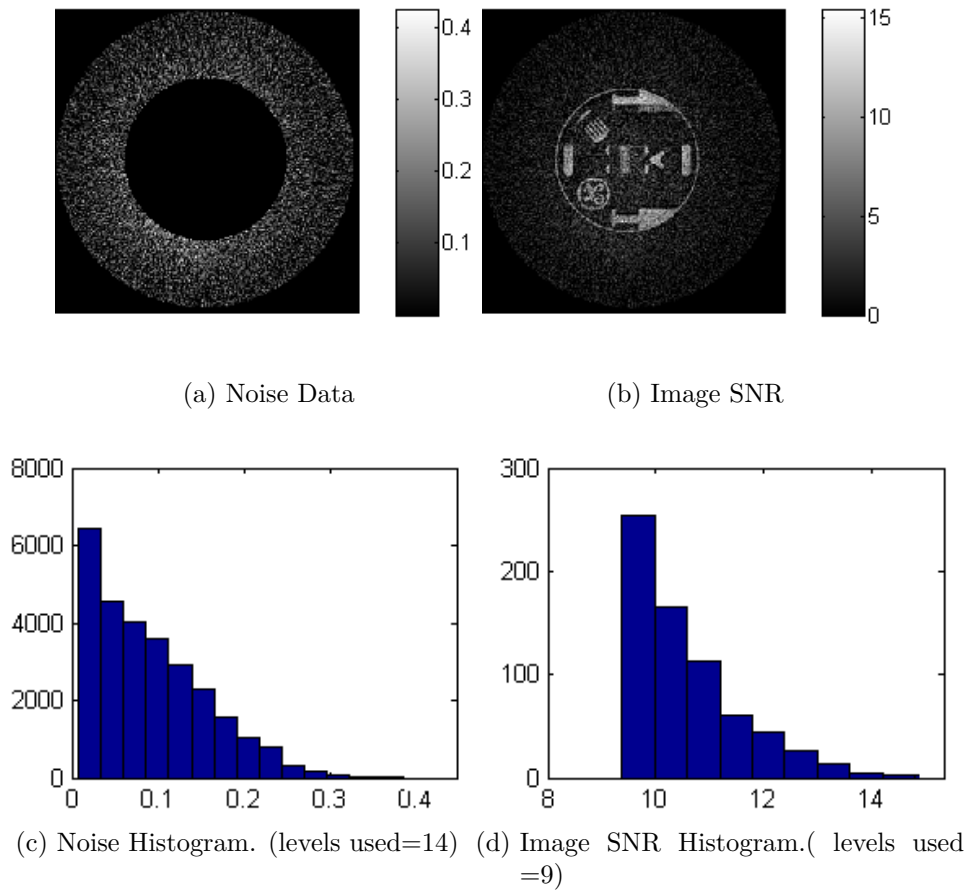


Figure 4.36: Sphere Logo for $b=100$.

Figure 4.37 a shows the noise in the image. In Figure 4.37 the histogram of the noise is shown, the histogram shows that the noise distribution can be approximated by a Gaussian function. In Figure 4.37 (c) the image SNR is shown. Most pixels had $\text{SNR} > 3$ which made it possible to approximate the noise by a Gaussian distribution. Figure 4.37 (d) shows the data SNR histogram. All of the pixels had SNR greater than 9 which confirm the conclusion was made before about the noise that it can be approximated by the Gaussian distribution.

Figure 4.37: Noise Study $b=100$.

To compare the performance of the algorithm, we computed the SNR of each image vs Beta. The results are shown in Figure 4.38. The SNR improved by 8% at Beta =0.3.

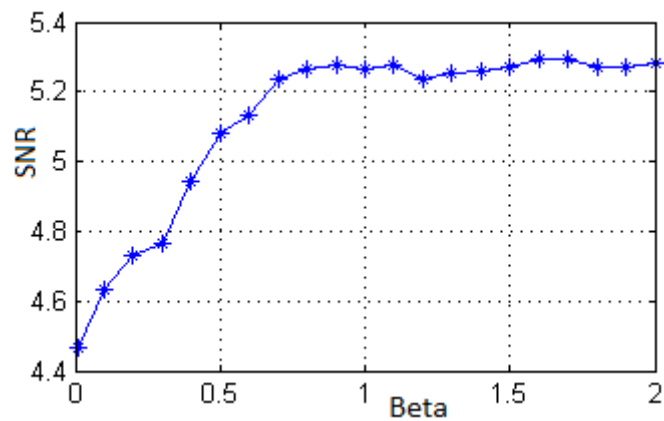
Figure 4.38: Logo SNR vs Beta for $b=100$.

Figure 4.39 shows the reconstructed images. The images had less noise as Beta was increased.

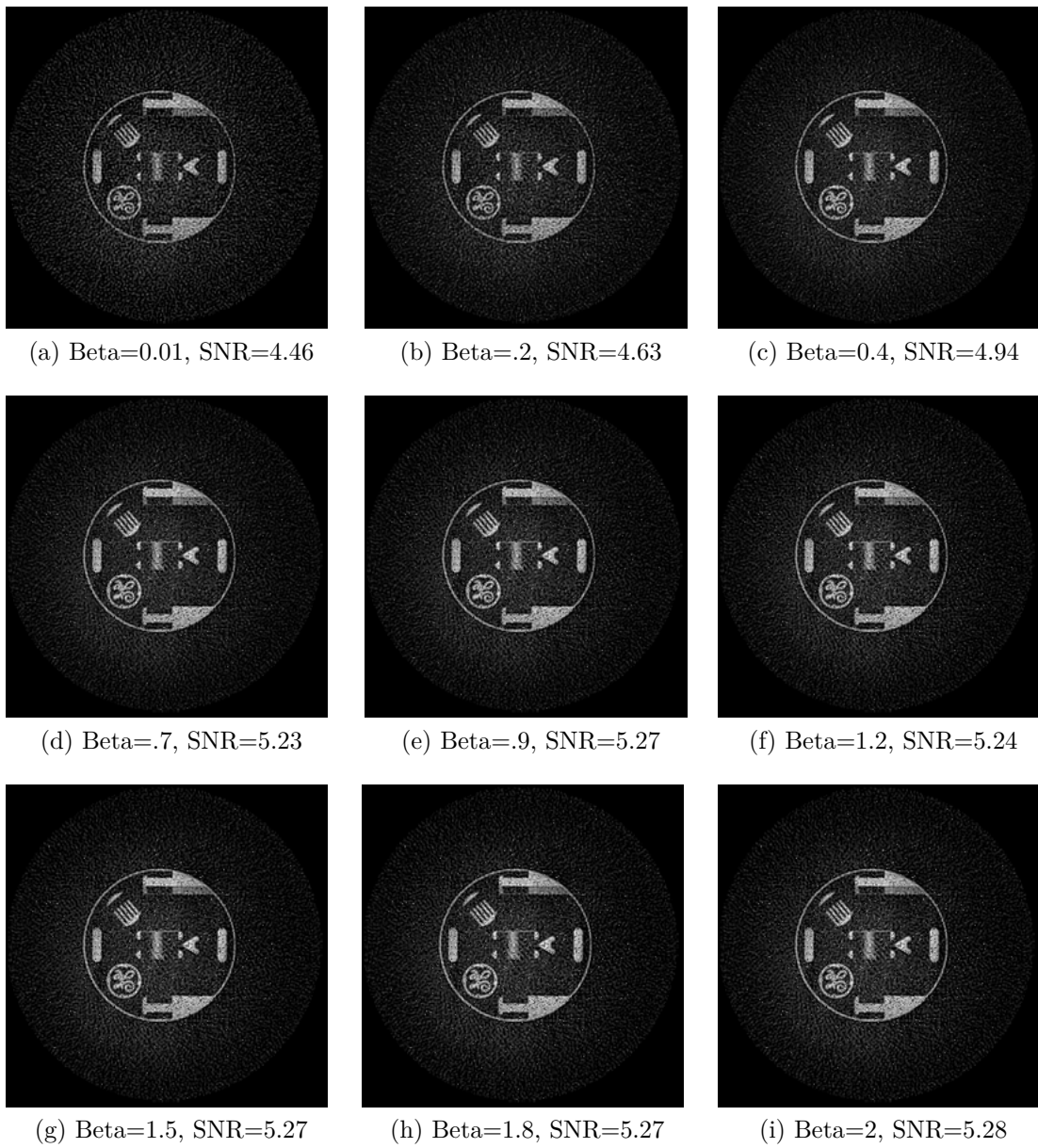


Figure 4.39: Logo Results for $b=100$.

4.2.3.3 Logo Phantom b=200

This Section shows the results when $b=200$. Figure 4.40 (a) and (b) show the prior image and the DWI, in 4.40 (c) the histogram of the difference between the DWI and the prior is shown, Most of the pixels fall within a 20% difference which indicates the high similarity between the DWI and the prior image.

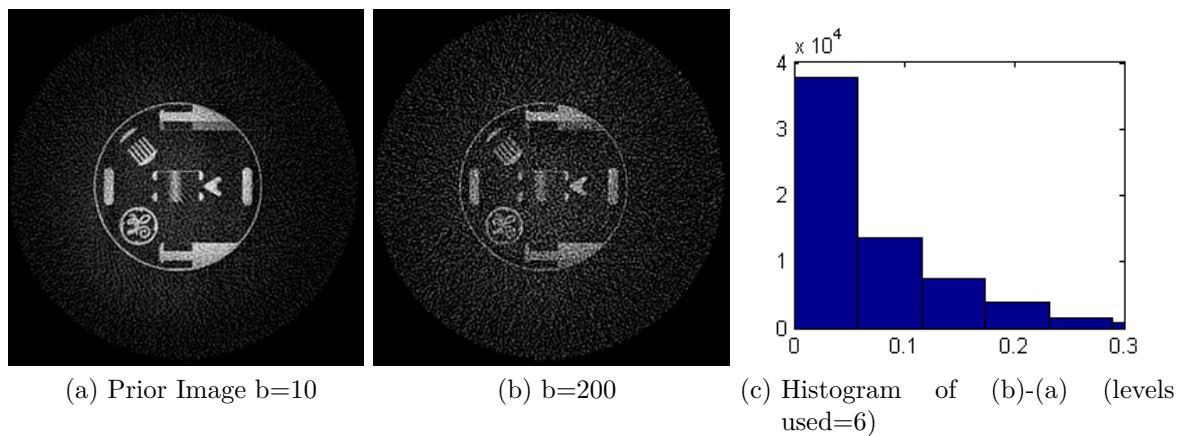
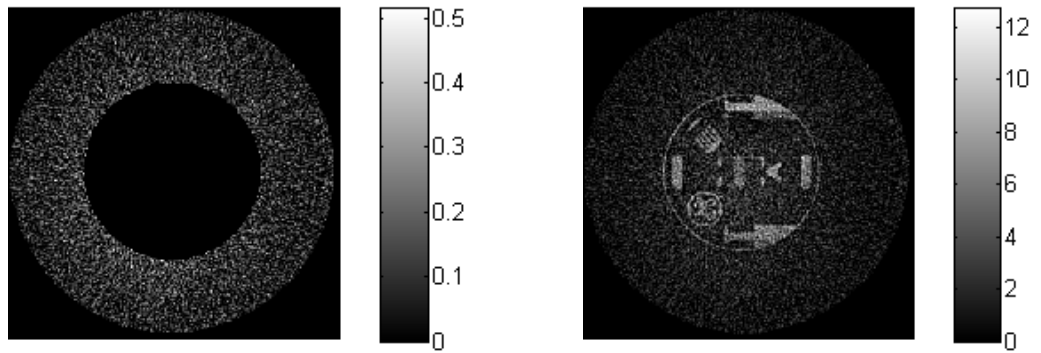


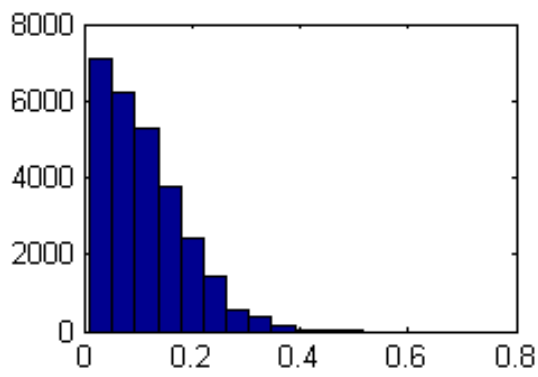
Figure 4.40: Sphere Logo b200.

Figure 4.41(a) and (c) show the noise of the image and its histogram. The noise histogram assumes a Gaussian shape. In Figure 4.41 (b) the image SNR is shown. Most of the pixels had SNR higher than 3. In figure 4.41 (d) the phantom data SNR histogram is shown. All pixels had an SNR that was greater than 7. This would make the Gaussian approximation of the noise a valid assumption.

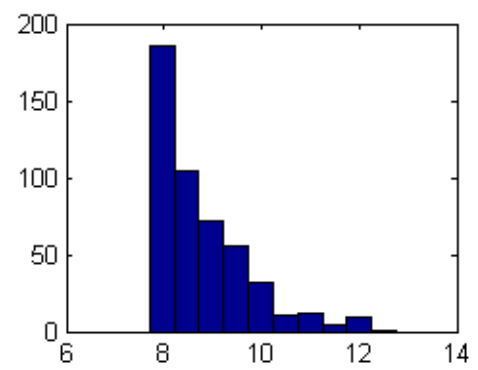


(a) Noise Data

(b) Image SNR



(c) Noise Histogram (levels used=12)



(d) Image SNR Histogram (levels used=12)

Figure 4.41: Logo Noise Study for $b=200$.

Figure 4.42 shows the SNRs and the improvement in it.

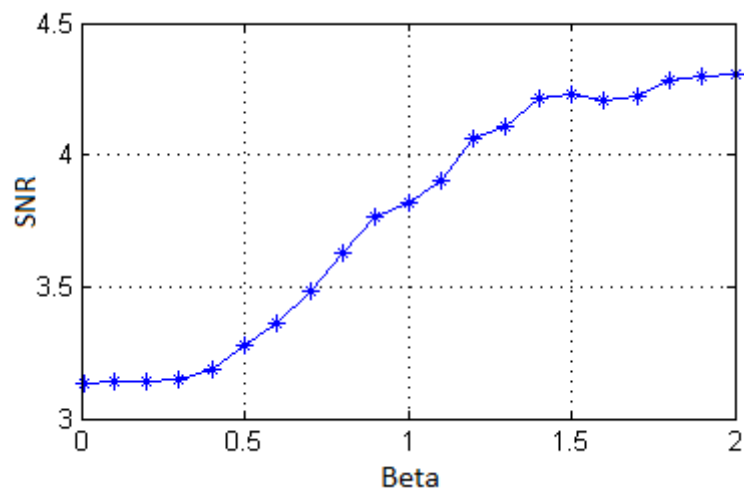
Figure 4.42: Logo SNR vs Beta for $b=200$.

Figure 4.43 shows the set of images and the improvement in the SNRs.

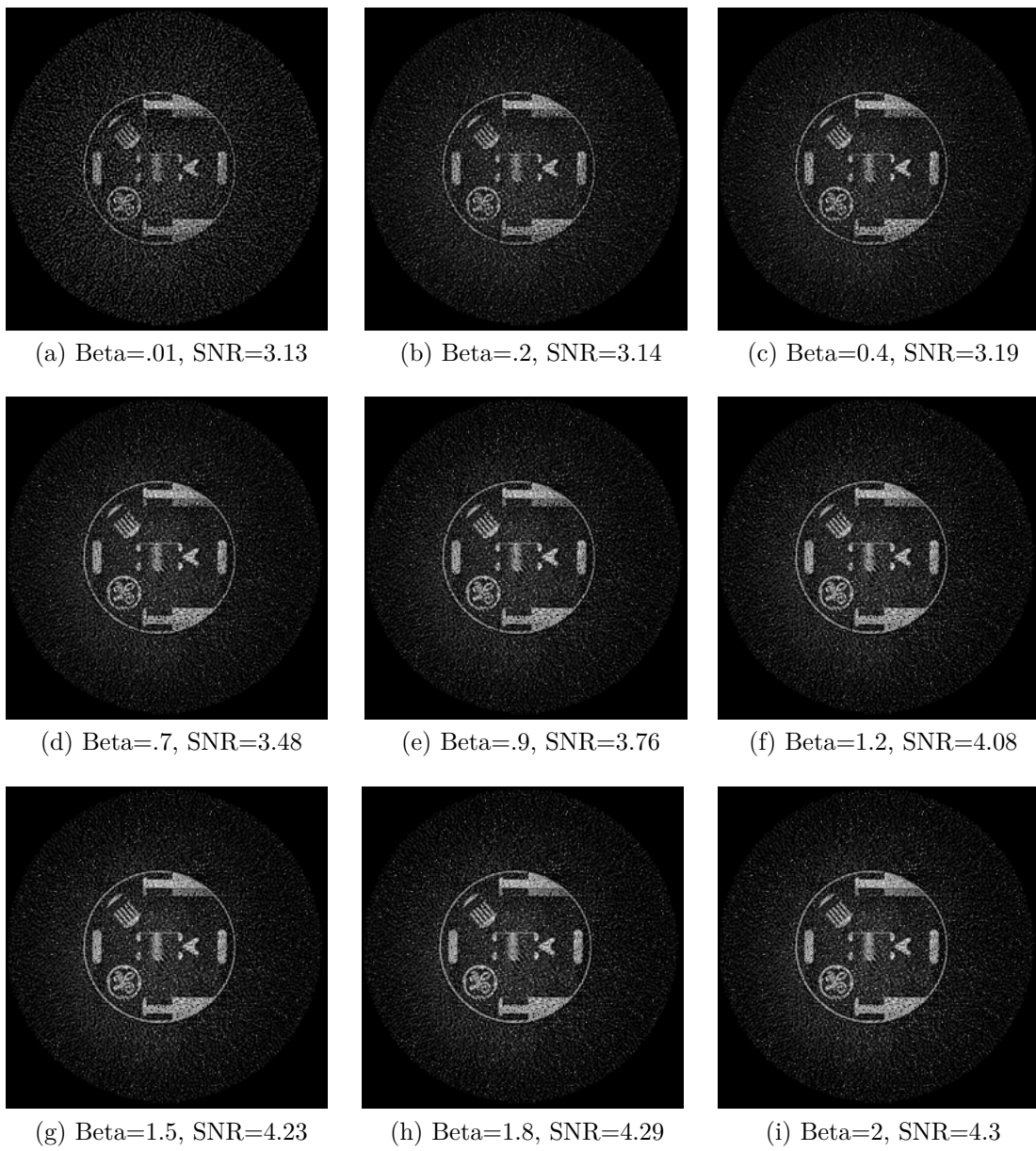


Figure 4.43: Logo Results for $b=200$.

4.2.3.4 Logo Phantom $b=400$

This Section shows the results when $b=400$. Figure 4.44 (a),(b) show the prior image and the DWI. In 4.44 (c) the histogram of the difference between the DWI and the prior is shown. Most of the histogram bins fall below 20% with a few of the bins fall below 40%. This indicates good similarity between the DWI image and the b_0 image although the similarity is not as much as it is for the previous cases when the b -value is smaller.

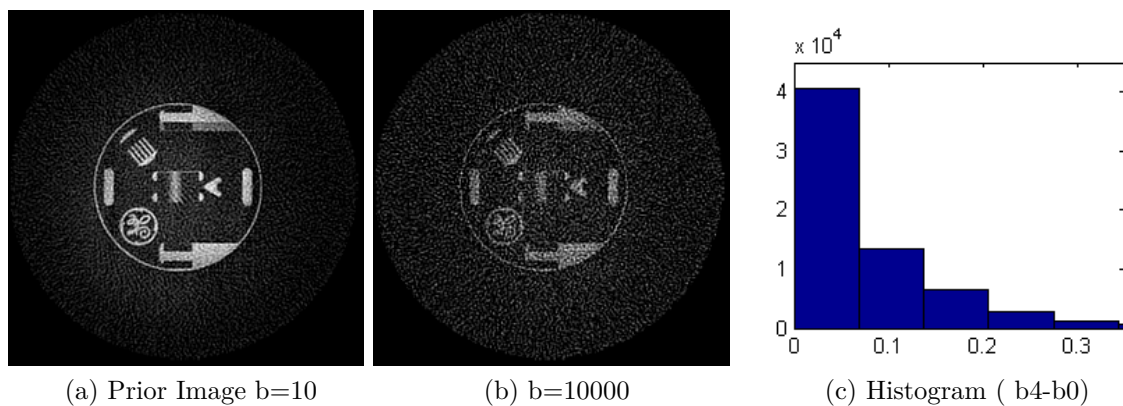
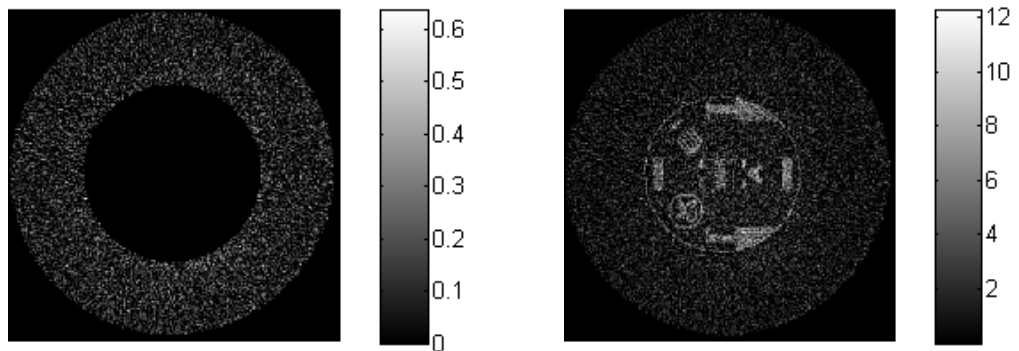


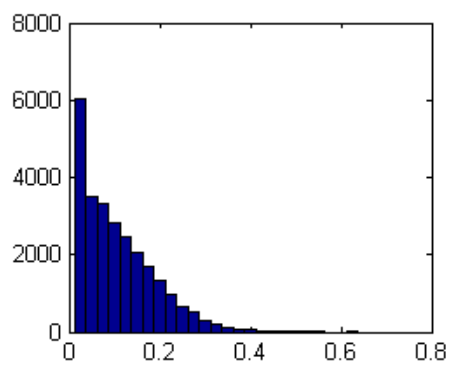
Figure 4.44: Sphere Logo $b400$.

Figure 4.45 (a) shows the noise in the DWI. In Figure 4.45 (c) the histogram of the noise is shown and it has a Gaussian shape. In Figure 4.45 (b), the image SNR is shown. The histogram shows that most of the pixels have an SNR > 3 . In Figure 4.45 (d) the image SNR histogram of the image is shown. All of the data phantom pixels have SNR > 7 .

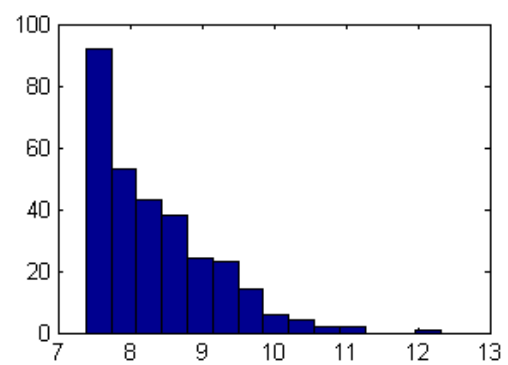


(a) Noise Data

(b) Image SNR.



(c) Noise Histogram (levels used=22)



(d) Image SNR Histogram (levels used=14)

Figure 4.45: b400 Noise Study

The EBA algorithm is run with different Beta values ranging between 0.01 and 2. For each Beta, the SNR of the image is computed and Figure 4.46 shows the improvement of SNRs with Beta.

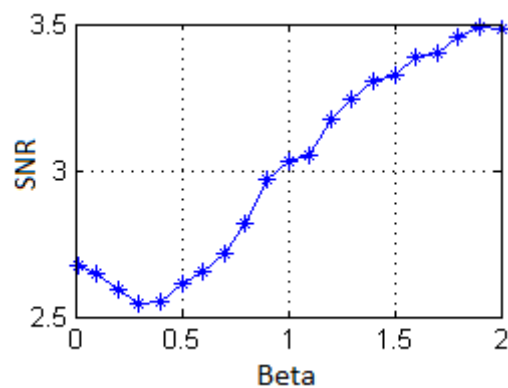


Figure 4.46: Logo SNR vs Beta for b=400.

Figure 4.47 shows the set of results from the EBA algorithm for different Beta values. The graph shows that beta enhances the image clarity.

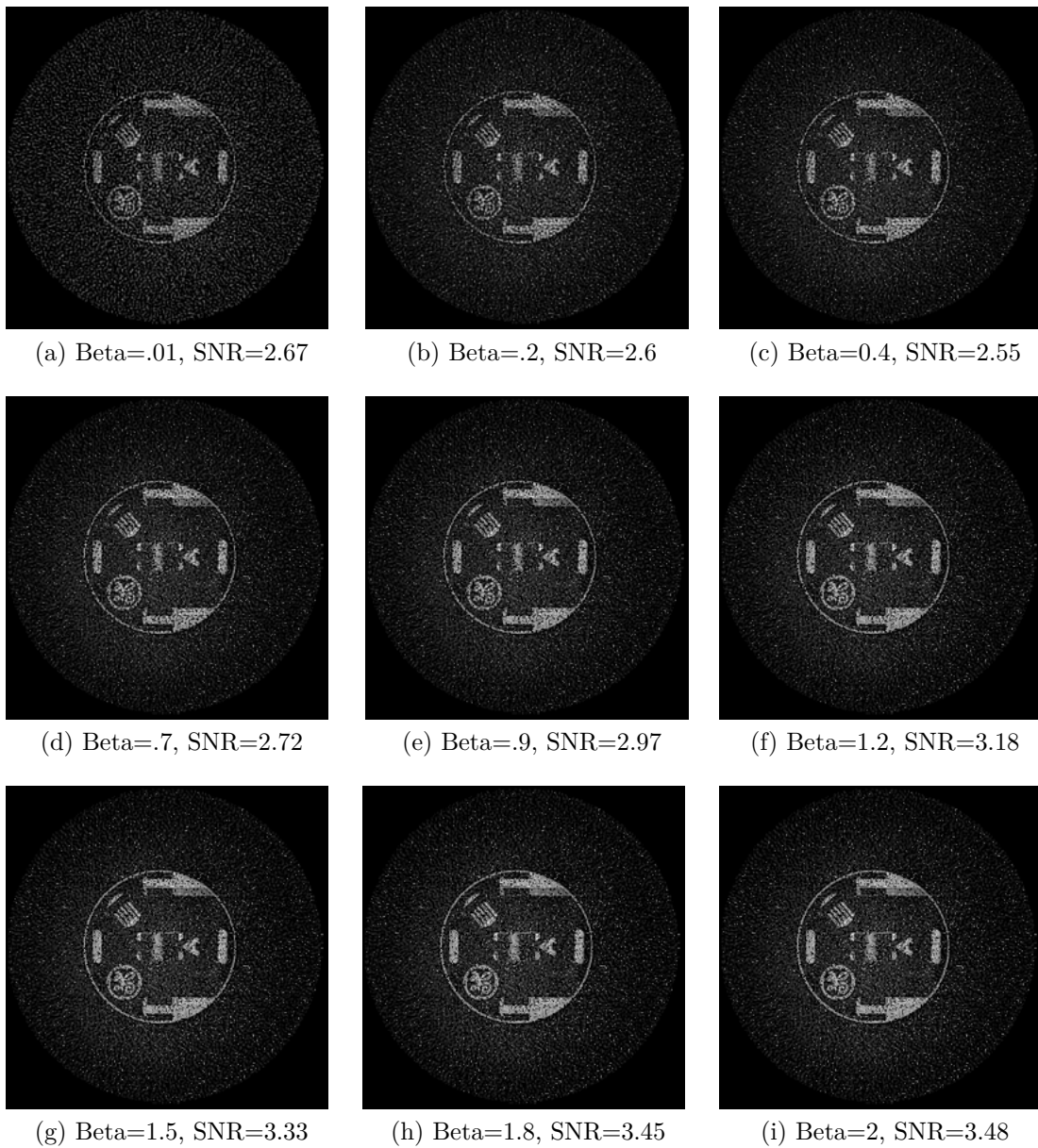


Figure 4.47: Logo Results for $b=400$.

4.3 Anatomical Results

The third part of the practical study is to test the algorithm on real anatomical data. We have two data sets. The first one is for a brain image, and the second one is a stack of 28 slices of Propeller images of a brain. We begin by showing results of the brain image.

4.3.1 Brain Results

In this Section, we show simulation results of DWI images of the brain. These images are obtained from www.radiopaedia.org. Figure 4.48 shows two images of the brain. The image in (a) has no pathology and Gaussian noise is added to it to generate the prior image. The image in (b) had pathology and Rician noise is added to (b) to form the DWI in (d) and Gaussian noise is added to (b) to form the DWI in (c).

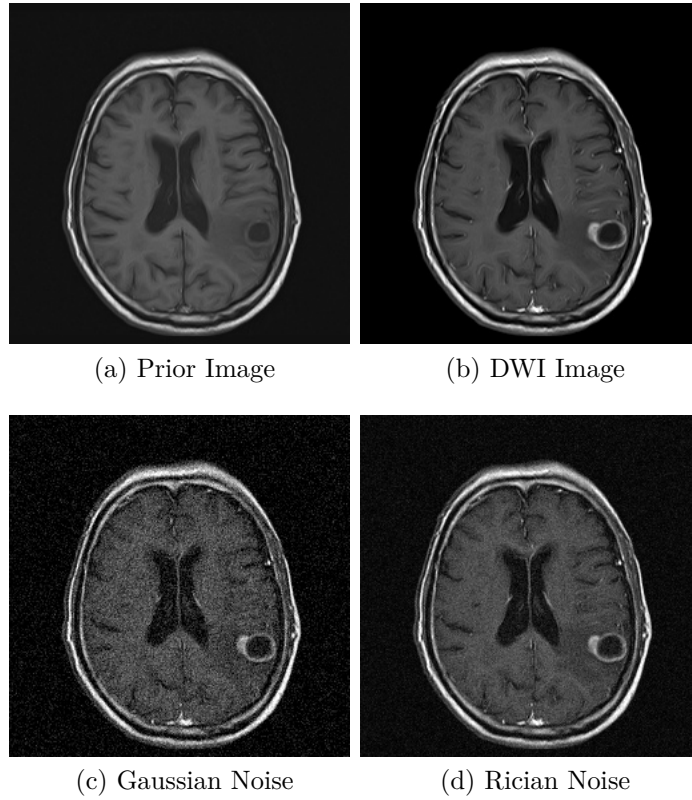


Figure 4.48: Brain Images With Pathology.

The goal of this experiment is to investigate the clinical performance of the EBA algorithm. We start by generating the histograms of the difference between the DWI and the prior images in Figure 4.48. Figure 4.49 shows histogram of each image.

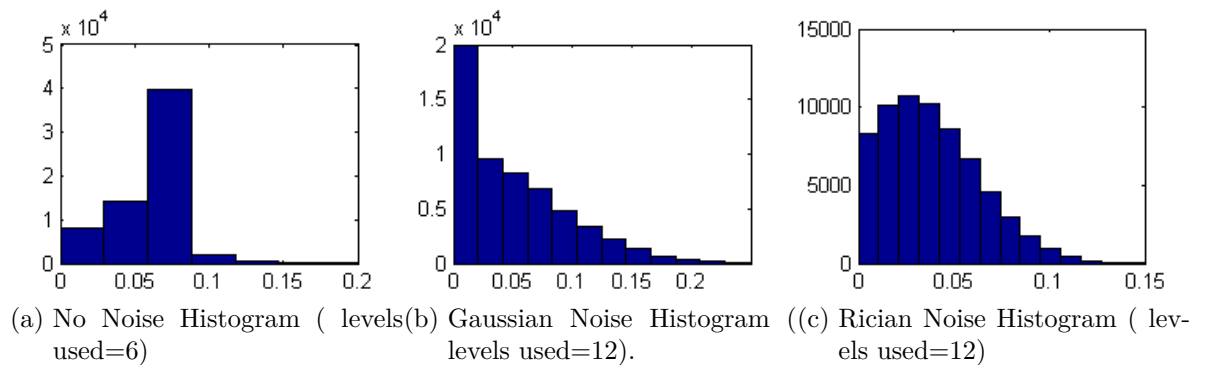


Figure 4.49: Pathology Histograms.

The histograms show that most of the pixels had small difference in the DWI image and the prior, indicating a proper choice for prior.

Figure 4.50 shows the results of the noise analysis of the DWI with Gaussian noise. Figure 4.50 (a) shows the DWI image with the Gaussian noise and in (c) its histogram of the noise is shown. The histogram assumes Gaussian shape as expected. In Figure 4.50 (b), the image SNR is shown. The SNR is higher than 3 for most pixels which confirm the Gaussian noise. In (d) the histogram of the image SNR is shown with almost all pixels having SNRs greater than 5 in agreement with [8].

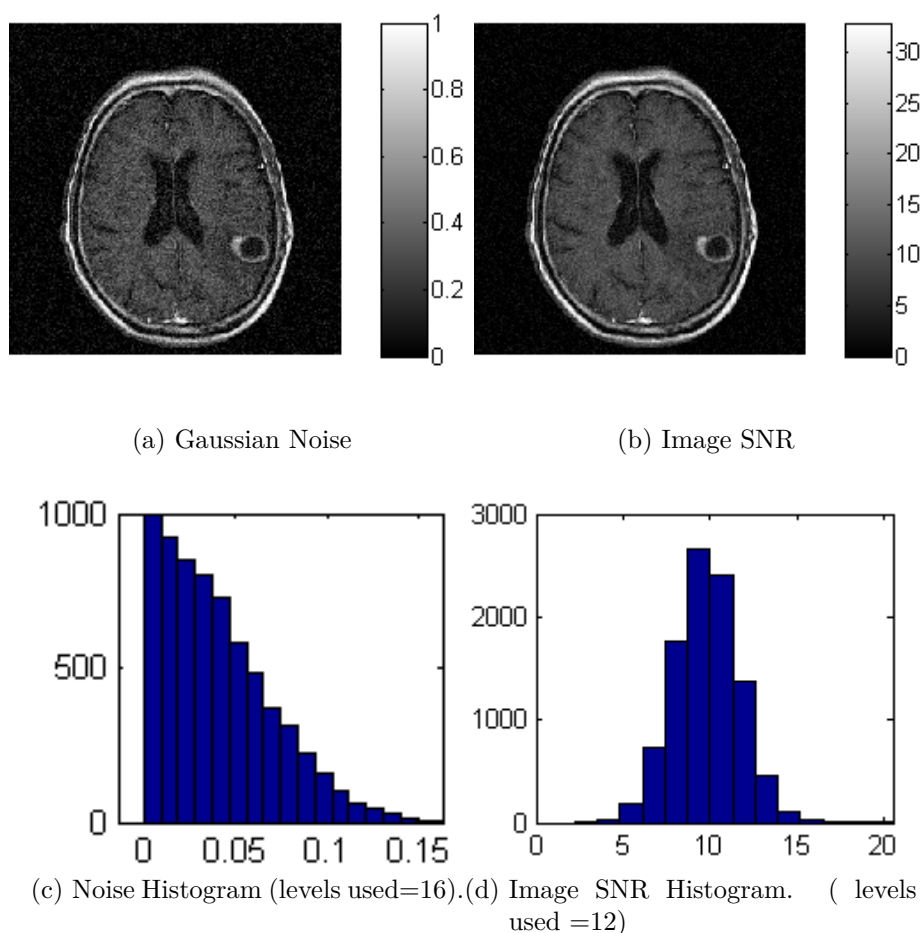


Figure 4.50: Pathology Gaussian Noise Study.

Figure 4.51 shows the noise study of the DWI with Rician noise. In (a) the image with Rician noise is shown, in (c) the histogram of the noise is shown, and it had the Rician distribution shape. In (b) the image SNR is shown. Most pixels had SNR greater than 3 which made it possible to approximate the noise by a Gaussian distribution. In (d) the histogram of the image SNRs is shown and it shows that almost all pixels had SNRs greater than 3 which agrees with the previous conclusion about the noise.

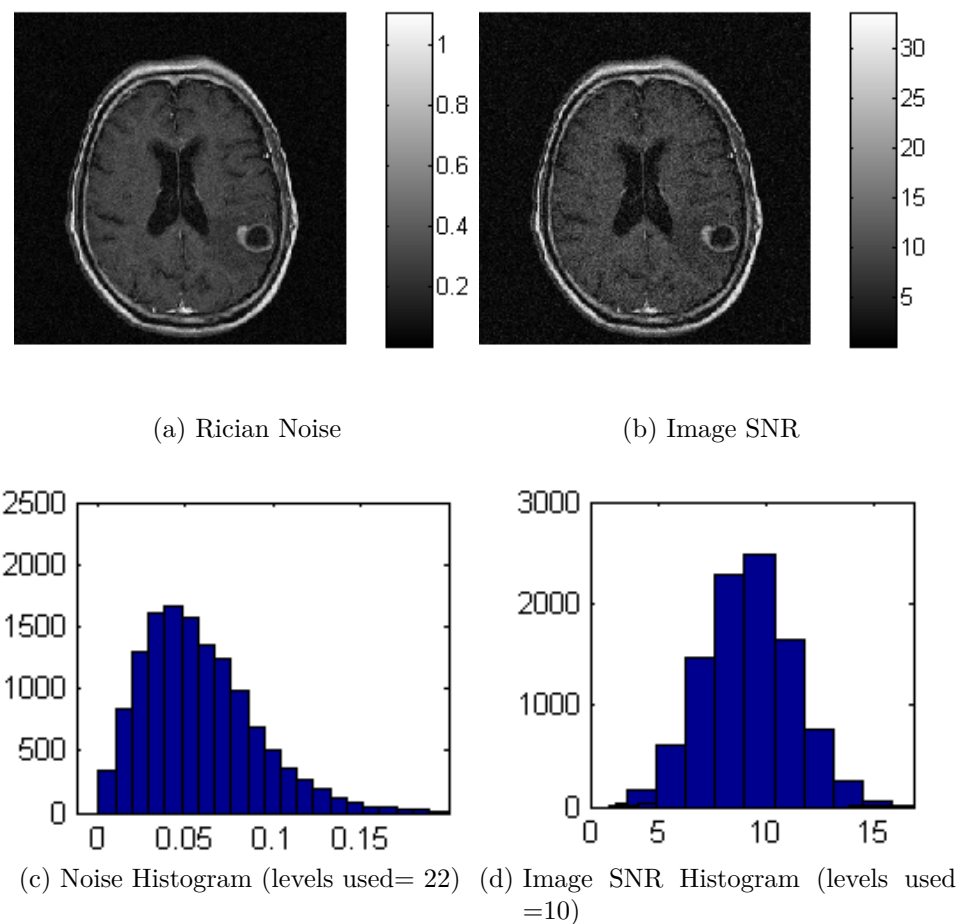


Figure 4.51: Pathology Rician Noise Study.

Next we show the results of the EBA algorithm. Figure 4.52 shows the improvement in the SNR vs Beta for both the Gaussian noise and the Rician noise. It is clear that as Beta increased the SNR increased as well.

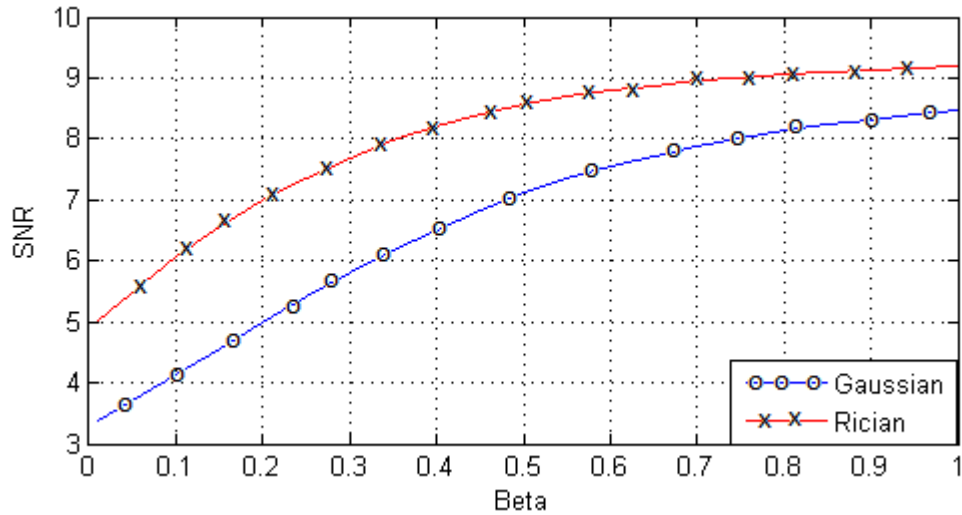


Figure 4.52: Brain Pathology SNR vs Beta

Figure 4.53 shows the reconstructed images when the noise was Gaussian for different values of Beta. The graph shows that as Beta increases, the image quality improves as well. However for large values of Beta, the algorithm wipes out the pathology area that is shown in Figure 4.53 (a) and makes it similar to that in the prior image. Our conclusion is to limit Beta to be less than 0.5. Figure 4.54 shows the reconstructed images with Rician noise. The same conclusion can be made about the quality of the images as in the previous case.

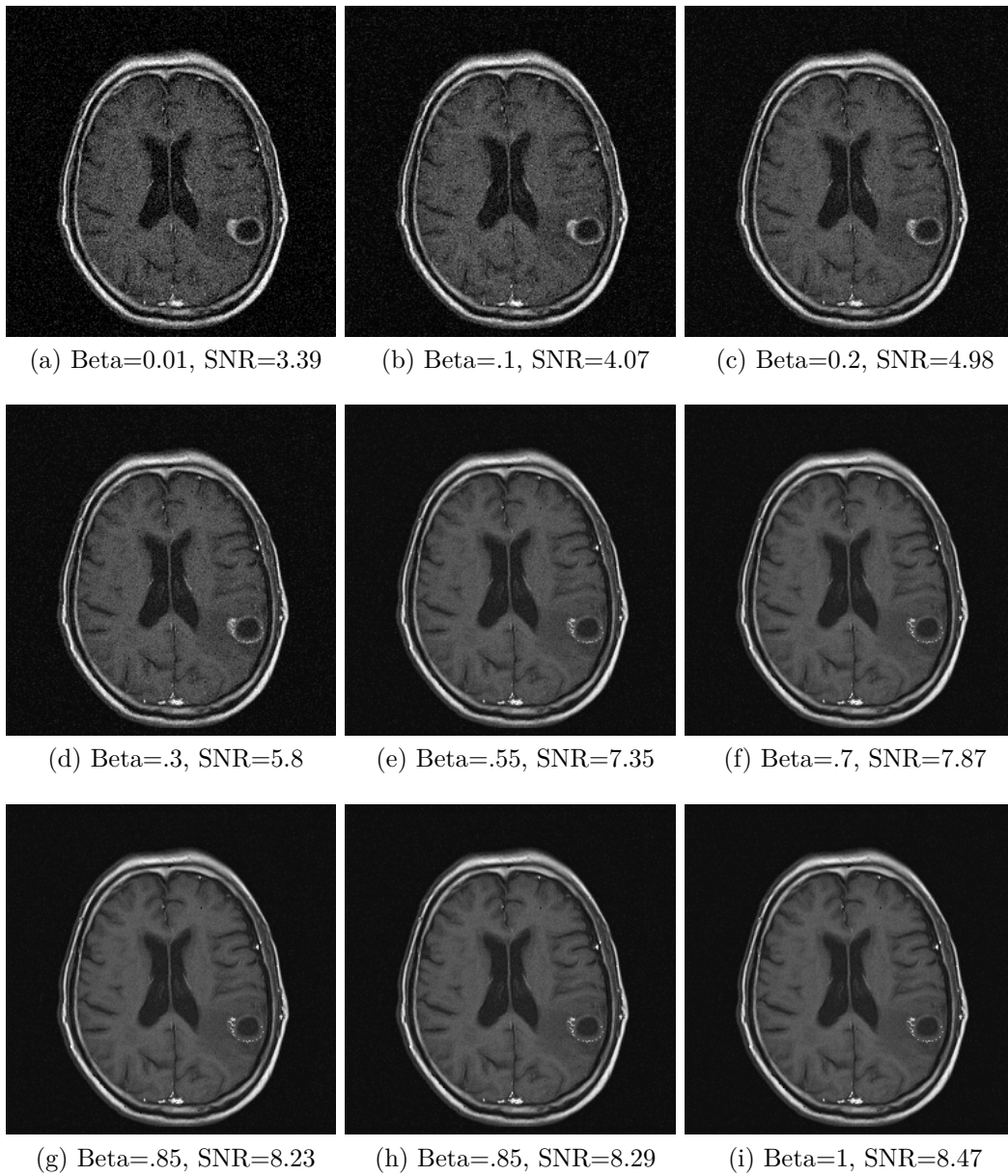


Figure 4.53: Brain Results With Gaussian Noise

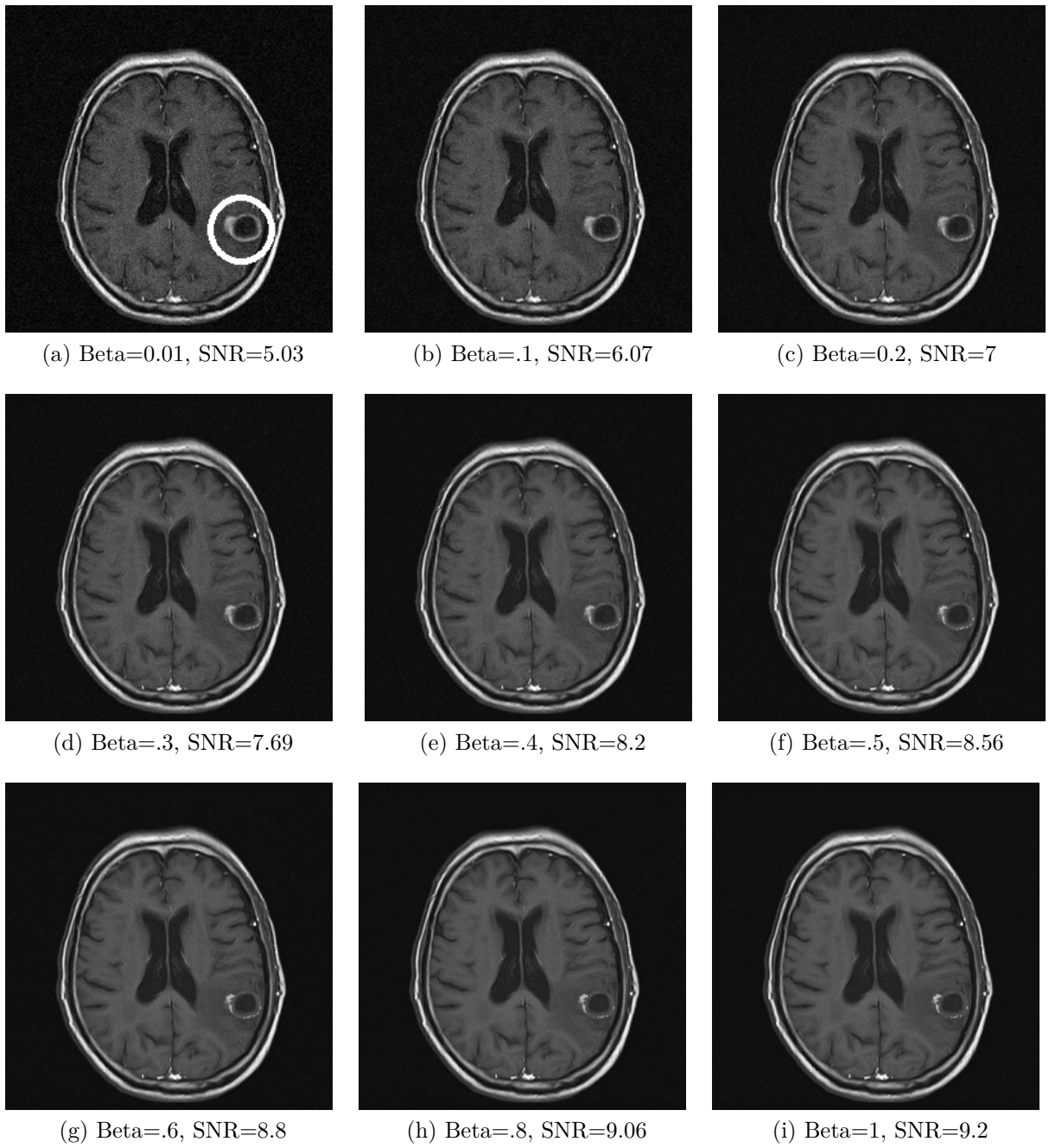


Figure 4.54: Brain Results with Rician Noise.

4.3.2 Propeller Images With Fixed Beta Value.

The MRI imaging technique is very sensitive to patient's motion. PROPELLER [21]“Periodically Rotated Overlapping Parallel Lines with Enhanced Reconstruction” MRI is designed to reduce the effect of patient voluntary and physiologic motion (breathing, flow). Figure 4.55 shows a filling pattern that is used to fill a radial k-space which is less sensitive to motion.

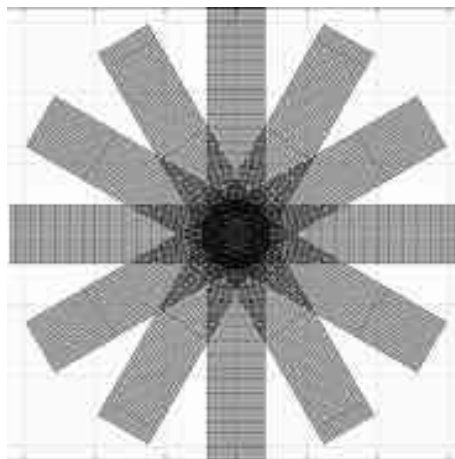
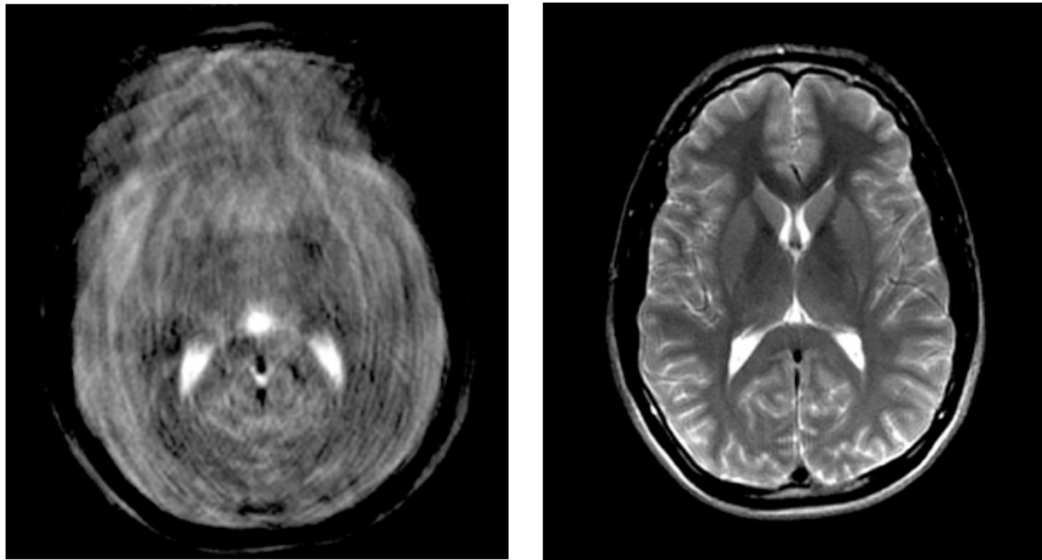


Figure 4.55: Propeller ¹

The oversampling of the k-space center typical for radial k-space filling yields increased SNR and high tissue contrast.

Figure 4.56 (a) shows MRI head image without Propeller; Figure 4.56 (b) shows the same head MRI image with Propeller motion correction.

¹Courtesy of GE healthcare.



(a) MRI head image without Propeller

(b) MRI head image with Propeller

Figure 4.56: Propeller Motion Artifact Correction. ²

In this Section we apply the EBA method on four data sets of Propeller brain images. For each image, Beta will have a fixed value, hence we call it "Fixed Beta EBA". We will show three cases where the algorithm improve the quality of the images, and one case where the algorithm cannot improve the quality of the reconstructed DWI images.

²Courtesy of Siemens healthcare.

4.3.2.1 Slice 5

The first example is for a case where the algorithm works. Figure 4.57 shows the b0 image in (a) and the DWI image in (b) for slice 5. The histogram of their difference is shown in (c). Most of the bins of the histogram are less than 20% difference and the histogram is not spread out. This makes the b0 image a good prior for this case.

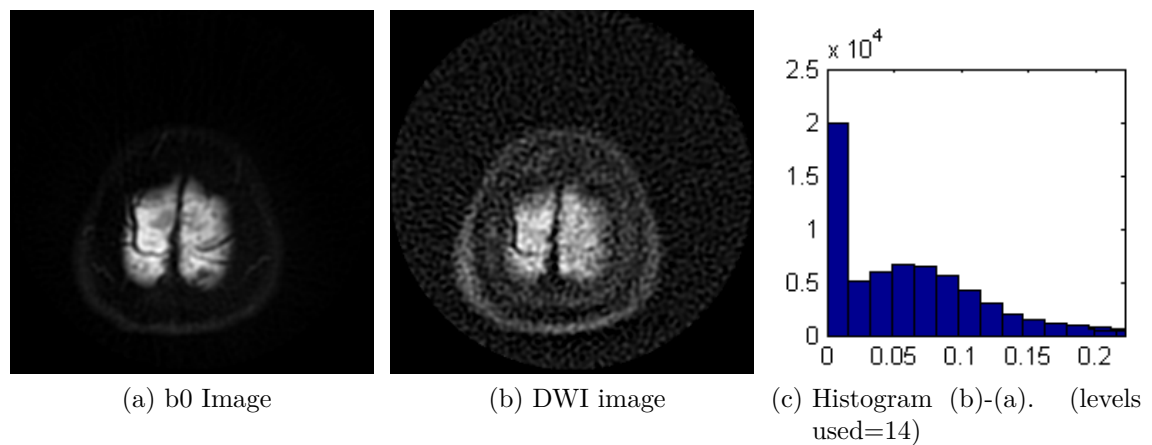


Figure 4.57: Propeller Images Slice 5.

In Figure 4.58 we show the results of the analysis of the noise in this slice. Figure 4.58 shows the results we had. In Figure 4.58, (a) is the DWI image and (b) is the image SNR where the majority of the pixels have an SNR greater than 3, enabling us to assume Gaussian noise in our image. In Figure 4.58 (c), we show the noise histogram of the image, fitting with a Rician distribution. In Figure 4.58 (d), the image data SNR is shown, a lot of pixels had an SNR greater than 3 at which the Gaussian noise assumption is valid.

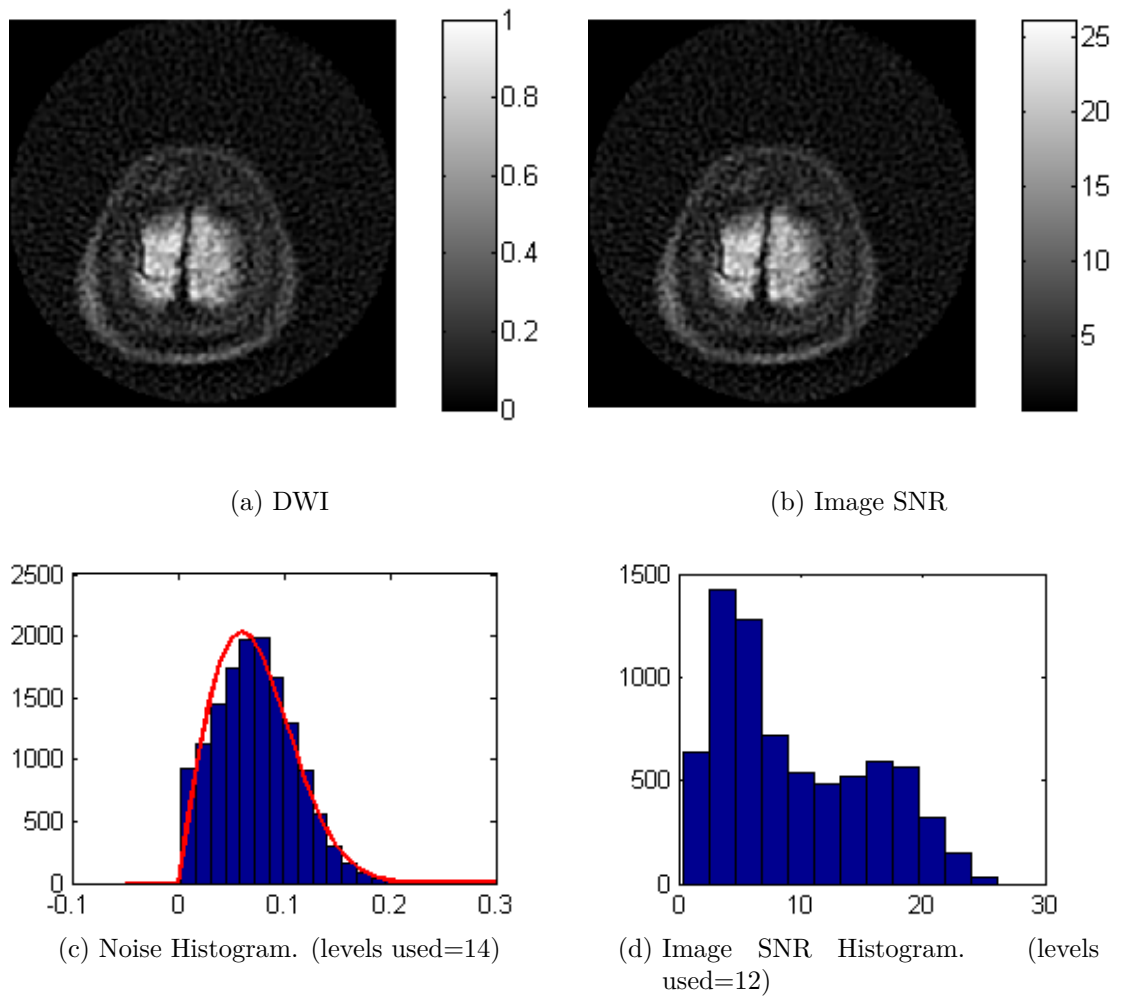


Figure 4.58: Propeller Images Slice 5 Noise Analysis Data.

Next we show the reconstructed images using the EBA method. Figure 4.59 shows the reconstructed images for different Beta values. The figure shows that as Beta increases the images have less noise. Figure 4.60 summarizes the improvement in the SNRs with a graph. At Beta=0.5, the SNR improves by 60%.

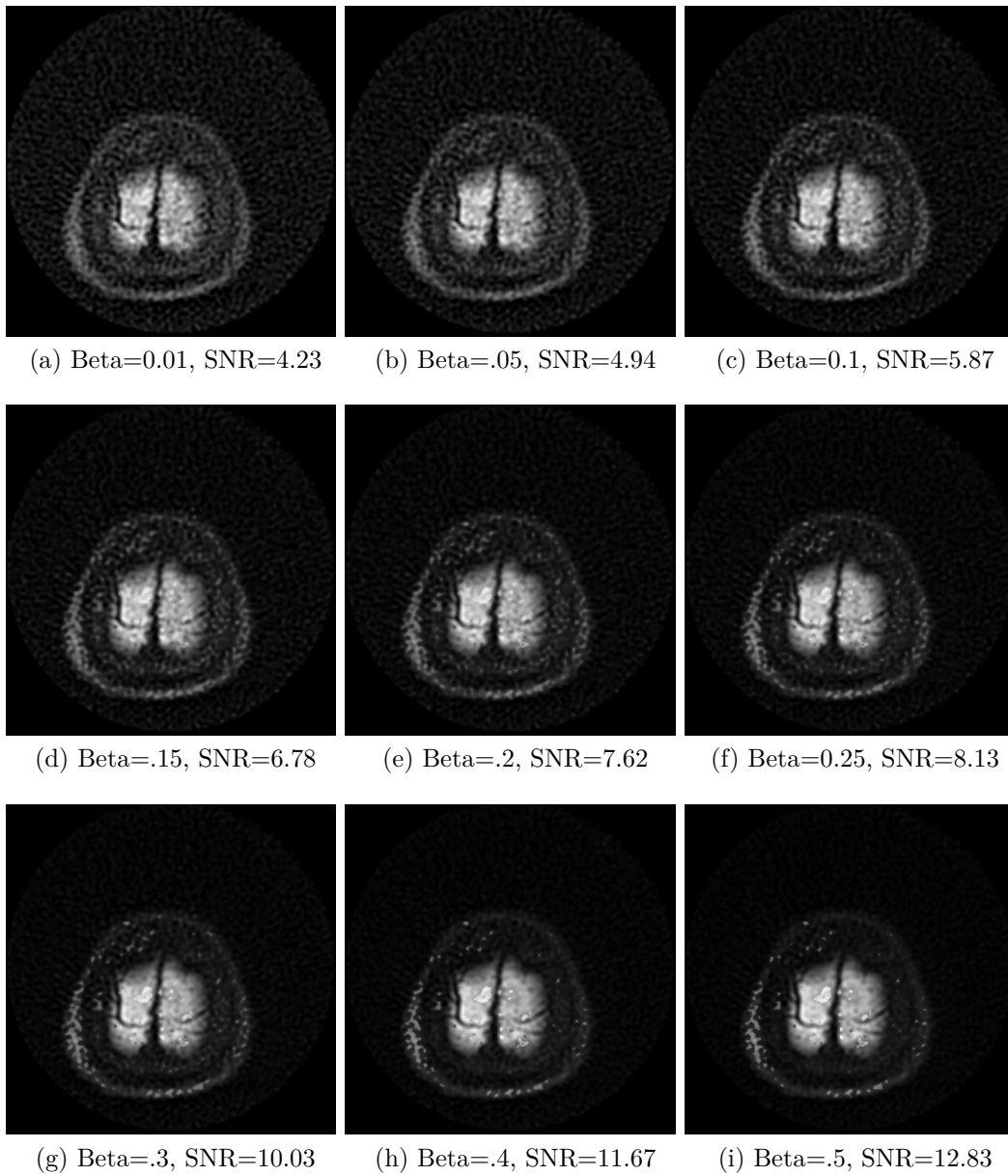


Figure 4.59: Propeller Images Slice5 Results.

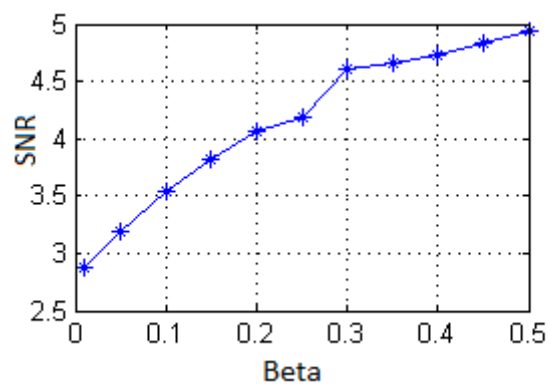


Figure 4.60: Slice 5 SNR vs Beta.

4.3.2.2 Slice 14

The next example is for slice 14 where the algorithm does not perform well. Figure 4.61 shows the b0 image in (a) and the DWI image in (b). The histogram of their difference is shown in (c). The histogram is spread out, and a lot of pixels have difference in the intensity greater than 20%. This indicates that b0 image is not a good prior for this image.

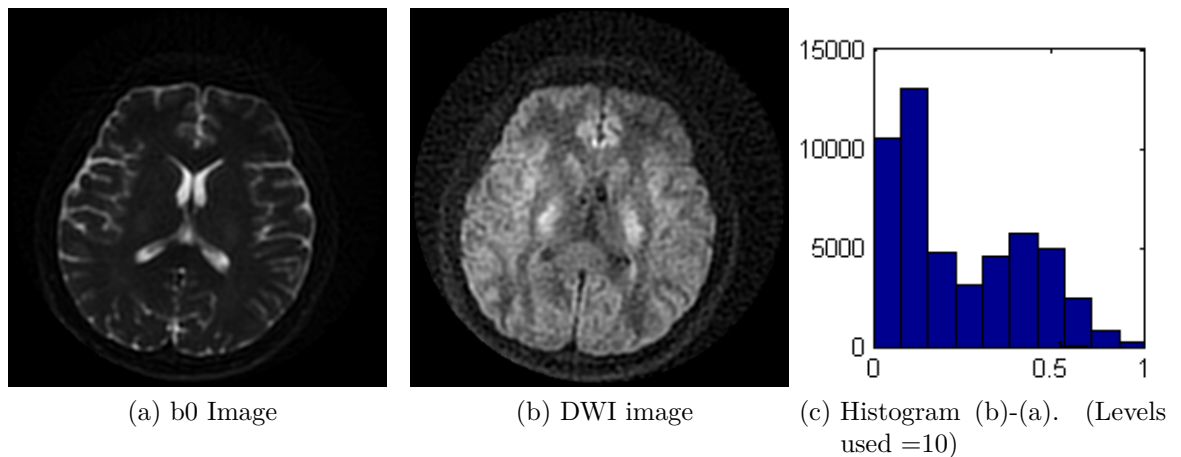


Figure 4.61: Propeller Images Slice 14.

Next we show the reconstructed images generated by the EBA method in Figure 4.62. The algorithm performs poorly and changes the structure of the image. As Beta is increased, the image becomes more like the average of the two images at some pixels where the difference is not very large, and it mimics the prior image at other pixels where the difference is very large.

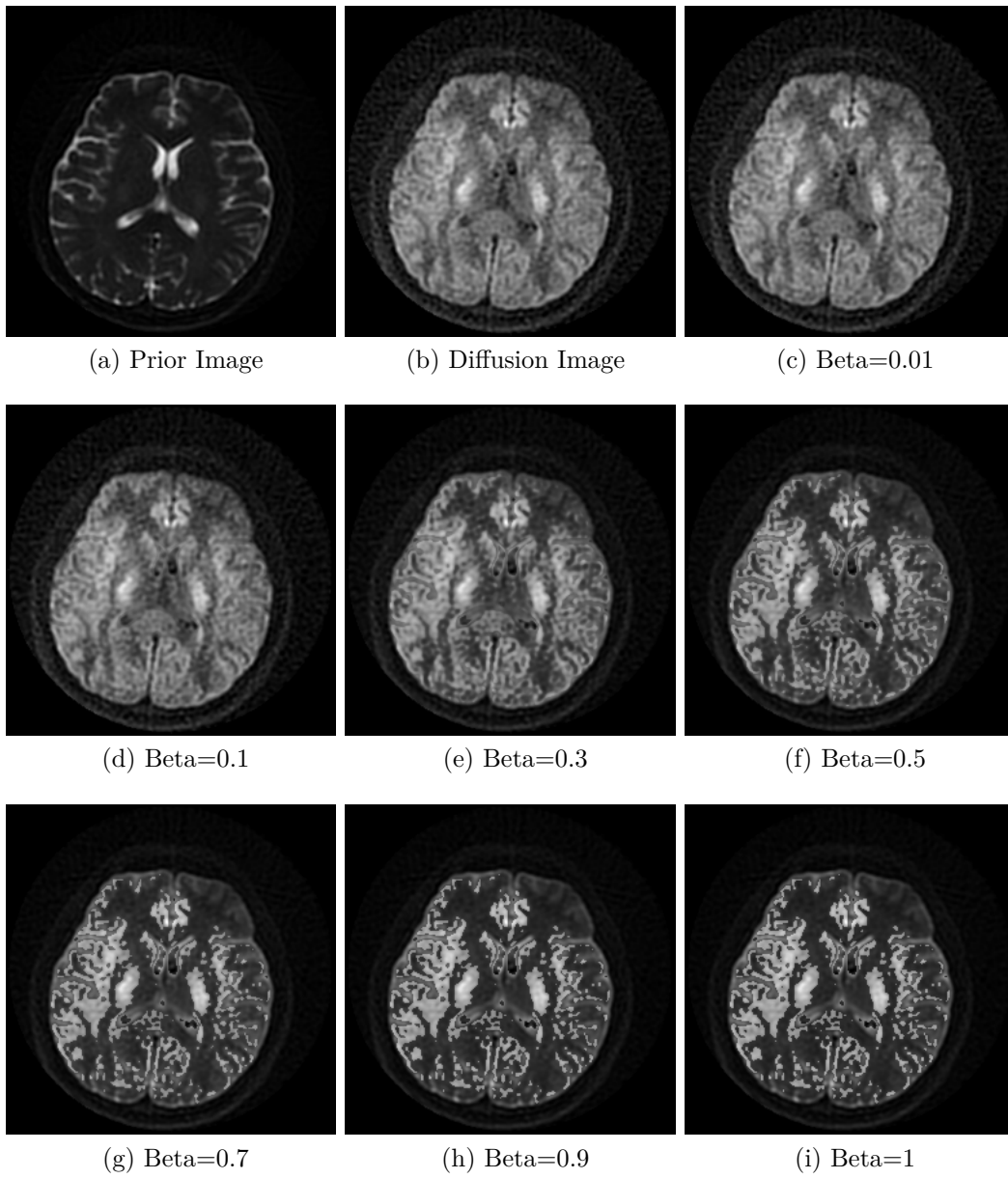


Figure 4.62: Propeller Images Slice 14 Results.

4.3.2.3 Slice 27

In the next example we show the results of slice 27. Figure 4.63 shows the b0 image in (a) and the DWI image (b). The histogram of their difference is shown in (c). It shows that most pixels has less than 20% difference in the intensity and is not spread out. This indicated that the b0 image was a good prior.

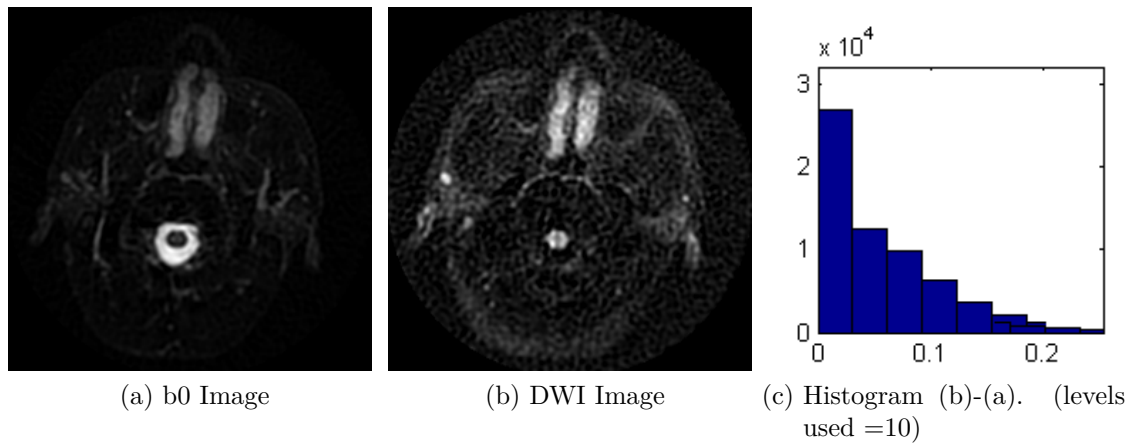


Figure 4.63: Propeller Images Slice27.

In Figure 4.64 we show the results of the noise analysis of this slice. In Figure 4.64, (a) is the DWI image and (c) is the histogram of the noise. The pixels of the noise are used to estimate the mean of a Rician distribution which is fit to the histogram. In Figure 4.64 (b), the image SNR has many pixels with $\text{SNR} > 3$ which validates the Gaussian noise assumption. In Figure 4.64 (d), we show the image SNR histogram. Its first bin lied under the $\text{SNR}=2$, corresponding to the dark spots in the image while the remaining pixels with $\text{SNR} > 2$.

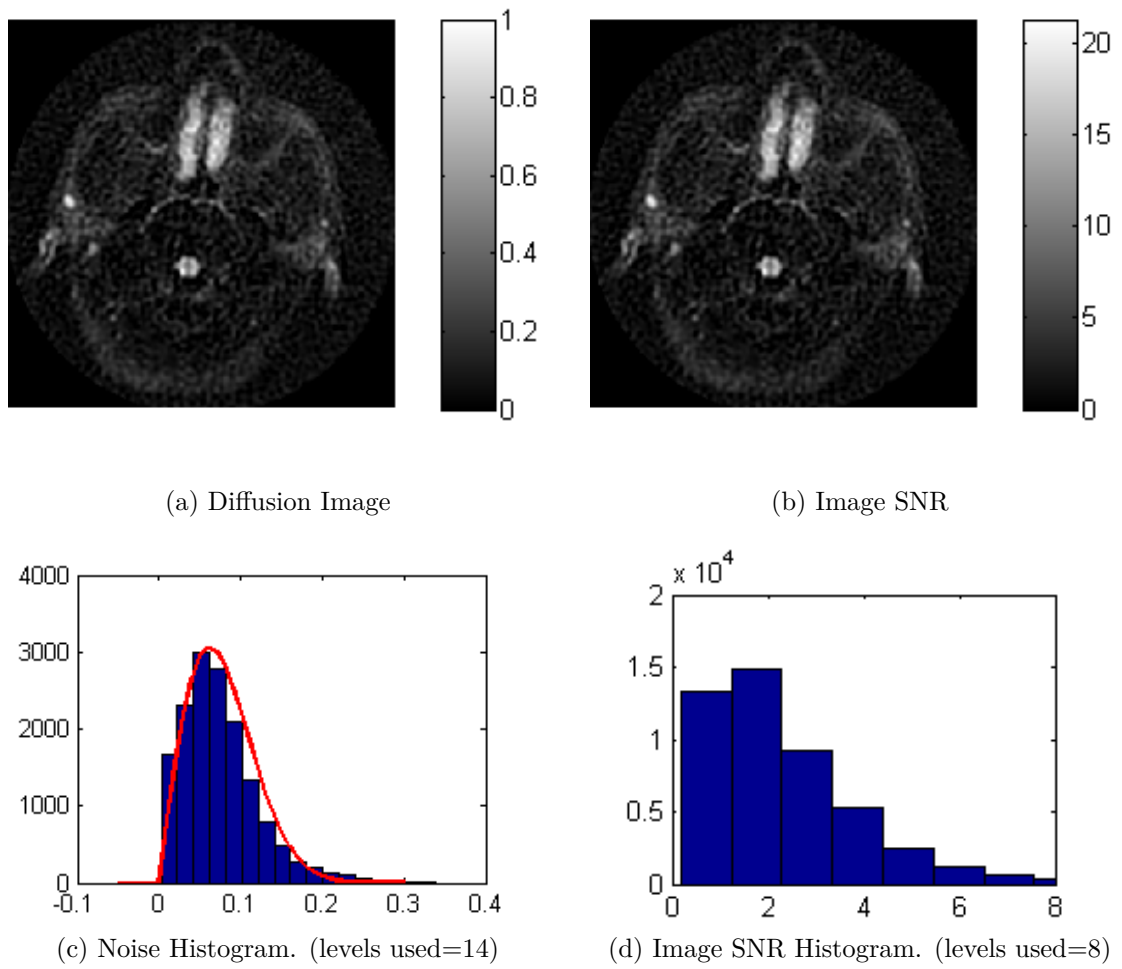


Figure 4.64: Propeller Images Slice 27 Noise Analysis Data.

For the same data set, we show the reconstructed images with the EBA method in Figure 4.65 which again shows the reduction in noise as Beta is increased.

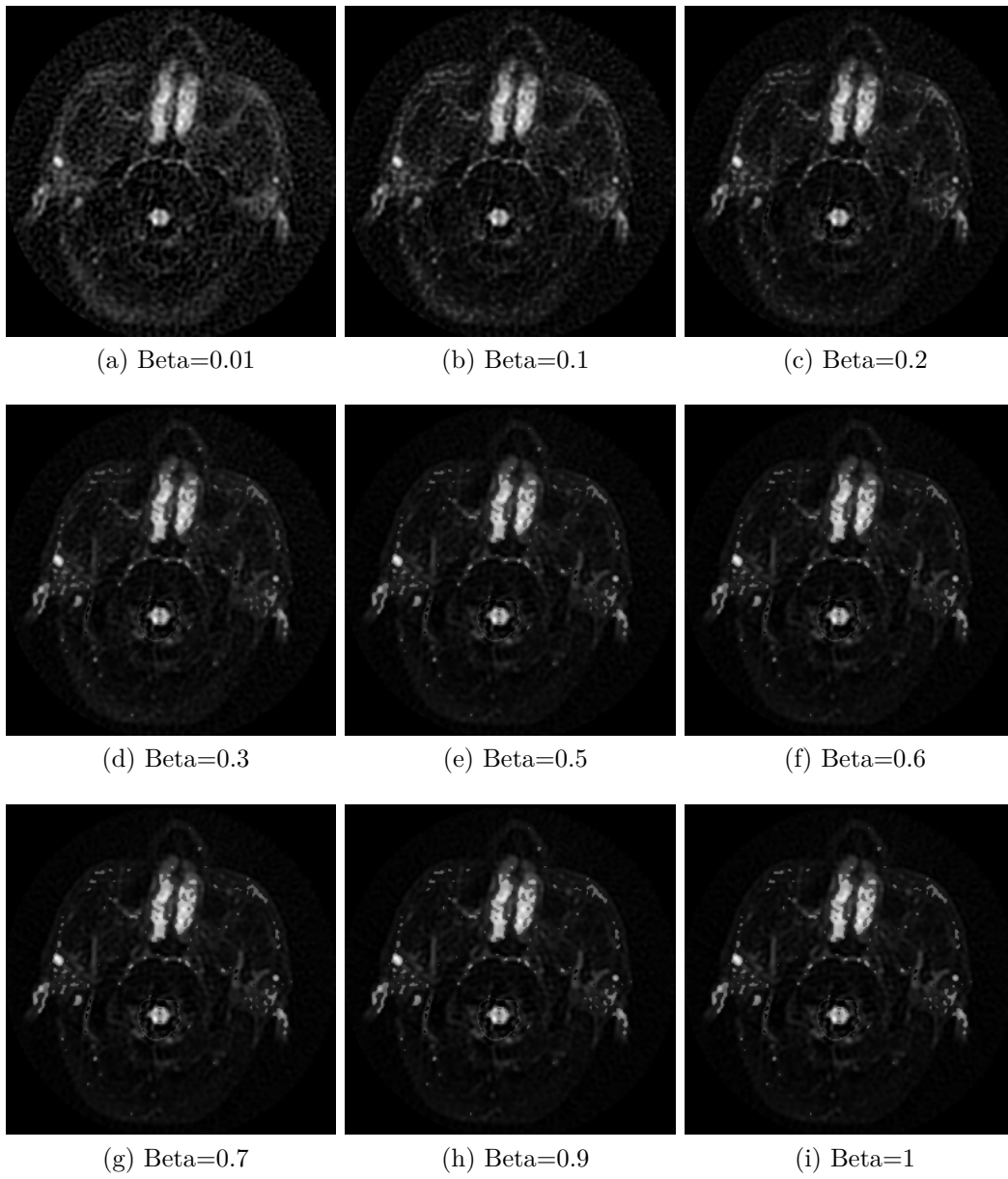


Figure 4.65: Propeller Images Slice 27 Results.

4.3.2.4 Slice 28

Finally, we show the results of slice 28. Figure 4.66 shows the b0 image in (a) and the DWI image (b). Their histogram shown in (c) is not spread out and has most of the pixels with differences less than 20%. This indicates that b0 is a good prior for this image.

In Figure 4.67 we show the results of the noise analysis.

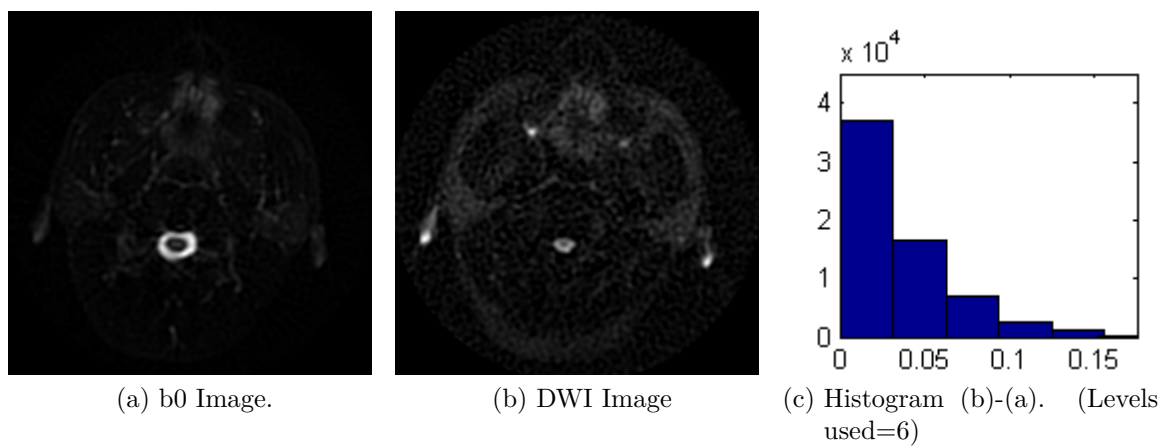


Figure 4.66: Propeller Images Slice 28.

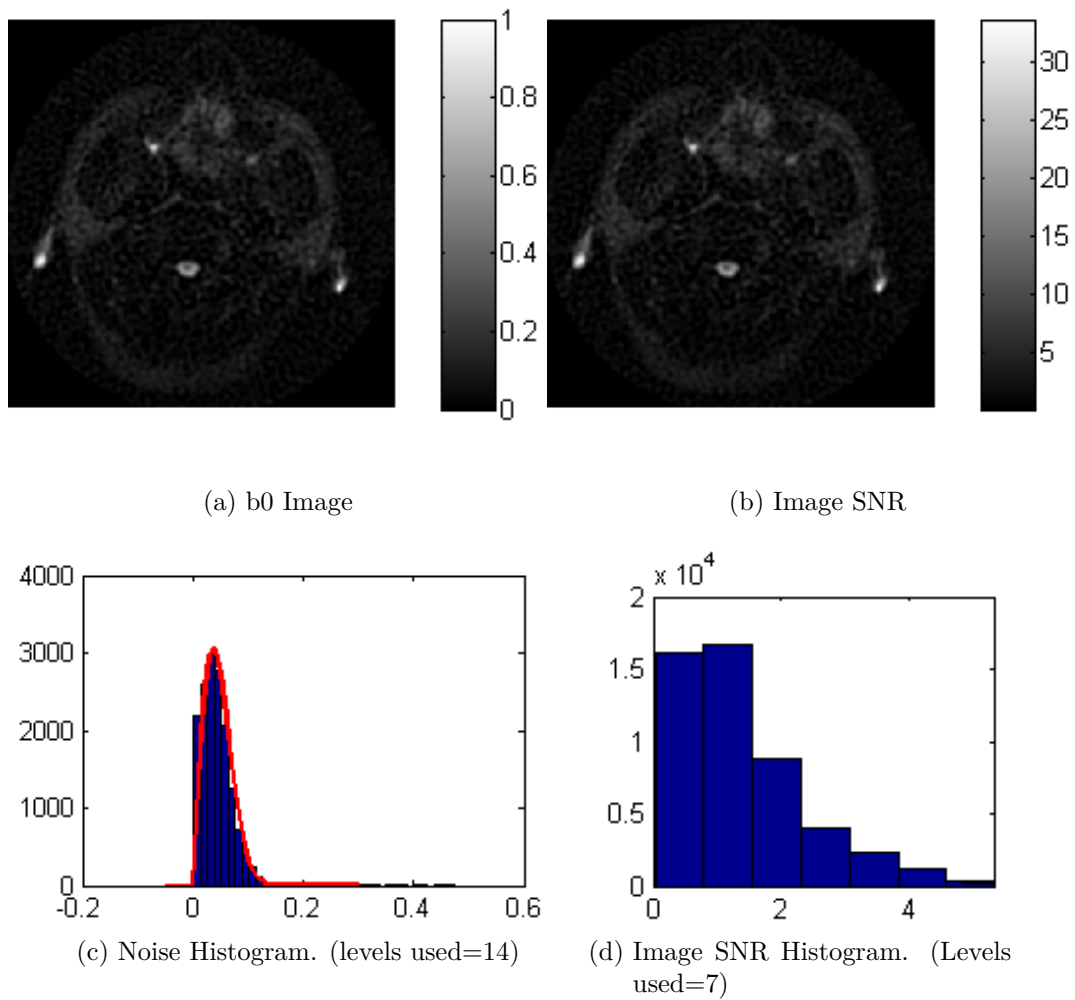


Figure 4.67: Propeller Images Slice 28 Noise Analysis.

We show the DWI Figure 4.67 (a) and its noise histogram in Figure 4.67 (c) is fit to a Rician distribution. The image SNR in Figure 4.67 (b) has many pixels with SNR greater than 3 which enables us to approximate the noise with the Gaussian distribution. The image data SNR histogram in Figure 4.67 (d) further confirms a lot of pixels had $\text{SNR} > 3$. The improvement in the SNRs is demonstrated in Figure 4.68. When beta is equal to 0.3, the SNR increased by 23%. Next we show the reconstructed images for slice 28 generated by the EBA method in Figure 4.69

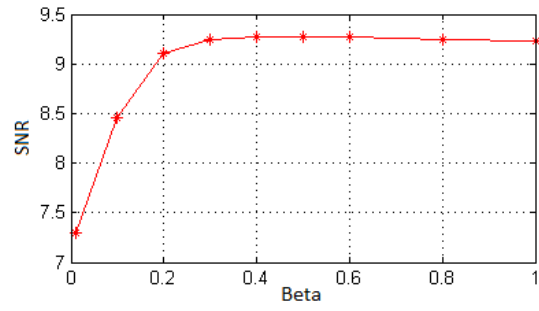


Figure 4.68: Slice 28 SNR versus Beta for Data from Slice 28.

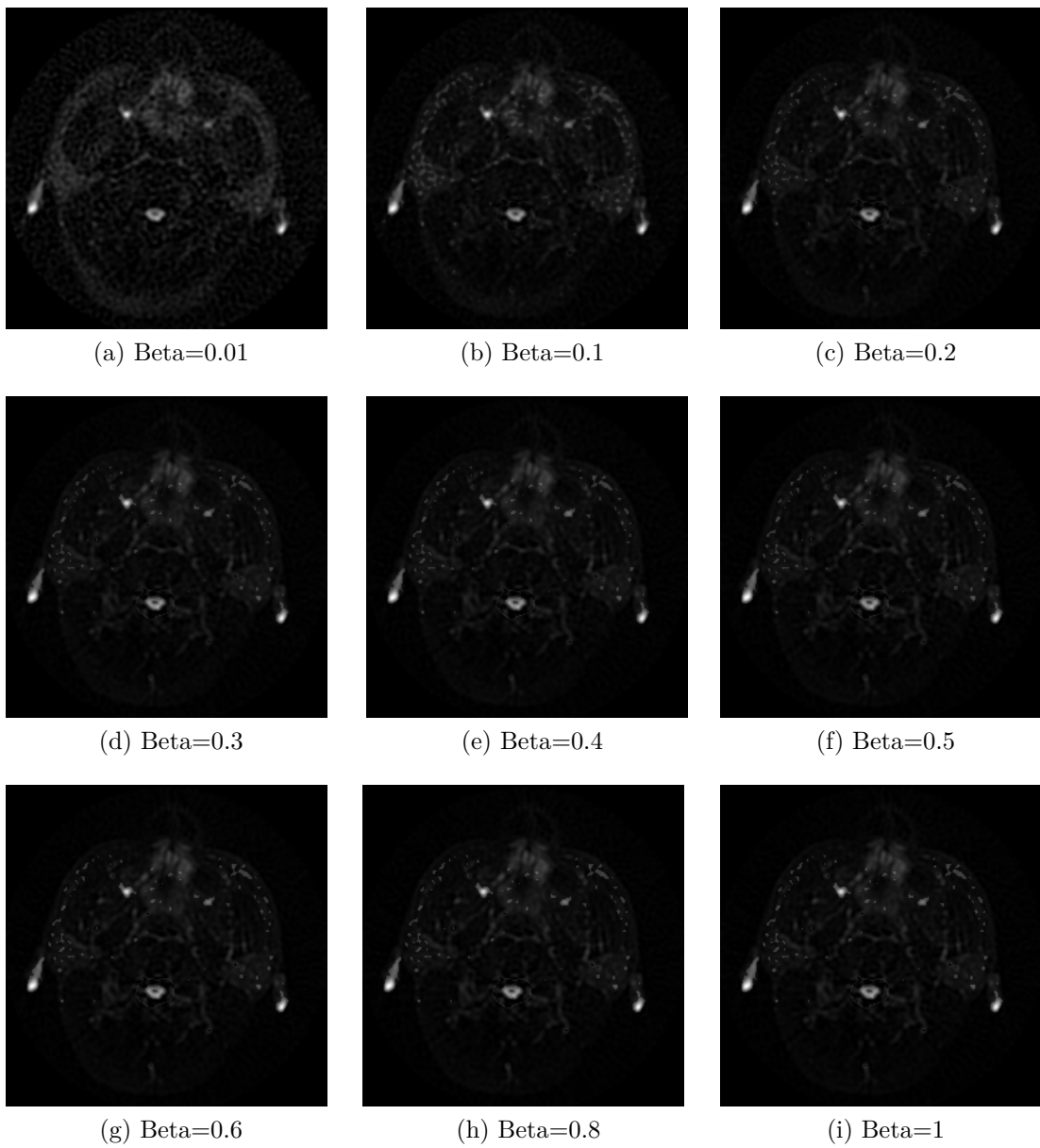


Figure 4.69: Propeller Images Slice 28 Results.

4.3.3 Propeller Data with Variable Beta Value.

In this Section, we propose the use of variable Beta size. Previously, we have showed that the algorithm will perform poorly if the histogram of the difference between the DWI image and the prior has many bins with differences greater than 20%. Under this condition, the algorithm will produce a poorly reconstructed image. In this Section, we investigate the use of variable beta values. The beta value will change for each pixel in the image depending on the difference in the intensity between the corresponding pixels in the DWI and the prior image. If the difference is small, large beta will be used. As the difference increases, beta will be decreased. We have implemented this idea for several slices.

4.3.3.1 Slice 17

We begin with slice 17 where the DWI and the prior are not similar. Figure 4.70 shows the prior image in (a) and the DWI image in (b). The histogram of the difference between the DWI image and the prior image in (c) is spread out and has a lot of pixels with more than 20% difference in the intensity, indicating that the b0 image is not a good prior.

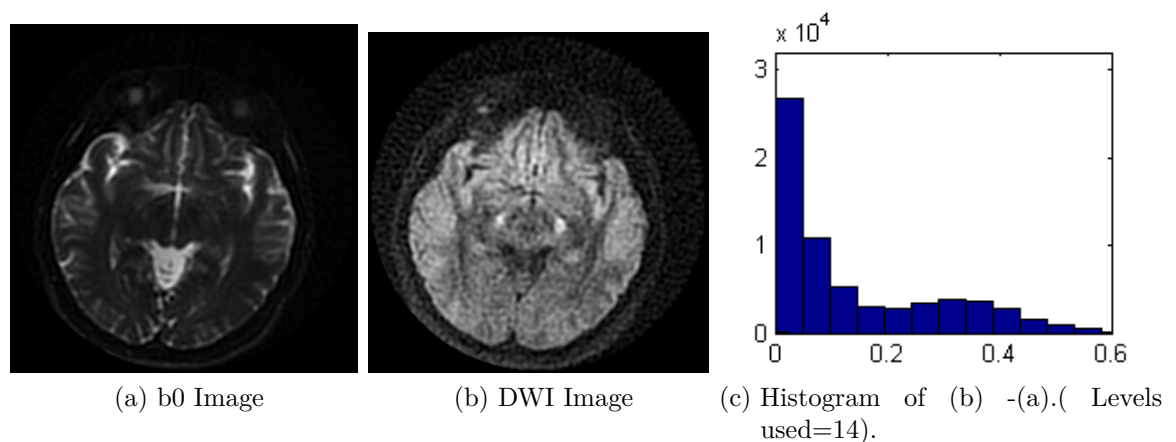


Figure 4.70: Propeller Slice 17 Images.

In Figure 4.71 we show the reconstructed DWI for slice 17 with fixed $\beta=0.4$ in (a) and with variable β in (b). The reconstructed image in Figure 4.71 (a) is obviously poor in quality while the image in 4.71 (b) has preserved the features of the original DWI.

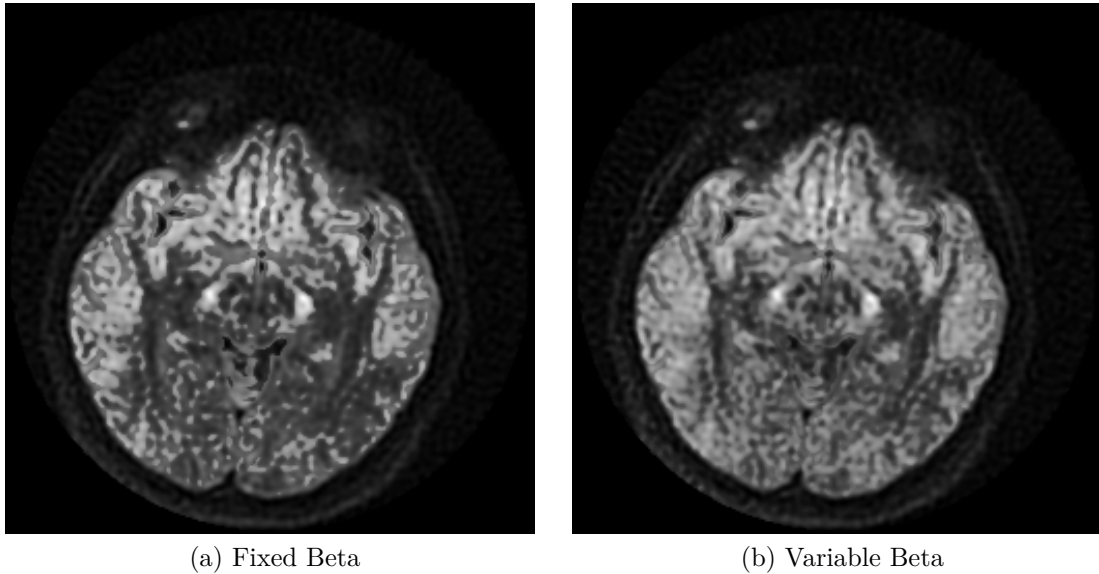


Figure 4.71: Propeller Slice17 Results.

4.3.3.2 Slices 19 and 21

In this Section, we show another two cases where the difference between the DWI and the prior image is large. Figure 4.72 (b) and (e) show the DWI images of slice 19 and slice 21 respectively. Their prior images are displayed in Figure 4.72 (a) and (d). The histograms of the difference image between the prior and the DWI in Figures (c) and (f) are spread out have many pixels with difference greater than 20%. This indicates that the b0 images are not good priors.

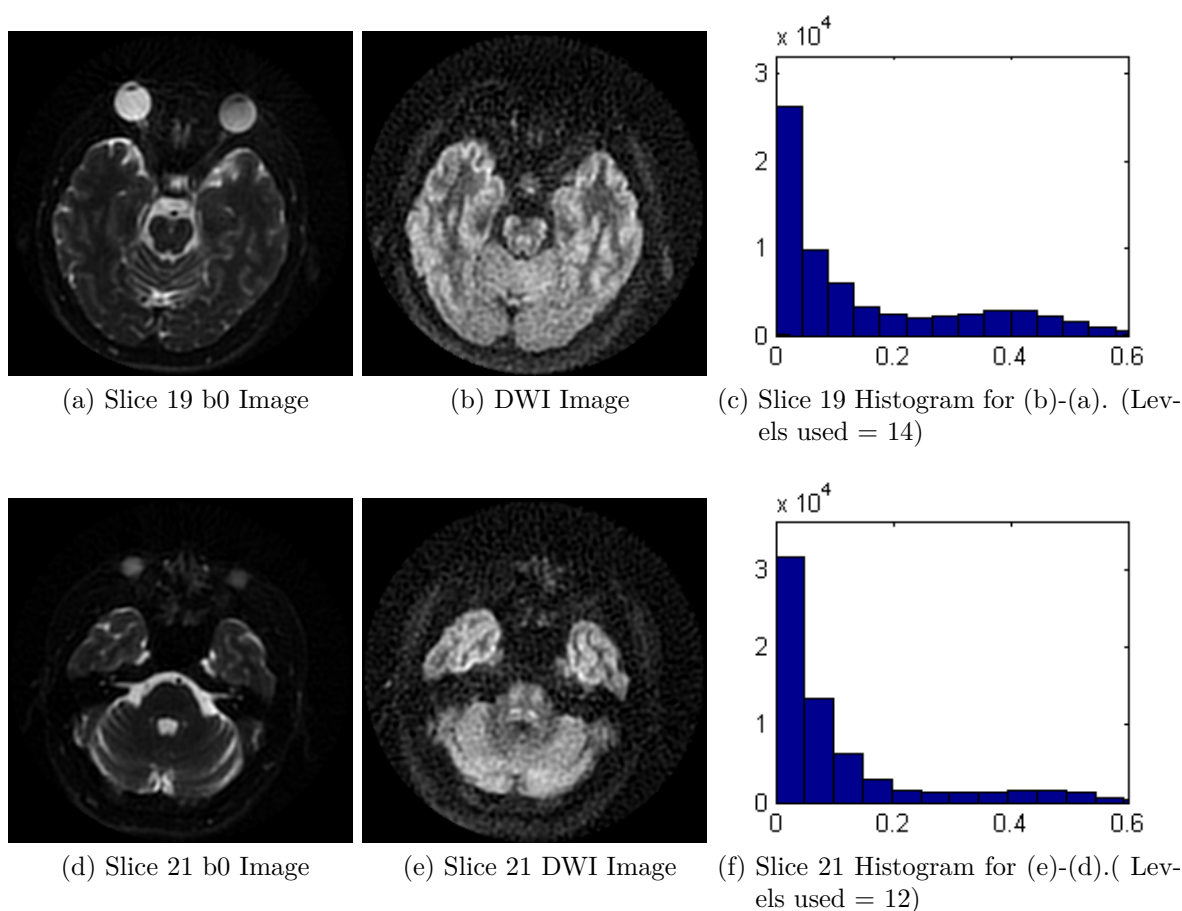
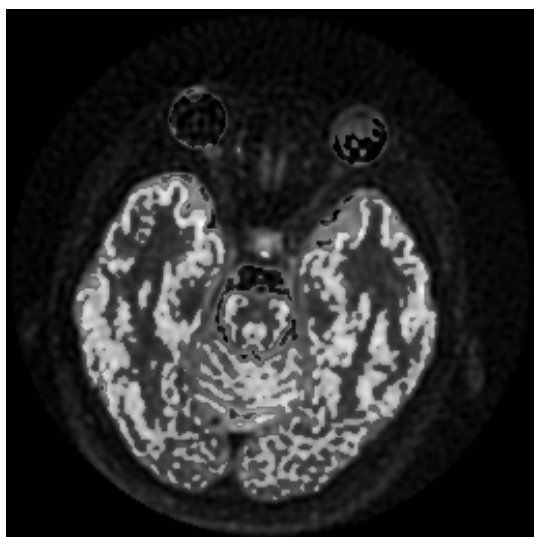


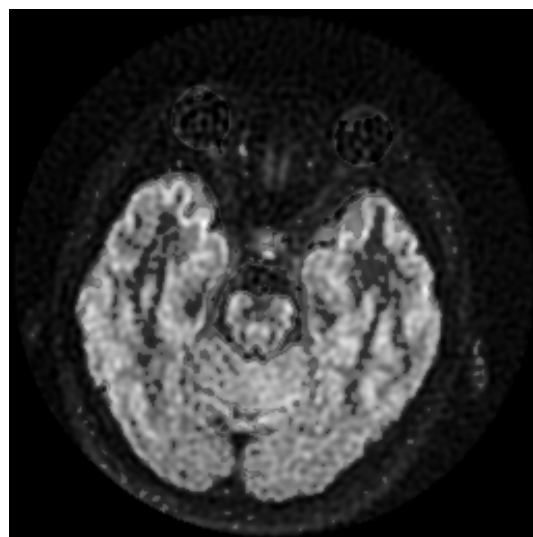
Figure 4.72: Propeller Slices 19 and 21 Images.

In Figure 4.73, we show the reconstructed images with fixed beta and variable beta. In Figure 4.73 (a), the image of slice 19 reconstructed with fixed beta= 0.6 is distorted by the algorithm. In Figure 4.73 (b) the same image is reconstructed but using variable Beta. The image is not distorted severely by the

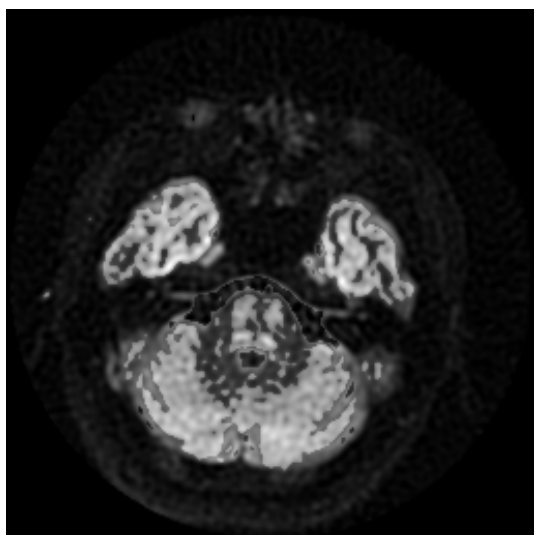
reconstruction algorithm and is close to the original DWI. Similar observation can be seen in the data for slice 21 in Figures 4.73 (c) and 4.73 (d).



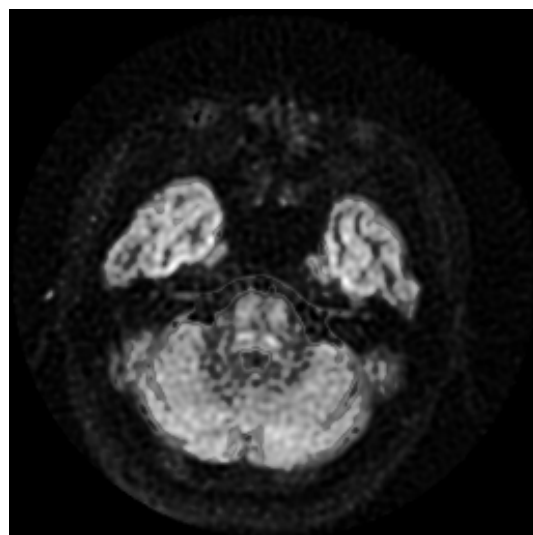
(a) Slice 19 DWI Image with Fixed Beta = 0.3



(b) Slice 19 DWI Image with Variable Beta



(c) Slice 21 DWI Image with Fixed Beta = 0.3



(d) Slice 21 DWI Image with Variable Beta

Figure 4.73: Propeller Slice19 and Slice 21 Results.

4.3.3.3 Slice 28

Finally we consider a case where the difference between the DWI and the prior was less than 20% for most of the pixels. Figure 4.74 shows data for slice 28 the prior image in (a), the DWI in (b) and the histogram of their difference in (c). The histogram in Figure 4.74 (c) is not spread out with most pixels have difference less than 20%. Consequently, the reconstructions with fixed $\beta = 0.4$ and variable β in Figure 4.75 do not have any difference since the algorithm has adjusted the variable β to the same fixed value.

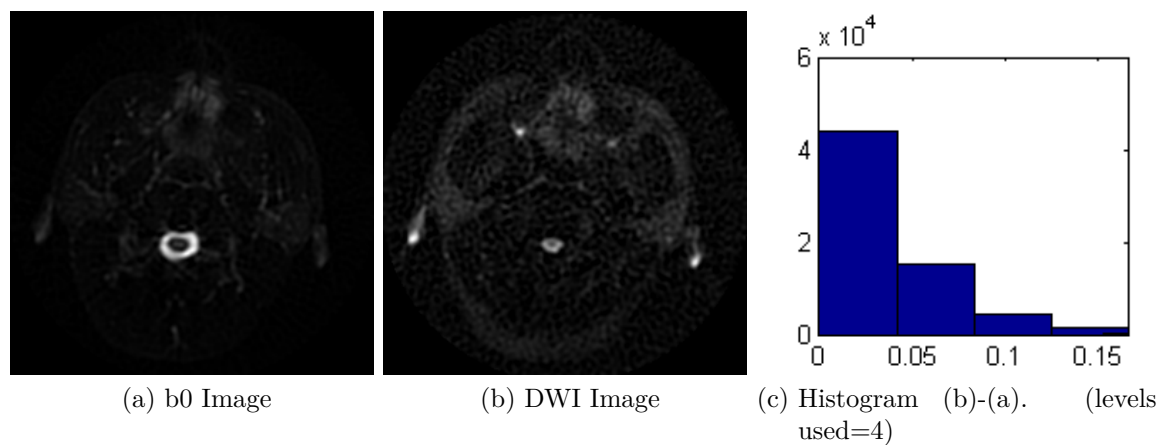


Figure 4.74: Propeller Slice 28 Images.

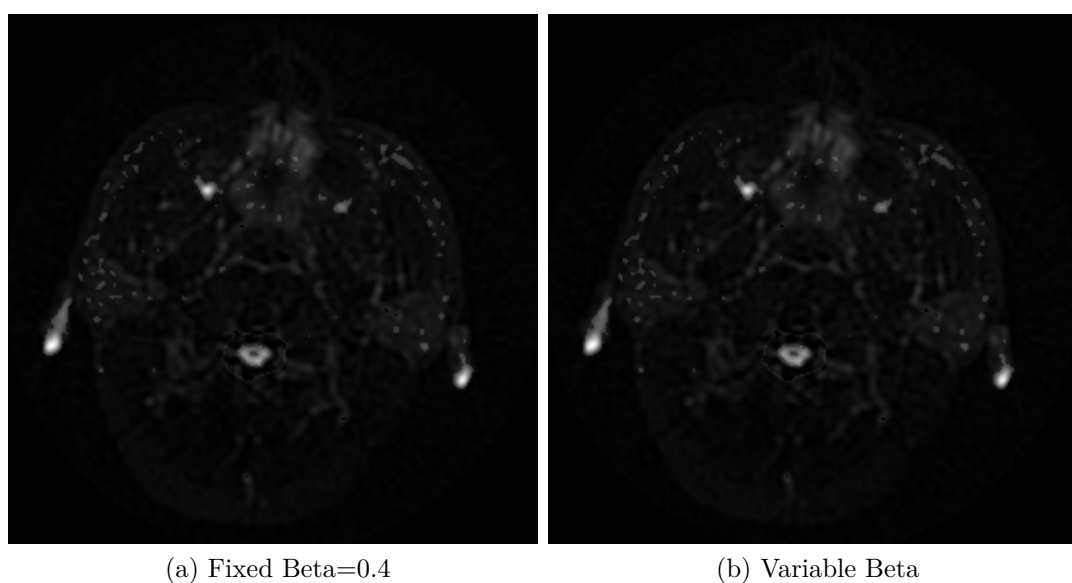


Figure 4.75: Propeller Slice 28 Variable Beta

Chapter 5

Conclusion and Future Work

In this thesis, we have explained the concepts behind DWI and some of its applications. DWI images are corrupted with noise. The main goal of this thesis is to enhance DWI image quality. In the first chapter we discuss the concepts of DWI and the measurement procedure for image generation. We also describe some of the diffusion parameters such as the b-value. Then we examine the noise distribution in DWI which has been shown to be Rician. The Rician distribution can be approximated by a Gaussian distribution when the SNR is high enough. Particularly, we demonstrate that when the $SNR > 3$, the Gaussian approximation will be valid.

We then describe three contemporary denoising methods for DWI images. The first method has been proposed by Guleryuz (2007). It makes use of the sparseness in the image representation in the DCT to remove the noise. Manjón and Coupé (2013) have proposed the second method NLM which makes use of the self-similarity in the image for filtering. The last method has been proposed by Manjón and Pierrick Coupe. It depends on the PCA of image slices.

In Chapter two, we explain the Bayesian Approach in denoising images. The Bayesian approach makes use of another image to improve the reconstruction of the DWI images. In MRI scanners, each time a scanner is used to produce a diffusion image at a certain b-value while it will also generate another image that has no diffusion. This image is called the b0 image. The b0 image has less noise than the DWI image and hence can be used as a prior to denoise the DWI image.

Since the b0 image is used in our Bayesian approach, we need to define the joint entropy between the b0 image and the DWI image and demonstrate that it can be used as a measure of similarity between two images. The closer is the DWI image to the b0 image, the smaller the joint entropy is. Joint entropy is used as a regularization term in the image reconstruction, and hence we called the algorithm the Entropy Bayesian Approach (EBA).

In chapter three, we compare the performance of the EBA algorithm to that of the other three denoising methods discussed in chapter one. The processing time of the EBA is about the same as that of the NLM and PCA, but less than that of the DCT algorithm. The EBA and the PCA have performed the best in keeping the fine details of the image.

In Chapter four, we have validated the EBA algorithm performance with the simulation and experimental results. Simulations have been performed using prior images with added Gaussian noises at different levels as the DWI images. We then apply EBA and observe the effect of beta on the results. For the experimental validation, we have used real DWI intensity images. These images are generated from phantom and clinical images that are obtained from GE healthcare.

We have used two different phantoms to produce the DWI images. These images are collected at different b-values. The noise distribution is empirically studied. We have measured the changes in the noise distribution with the b-value.

For low b-values, since the SNR is higher than 3 and the noise can be assumed to have Gaussian distribution. When the b-value is higher than 2500, the SNR becomes less than 3 and the noise assumes Rician distribution.

The Bayesian approach performs well if the prior image is similar to the DWI image. However, the b0 image is not always similar to the DWI image. When they are different, the EBA algorithm fails to improve the DWI image significantly. We use the histogram of the difference between the DWI and the b0 image as a measure of similarity. We verify the performance of the EBA algorithm with respect to this similarity and find a threshold for the algorithm effectiveness. A 20% intensity difference between the DWI and the b0 image is chosen to be the threshold for the algorithm to improve the DWI. When the difference in intensity is higher than 20%, we vary the value of beta according to the intensity difference. Comparing to the results from the fixed beta size, we illustrate the variable beta can improve performance for high intensity difference.

The EBA algorithm is also validated on real clinical images. We apply it to 28 image slices. For some slices, when the b0 image is similar to the DWI image (such as in slices 27 and 28), the algorithm performs very well. When the two images are not very similar (such as slice 14), the algorithm corrupts the original image with the prior image. To improve the performance of the EBA, we adopt variable beta approach. If the difference between the corresponding pixels in the b0 image and the DWI is less than 20%, beta is fixed to .5. If this difference is larger than 20%, beta is reduced proportional.

The limitation of the EBA method is that it requires the right prior. This means that the prior image, i.e the b0 image, should be similar to the DWI. Otherwise, the algorithm cannot improve the quality of the DWI significantly.

For future work, the algorithm will be tested with complex images, which is the case in MRI images. Unlike image intensity which has Rician noise, complex images have Gaussian noise and hence no need to approximate the noise distribution. This is impossible for the current experimental work since GE can only provide intensity images but not complex images.

So far, the size of beta is fixed for the images for the current EBA algorithm. Further implementation will be done to automatically choose the value of beta based on the noise level and difference between the DWI and the b0 image.

To improve the performance of the EBA algorithm, the prior can be pre-processed with a low pass filter to smooth the image before it is used in the EBA algorithm. We can also combine this method with the contemporary methods especially the PCA algorithm. The denoised image from the PCA can be used as the prior image, replacing the b0 image. The optimization can also be done using a block around each pixel rather than the pixel itself. This will have the advantage of enabling better estimate of the noise around the pixel and hence, an accurate measure of the true pixel value can be obtained.

References

- [1] Roland Bammer. Basic Principles of Diffusion-Weighted Imaging. *European journal of radiology*, 45(3):169–184, 2003.
- [2] Mokhtar S Bazaraa, Hanif D Sherali, and Chitharanjan Marakada Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley & Sons, 2013.
- [3] Christian Beaulieu. The Basis of Anisotropic Water Diffusion in the Nervous System—a Technical Review. *NMR in Biomedicine*, 15(7-8):435–455, 2002.
- [4] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. A Review of Image Denoising Algorithms, with a New One. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.
- [5] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. Wiley-Interscience, 2006.
- [6] Yu-Hong Dai. Nonlinear Conjugate Gradient Methods. *Wiley Encyclopedia of Operations Research and Management Science*, 2011.
- [7] Gonzalez. *Digital Image Processing Using MATLAB*. New Jersey, Prentice Hall., 2003.

- [8] Håkon Gudbjartsson and Samuel Patz. The Rician Distribution of Noisy MRI Data. *Magnetic Resonance in Medicine*, 34(6):910–914, 1995.
- [9] Onur G Guleryuz. Weighted Averaging for Denoising with Overcomplete Dictionaries. *Image Processing, IEEE Transactions on*, 16(12):3020–3034, 2007.
- [10] Mark Holden, Derek LG Hill, Erika RE Denton, Jo M Jarosz, Tim CS Cox, Torsten Rohlfing, Joanne Goodey, and David J Hawkes. Voxel Similarity Measures for 3-D Serial MR Brain Image Registration. *Medical Imaging, IEEE Transactions on*, 19(2):94–102, 2000.
- [11] Keyanoosh Hosseinzadeh and Samuel David Schwarz. Endorectal Diffusion-Weighted Imaging in Prostate Cancer to Differentiate Malignant and Benign Peripheral Zone Tissue. *Journal of Magnetic Resonance Imaging*, 20(4):654–661, 2004.
- [12] Dow-Mu Koh and David J Collins. Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. *American Journal of Roentgenology*, 188(6):1622–1635, 2007.
- [13] Thomas C Kwee, Taro Takahara, Reiji Ochiai, Rutger AJ Nieuvelstein, and Peter R Luijten. Diffusion-Weighted Whole-Body Imaging with Background Body Signal Suppression (DWIBS): Features and Potential Applications in Oncology. *European radiology*, 18(9):1937–1952, 2008.
- [14] Denis Le Bihan, Jean-François Mangin, Cyril Poupon, Chris A Clark, Sabina Pappata, Nicolas Molko, and Hughes Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of magnetic resonance imaging*, 13(4):534–546, 2001.
- [15] Frederik Maes, Andre Collignon, Dirk Vandermeulen, Guy Marchal, and Paul

- Suetens. Multimodality image registration by maximization of mutual information. *Medical Imaging, IEEE Transactions on*, 16(2):187–198, 1997.
- [16] Ashkan A. Malayeri, Riham H. El Khouli, Atif Zaheer, and Michael A. Jacobs. Principles and Applications of Diffusion-weighted Imaging in Cancer Detection, Staging, and Treatment Follow-up. *RadioGraphics 2011*, 31(6):1773–1791, 2011.
- [17] José V Manjón, Pierrick Coupé, Luis Concha, Antonio Buades, D Louis Collins, and Montserrat Robles. Diffusion Weighted Image Denoising Using Overcomplete Local PCA. *PloS one*, 8(9):e73021, 2013.
- [18] V. Manjón, José, Pierrick Coupé, Antonio Buades, D. Louis Collins, and Montserrat Robles. New Methods for MRI Denoising based on Sparseness and Self-Similarity. *Medical Image Analysis*, 16(1):18–27, January 2012.
- [19] Robert D Nowak. Wavelet-based Rician Noise Removal for Magnetic Resonance Imaging. *Image Processing, IEEE Transactions on*, 8(10):1408–1419, 1999.
- [20] Johan Nuyts. The Use of Mutual Information and Joint Entropy for Anatomical Priors in Emission Tomography. 6:4149–4154, 2007.
- [21] James G Pipe et al. Motion Correction with PROPELLER MRI: Application to Head Motion and Free-Breathing Cardiac Imaging. *Magnetic Resonance in Medicine*, 42(5):963–969, 1999.
- [22] Josien PW Pluim, JB Antoine Maintz, and Max A Viergever. Mutual-Information-Based Registration of Medical Images: a Survey. *Medical Imaging, IEEE Transactions on*, 22(8):986–1004, 2003.
- [23] Jing Sheng, Dong Liang, Jing Tang, Donglai Huo, and Lei Ying. Bayesian

- Reconstruction of DW-PROPELLER Images Using Joint Entropy. Proceedings of International Society of Magnetic Resonance in Medicine Scientific Meeting, :2861, 2009.
- [24] Stephen M Smith, Mark Jenkinson, Mark W Woolrich, Christian F Beckmann, Timothy EJ Behrens, Heidi Johansen-Berg, Peter R Bannister, Marilena De Luca, Ivana Drobnyak, David E Flitney, et al. Advances in Functional and Structural MR Image Analysis and Implementation as FSL. *Neuroimage*, 23:S208–S219, 2004.
- [25] Laurent Sorber, Marc Van Barel, and Lieven De Lathauwer. Unconstrained Optimization of Real Functions in Complex Variables. *SIAM Journal on Optimization*, 22(3):879–898, 2012.
- [26] Stejskal and Tanner. Spin Diffusion Measurements: Spin Echoes in the Presence of Time-Dependent Field Gradient. *Journal of Chemical Physics*, 42(1):288–292, 1965.
- [27] Jing Tang and Arman. Rahmim. Bayesian PET Image Reconstruction Incorporating Anato-Functional Joint Entropy. *Physics in medicine and biology*, 1:123–140, 2009.
- [28] Christopher G Thomas, Richard A Harshman, and Ravi S Menon. Noise Reduction in BOLD-Based fMRI using Component Analysis. *Neuroimage*, 17(3):1521–1537, 2002.
- [29] Wolfgang Wein, Shelby Brunke, Ali Khamene, Matthew R Callstrom, and Nassir Navab. Automatic CT-Ultrasound Registration for Diagnostic Imaging and Image-Guided Intervention. *Medical image analysis*, 12(5):577–585, 2008.
- [30] Nicolas Wiest-Daesslé, Sylvain Prima, Pierrick Coupé, Sean Patrick Morrissey, and Christian Barillot. Rician Noise Removal by Non-Local Means Filter-

- ing for Low Signal-to-Noise Ratio MRI: Applications to DT-MRI. 1:171–179, 2008.
- [31] Gaohang Yu, Yanlin Zhao, and Zengxin Wei. A Descent Nonlinear Conjugate Gradient Method for Large-Scale Unconstrained Optimization. *Applied mathematics and computation*, 187(2):636–643, 2007.
- [32] Yang-Ming Zhu. Volume Image Registration by Cross-Entropy Optimization. *Medical Imaging, IEEE Transactions on*, 21(2):174–180, 2002.
- [33] Barbara Zitova and Jan Flusser. Image Registration Methods: a Survey. *Image and vision computing*, 21(11):977–1000, 2003.

Mustafa Farrah Curriculum Vitae

mfarah@uwm.edu

- Education:
 - PhD in Electrical Engineering, May 2015, GPA: 3.72/4.0 University of Wisconsin-Milwaukee – Milwaukee, WI – USA
 - MS in Electrical Engineering, May 2004, GPA: 3.67/4.0 University of Wisconsin-Milwaukee – Milwaukee, WI - USA
 - BS in Electronics Engineering, June 2000, GPA: 76/100 Al-Quds University – Jerusalem, Israel Dissertations
- PhD dissertation title: “New Image Reconstruction Techniques of DWI MRI Images Using Bayesian Approach.”
- MS dissertation title: “Statistical Behavior of Artificial Neural Networks.”
- Computer Skills: UNIX, LYX, Latex, Fortran, C/C++ language, Assembly Language, Matlab, Mathcad, Pspice, Microsoft Office, Windows, DOS, System View, NeuroSolution.
- Teaching Experience
 - University of Wisconsin-Milwaukee. Milwaukee, WI. USA
 - Lecturer: Analytical and Mathematical Methods in Engineering, Sep 2010-Present, Jan 2008-May 2008.
 - Lecturer: Electrical Circuits 2, July 2010-Aug 2010, July 2009- Aug 2009, July 2008- Aug 2008, July 2006-Aug 2006.

- Teaching Assistant: Analytical and Mathematical Methods in Engineering, Sep 2003- Jan 2004.
- Teaching Assistant: Electro-mechanic Energy Conversions, Sep 2002-May 2003.
- Teaching Assistant: Microprocessors, Sep 2002- Jan 2004. University of Wisconsin-Madison, Madison, WI. USA
- Lab Instructor: Advanced Electrical Engineering Lab, Sep 2008-May 2009
- Wajdi Institute of Technology, Jerusalem, Israel.
- Lecturer: Introduction to Electrical Engineering. Sep 2005- Jan 2006. Palestine Polytechnic University, Hebron, West Bank.
- Lecturer: Introduction to Electrical Engineering. Sep 2005- Jan 2006. Signals and data communications Introduction to Electrical Circuits Analysis Introduction to Microprocessors Assembly Language and computer Architectures
- Research Experience: University of Wisconsin-Milwaukee. Milwaukee, WI. USA
 - Investigating new image reconstruction methods of Diffusion weighted images using Bayesian approach and implementation of the methods in Matlab. “ Cooperating with GE “ Project Assistant: Estimation of head position and movement tracking of head in MRI machines using Retro grate reflector targets and Kalman filter using Matlab programming. Sep 2006-June 2007.
 - Conducted research on classification problems in the investigation and

comparison of different neural network problems and comparing different training methods. Sep 2003- May 2004.

- Publications: In Progress: Bayesian image reconstruction of DWI images using joint entropy and T2 weighted images as a prior.
- O'Connor KM, Armstrong BSR, Watts SR, Farrah MMA, Bottum MC. An innovative diagnostic tool for reducing traumatic knee injuries. In Proceedings of the 31st Annual Meeting of the American Society of Biomechanics; Palo Alto, CA. 2007.
- Awards: Chancellor Award Recipient for the following years: 2003, 2006, 2007, 2008, 2009, 2010
- Languages: Arabic, English, and Hebrew.