Wavelet-Based Rician Noise Removal for Magnetic Resonance Imaging

Robert D. Nowak

Department of Electrical Engineering
Michigan State University
East Lansing, MI 48824–1226
Fax: (517) 353–1980
Email: nowak@egr.msu.edu
Web: http://www.egr.msu.edu/~nowak/

Submitted to *IEEE Transactions on Image Processing*
August 1997

Revised January 1998

EDICS Numbers: IP 2.4 Magnetic Resonance Imaging, 1.6 Multiresolution Processing

Abstract — It is well-known that the noise in magnetic resonance magnitude images obeys a Rician distribution. Unlike additive Gaussian noise, Rician noise is signal-dependent and consequently separating signal from noise is a difficult task. Rician noise is especially problematic in low signal-to-noise ratio (SNR) regimes where it not only causes random fluctuations, but also introduces a signal-dependent bias to the data that reduces image contrast. This paper studies wavelet-domain filtering methods for Rician noise removal. We derive a novel wavelet-domain filter that adapts to variations in both the signal and the noise. The new wavelet-domain filter reduces Rician noise contamination in both high and low SNR regimes.

I. Introduction

In magnetic resonance imaging (MRI), there is an intrinsic trade-off between the signal-to-noise ratio (SNR), spatial resolution, and acquisition time required by the intended clinical/research application [1]. Therefore, given physiological or research paradigm constraints, achievable SNR can be limited. The SNR in most clinical applications of MRI is relatively high. This is accomplished, explicitly or implicitly, by averaging. Two types of averaging take place in MRI data acquisition:

1. The discrete nature of the data acquisition process causes spatial volume averaging.

Permission to publish abstract separately is granted.

Supported by the National Science Foundation, grant no. MIP–9701692, and Michigin State University, research initiation grant no. 71–4564.

2. In some applications it is common to acquire several measurements at the same k-space[1] location and average them to reduce noise.

The two averaging operations are interrelated. Higher resolution images can be obtained by higher sampling in k-space. However, achieving a desired SNR at higher spatial resolutions may require additional time-averaging, and consequently a longer acquisition time. The acquisition time (and hence SNR and resolution) is limited in practice due to practical constraints such as patient comfort and system throughput, and by physiological constraints arising in dynamic applications such as lung and cardiac imaging.

Roughly speaking, we can divide MRI into two regimes: relatively low resolution[2], high SNR imaging and high resolution, low SNR imaging. The high SNR enjoyed in many applications comes at the expense of limited spatial and/or temporal resolution. Higher resolution imaging is possible, but only at a cost of lower SNR or much longer data acquisition times. The ideas in the paper are motivated by interest in high resolution (spatial and/or temporal) MRI for applications such as angiography, functional MRI, and myocardial perfusion imaging. The relevance of this problem is also attested to by the fact that low SNR is the subject of a large number of papers in the MRI literature, *e.g.,* [2, 3, 4, 5, 6, 7, 8, 9, 10]. Hence, we are interested in studying noise reduction methods for MRI.

Magnetic resonance (MR) image reconstruction data are commonly modeled by the Rician distribution [7, 5], and we assume this model throughout the paper, unless otherwise noted. The Rician model is widely accepted in the MRI community, and we have experimentally verified that it does indeed provide an excellent match to the observations.

Rician noise causes random fluctuations in the data and introduces a bias to the MR image that reduces image contrast, as demonstrated in Sections IV and V. The Rician noise is a signal-dependent noise, rather than a simple additive noise. In low intensity regions of the magnitude image the noise distribution tends to the Rayleigh distribution. In regions of high intensity, the noise tends to a Gaussian distribution. Because the noise is signal-dependent, noise removal is very difficult in MRI, especially at low SNRs. Not only must we contend with the usual problems associated with noise reduction such as over smoothing, we face the additional complexity of a

[1] The most common type of MRI involves sampling the object of interest in the frequency-domain or *k-space* as it is referred to in the MRI literature.

[2] We stress that the notions of “low resolution” and “high resolution” are meant in a relative sense. Although resolutions commonly used in existing applications may be considered acceptable and quite useful, additional information, that may be clinically relevant, is contained at higher resolutions.

spatially varying noise behavior that depends on the local intensity in the MR image.

Rician noise degrades images in both qualitative and quantitative senses and hinders image analysis, interpretation, and feature detection. For example, it is shown in [11] that the bias due to Rician noise reduces detectability in low SNR MRI. Consequently, it is highly desirable to develop filtering methods that remove this noise. In this paper we develop a novel wavelet-domain filtering procedure for Rician data that removes noise from the MRI data.

Other wavelet-domain filters for MRI noise removal have been proposed in the literature. Filters based on wavelet coefficient soft-thresholding [12] and related thresholding [13] have been applied to MRI. Neither of these methods exploit the Rician nature of the noise, nor do they account for the inherent bias in the MR magnitude image. The wavelet-domain filter described in this paper is designed to specifically handle Rician noise. The wavelet-domain filtering procedure preserves key image details and features, reduces random noise fluctuations in the image, and removes the bias introduced by the Rician noise.

The paper is organized as follows. In Section II we review the noise statistics encountered in MRI. In Section III we review the discrete wavelet transform and wavelet-domain filtering. In Section IV we derive two wavelet-domain noise removal procedures for MRI. We show that careful consideration of the Rician nature of the noise is especially necessary in low SNR regimes. In Section V we provide several simulated and real-data experiments. In Section VI we draw some conclusions and indicate directions of ongoing and future work.

II. Noise in MRI

The most common method for acquiring MR images is to sample the object of interest in the frequency-domain [7]. The raw frequency-domain[3] measurements are denoted by

$$Y(\mu,\nu) = Y_{Re}(\mu,\nu) + i\,Y_{Im}(\mu,\nu),$$

where μ and ν denote horizontal and vertical frequencies, $i=\sqrt{-1}$, and Y_{Re} and Y_{Im} denote the real and imaginary components of the data. The measurements can be decomposed into a signal and noise component

$$Y(\mu,\nu) = S(\mu,\nu) + N(\mu,\nu),$$

where $S(\mu,\nu)$ is the signal of interest and $N(\mu,\nu)$ is a complex Gaussian white noise. The noise is primarily due to thermal noises in the patient [3, 4, 8, 14].

[3]The measurement space is also commonly referred to as *k-space* in the literature [7].

The most common reconstruction technique in MRI is to compute the inverse discrete Fourier transform (DFT) of the raw data Y [7, 11]. Let y denote the inverse DFT of Y. Due to phase errors which are difficult to control, the signal component of the measurements occurs in both real and imaginary channels [7, 11]. The value of the m,n-th pixel is complex and is given by

$$y[m,n] = (s[m,n]\cos(\theta[m,n]) + n_{Re}[m,n]) + i\,(s[m,n]\sin(\theta[m,n]) + n_{Im}[m,n]) \tag{2.1}$$

where s denotes the signal of interest, n_{Re} and n_{Im} denote Gaussian white noises with variance σ^2 in the real and imaginary channels[4], and $\theta[m,n]$ represents the phase error in the m,n-th pixel. Because of the phase error, the magnitude of y is typically used to form the image reconstruction. Define

$$x[m,n] \triangleq |y[m,n]| = \left[(s[m,n]\cos(\theta) + n_{Re}[m,n])^2 + (s[m,n]\sin(\theta) + n_{Im}[m,n])^2\right]^{\frac{1}{2}} \tag{2.2}$$

The magnitude image $x[m,n]$ is real-valued and can be displayed for visual inspection or used for automated computer analysis. Alternatively, phase-correction methods can be applied thereby enabling direct extraction of the real signal from the complex data [11, 15], but such approaches present their own difficulties and are not often used in practice. In this paper, we focus on the more common magnitude image reconstruction (2.2).

Since the magnitude reconstruction is simply the square-root of the sum of two independent Gaussian random variables, the noise in the magnitude image x is described by a Rician distribution [16]. We can regard x as an observation of the signal s with an additive noise η

$$x[m,n] = s[m,n] + \eta[m,n] \tag{2.3}$$

The additive noise is not a simple white noise in this model. Rather, $\eta[m,n] = x[m,n] - s[m,m]$ and the structure of this noise depends on $s[m,n]$. Hence, we say that the Rician noise is a signal-dependent noise. In low intensity (dark) regions of the magnitude image, the noise distribution tends to the Rayleigh distribution. In high intensity (bright) regions, the noise tends to a Gaussian distribution.

The objective of this paper is to develop a filtering method to estimate the noise-free signal s from the magnitude image x. We employ wavelet-domain filtering methods to tackle this especially difficult problem. Wavelet-domain filtering can automatically adapt to spatial variations in the

[4]The noise in the reconstructed data x is, again, Gaussian white noise since the DFT is an orthogonal transformation.

signal behavior to remove noise and preserve important signal details like edges. Moreover, we show that wavelet-domain filtering can also adapt to spatial variations in the noise distribution. Adaptation to the noise distribution is especially important in low SNR applications.

III. Wavelet-Domain Filtering

We choose to work in the wavelet-domain because the orthogonal wavelet transform tends to concentrate the energy of the desired signal s into a small number of coefficients. Hence, the wavelet transform of the noisy image consists of a small number of coefficients with high SNR (which we keep) and a large number of coefficients with low SNR (which we discard). After discarding the noisy coefficients, we reconstruct the image using the inverse wavelet transform. In this way we can filter the signal from the noise. Of course, a similar procedure could be carried out using any orthogonal signal representation, including the Fourier transform. However, the wavelet transform enables the filtering procedure to adapt to the spatial variations in the signal frequency content and thereby balances the trade-off between noise removal and excessive smoothing. Fourier-domain filtering is a global operation that cannot adjust to such spatial variations and hence leads to excessive smoothing in regions where the image has high frequency content (edges, for example). Furthermore, it has been shown that noise removal, compression, and signal recovery methods based on wavelet coefficient shrinkage or wavelet series truncation enjoy asymptotic minimax-optimal performance characteristics and, moreover, do not introduce excessive artifacts in the signal reconstruction [17].

A. The Discrete Wavelet Transform

The 1-d discrete wavelet transform (DWT) represents a real-valued continuous-time signal $f(t)$ in terms of shifts and dilations of a lowpass *scaling function* $\phi(t)$ and bandpass *wavelet* $\psi(t)$. For special choices of these functions, the shifts and dilations form an orthonormal basis for $L_2(\mathbb{R})$ and we have the signal representation [18, 19]

$$f(t) \;=\; \sum_k c_k^J \, 2^{-J/2} \, \phi(2^{-J}t - k) \;+\; \sum_{j=-\infty}^{J} \sum_k d_k^j \, 2^{-j/2} \, \psi(2^{-j}t - k) \tag{3.1}$$

with *scaling coefficients*

$$c_k^J \;=\; 2^{-J/2} \int f(t) \, \phi(2^{-J}t - k) \, dt \tag{3.2}$$

and *wavelet coefficients*

$$d_k^j \;=\; 2^{-j/2} \int f(t) \, \psi(2^{-j}t - k) \, dt. \tag{3.3}$$

A key feature of the wavelet representation is multiresolution: the first term in (3.1) is an approximation to $f(t)$ at scale (resolution) J, while the second term consists of refinements at finer and finer scales $j \leq J$. Unlike the Fourier transform which is based on sinusoidal basis functions with poor spatial-domain localization, the wavelet basis functions are localized in both time and frequency, making the DWT an ideal decomposition for signals with spatial-varying frequency content. In the DWT, many choices are possible for the wavelet and scaling function. The choice of wavelet is discussed in greater detail in Section III.

The scaling and wavelet coefficients can be computed using an extremely efficient recursion. Using the fact that the scaling function and wavelet are related by the two-scale difference equations [18, 19]

$$\phi(t/2) = \sqrt{2} \sum_n h[n]\, \phi(t-n) \tag{3.4}$$

$$\psi(t/2) = \sqrt{2} \sum_n g[n]\, \phi(t-n), \tag{3.5}$$

with h and g discrete-time lowpass and highpass filters, respectively, the scaling and wavelet coefficients can be recursively computed

$$c_k^{j+1} = \sum_n h[n]\, c_{2k-n}^j \tag{3.6}$$

$$d_k^{j+1} = \sum_n g[n]\, c_{2k-n}^j . \tag{3.7}$$

These computations can be succinctly organized into a discrete-time filter bank of lowpass and highpass filters and decimators [18, 19].

Now consider a length-2^M discrete-time signal $\mathbf{s} = \left[s[1], s[2], \ldots, s[2^M]\right]^T$. The recursion (3.6)–(3.7) beginning with $c_k^0 = s[k]$ defines a discrete-time DWT of $\mathbf{s}$.[5] Iterating (3.6)–(3.7) J times, produces an J-scale DWT consisting of J sets of wavelet coefficients at scales $j = 1, \ldots, J$, and a single set of scaling coefficients at scale J. Due to the decimation involved, there are exactly $2^{(M-j)}$ wavelet coefficients d_k^j at each scale j, and $2^{(M-J)}$ scaling coefficients c_k^J. The scaling coefficients represent a low-resolution approximation to the signal. The wavelet coefficients represent the "details" needed to reconstruct the full-resolution signal. The maximum number of iterations is $J_{\max} = M$. Because of the special filter bank structure, the forward and inverse DWT can be implemented in $O(2N \times 2^M)$ computations for $2N$-length filters h and g. Remarkably, this cost is linear in the number of signal samples processed.

[5] Note that in the discrete-time case we regard the signal itself as the finest scale ($j = 0$) scaling coefficients. No finer detail is available so the summation over the scales in (3.1) ranges only from $j = 1$ to $j = J$.

The reason that the wavelet transform is so desirable is that the wavelet transforms of real-world signals and images tend to be very sparse, with a few large scaling and wavelet coefficients dominating the representation. That is, wavelet transforms tend to compress real-world signals. The compression property of the wavelet transform is attributed to the fact that the wavelet coefficients of polynomial signals are exactly zero [20]. This fact is a consequence of the "vanishing" moments of the wavelet functions. Let ψ_k^j denote the underlying discrete-time wavelet basis function at scale j and location k. Then

$$\sum_m m^n \, \psi_k^j[m] = 0, \qquad n = 0, \ldots, N-1. \tag{3.8}$$

The wavelet functions obtained from length-$2N$ filters h and g have N vanishing moments [18]. The number of vanishing moments of ψ_k^j determines the highest degree polynomial signal for which all wavelet coefficients are equal to zero. Higher degree ($\geq N$) polynomials will have non-zero wavelet coefficients. Therefore, if the noise-free signal is described by a polynomial of degree $< N$, then all wavelet coefficients are zero and the signal is completely described by the scaling coefficients alone. Noise on the other hand is not polynomial in structure. The noise tends to be evenly distributed over all wavelet and scaling coefficients. In fact, because the wavelet and scaling functions are orthonormal, white noise is identically distributed in each coefficient. Hence, a large amount of noise can be removed from the data by discarding all the wavelet coefficients and retaining only the scaling coefficients.

In more realistic situations the signal is not polynomial, but may be well-approximated by a piecewise polynomial function. Because the wavelet functions also have localized support, most of the wavelet coefficients of a piecewise polynomial signal will be zero except those corresponding to wavelets that have support near the breakpoints of the polynomial pieces that comprise the signal. These non-zero wavelet coefficients are crucial for reconstructing important signal details like edges. Therefore, the key to effective noise removal in the wavelet-domain is to determine which wavelet coefficients do not have significant signal energy and hence can be discarded without detrimental signal loss. Methods for making this determination are called *wavelet-domain filters.*

Before discussing wavelet-domain filtering, we note that images can be decomposed using separable 2-d wavelet bases, which are implemented using a 2-d filter bank where we alternate the application of h and g on rows and columns of the image [18, 19]. Given a $2^M \times 2^M$ image $s[k,l]$ we iterate (3.6)–(3.7) J times on rows and columns to obtain J sets of wavelet coefficients at scales $j = 1, \ldots, J$, and a single set of scaling coefficients at scale J. In this case the scaling coefficients correspond to a $2^{M-J} \times 2^{M-J}$ low-resolution approximation to the original image, and the wavelet coefficients represent detail images that can be combined with the scaling coefficients to reconstruct the full-resolution image using the inverse discrete wavelet transform. Using a 2-d filter bank, the computational cost of the forward and inverse 2-d DWT is $O(2N \times 2^M \times 2^M)$ for an $2^M \times 2^M$ image.

The discrete 2-d wavelet functions $\psi_{k,l}^{j,o}$ and scaling functions $\phi_{j,k,l}$ have several indices — j corresponding to the scale, o corresponding to wavelet orientation (horizontal, vertical, or diagonal), k, l corresponding to the position. In order to keep notation to a minimum, we will employ the abstract index I for these indices. Furthermore, we vectorize the 2-d wavelet functions, scaling functions, and the image by stacking the columns of each, and employ a single abstract spatial index m. Hence, the I-th scaling coefficient of the image s is computed

$$c_I = \sum_m \phi_I[m]s[m]. \tag{3.9}$$

Similarly the I-th wavelet coefficient is computed

$$d_I = \sum_m \psi_I[m]s[m]. \tag{3.10}$$

We collectively refer to the scaling and wavelet coefficients by vectors $\mathbf{c}$ and $\mathbf{d}$, respectively.

B. Wavelet-Domain Filtering

First we compute the 2-d DWT of an image $\mathbf{x}$. Let $\mathbf{c}$ and $\mathbf{d}$ denote the noisy scaling and wavelet coefficients of $\mathbf{x}$. The goal of wavelet-domain filtering is to obtain a better estimate of noise-free signal wavelet coefficients by filtering $\mathbf{d}$. We filter (or attenuate) the contribution of a particular wavelet ψ_I to the signal reconstruction by weighting the corresponding coefficient d_I by a number $0 \leq \alpha_I \leq 1$. That is, we modify $\mathbf{d}$ component-wise according to

$$\widehat{d}_I \;=\; \alpha_I d_I. \tag{3.11}$$

Setting $\alpha_I = 0$ completely removes the contribution of the wavelet function ψ_I; setting $\alpha_I = 1$ leaves it unaltered. Choosing $0 < \alpha_I < 1$ attenuates the contribution of the I-th wavelet function accordingly. The idea is to attenuate those wavelet coefficients which contain more noise than signal. Taking the inverse DWT of the filtered wavelet coefficients $\widehat{\mathbf{d}}$ and the scaling coefficients $\mathbf{c}$, we obtain an improved image estimate denoted $\widehat{\mathbf{s}}$. The new image $\widehat{\mathbf{s}}$ is called a *wavelet-domain filtered* version of $\mathbf{x}$. The collection of weights $\boldsymbol{\alpha} = \{\alpha_I\}$ is the *wavelet-domain filter.*

Clearly, the crucial issue in wavelet-domain filtering is the design of the filter $\boldsymbol{\alpha}$. Ideally we would like the wavelet-domain filtering procedure to be adapted to the local SNR in each wavelet coefficient, so that we can suppress wavelet coefficients with very low SNR. Let us assume that the noisy wavelet coefficient is an unbiased estimator of the value of the signal's wavelet coefficient in the ideal noise-free case, and denote its mean by $\delta_I = \mathrm{E}[d_I]$. That is, δ_I is the wavelet coefficient of the noise-free signal. Then the filter weight that minimizes the mean squared error of each coefficient is

$$\alpha_I^{\mathrm{MSE}} = \frac{\delta_I^2}{\delta_I^2 + \sigma_I^2} \tag{3.12}$$

where σ_I^2 is the variance of the noisy coefficient d_I. This optimal weight is easily deduced by minimizing $\mathrm{E}\left[(\delta_I - \alpha_I d_I)^2\right]$ with respect to α_I. We can interpret the filtering operation as a wavelet-domain analog of the classical Wiener filter [21]. Unfortunately, this optimal filter weight requires perfect knowledge of δ_I and σ_I^2 and therefore is not feasible in practice.

One filter that approximates (3.12) is given by [22, 23]

$$\alpha_I = \left(\frac{d_I^2 - \widehat{\sigma_I^2}}{d_I^2} \right)_+ \tag{3.13}$$

where d_I^2 is the square of the I-th wavelet coefficient, $\widehat{\sigma_I^2}$ is an estimate of the variance of the d_I, and $(\cdot)_+$ denotes the positive part of the argument (sets negative values to zero). Again, assuming that $\mathrm{E}[d_I] = \delta_I$, note that $\mathrm{E}[d_I^2] = \delta_I^2 + \sigma_I^2$. Hence, the $d_I^2 - \widehat{\sigma_I^2}$ is simply an estimate of the numerator in (3.12), and d_I^2 is an estimate of the denominator in (3.12). Therefore, (3.13) provides a reasonable approximation to (3.12), and we can interpret (3.13) as a data-adaptive, wavelet-domain, Wiener filter. The thresholding action of $(\cdot)_+$ guards against illogical negative weighting factors (note that the argument may be negative if the magnitude of the wavelet coefficient is small). This threshold was first introduced in [22], and it arises naturally in the minimization of a predictive sum of squared errors criterion.

The action of the filter (3.13) is to set small wavelet coefficients, with squared magnitude less than the estimated variance, to zero and to leave larger coefficients approximately unaltered. Fig. 1 depicts this nonlinearity by plotting $\widehat{d_I} = \alpha_I\, d_I$. Other wavelet-domain filtering operations, such as hard and soft-thresholds [12, 24], can also be employed. However, the operation given by (3.13) has been shown to provide better results than other methods in this and other imaging applications [23]. Therefore, we have chosen to work with this threshold here.

More aggressive noise removal is possible by weighting the estimated variance in (3.13). For example,

$$\alpha_I = \left(\frac{d_I^2 - \tau\, \widehat{\sigma_I^2}}{d_I^2} \right)_+ \tag{3.14}$$

where $\tau \geq 1$ is a user-defined parameter that adjusts the amount of noise reduction. Selecting $\tau > 1$, effectively raises the threshold level of the nonlinearity. For example, Donoho *et al* suggest a universal threshold for Gaussian noise removal that is greater than 1 [24]. Larger threshold levels can result in greater noise reduction, but may also lead to over smoothing. The choice of $\tau = 1$ provides a conservative lower bound on the threshold level. In MRI applications, we have found that $\tau = 2$ provides very good results in terms of squared error and visual quality (see Section V).

Another degree of flexibility in wavelet-domain filtering is the choice of the underlying wavelet. In our experiments we have found that the Haar wavelet [20] does a better job at preserving fine image details such as small vascular structure compared to other wavelets. The reason for this may be the fact that the Haar wavelet also has the most compact spatial support of all wavelets. That is, at each scale of analysis the Haar wavelet is most localized in space. On the other hand, wavelets with larger spatial support such as the Daubechies-6 wavelet [20] have better frequency localization and approximation properties. Filtering with wavelets with large spatial support generally produces smoother looking images. However, we have observed that such wavelets may also lead to excessive smoothing of fine details. This illustrates a fundamental trade-off in wavelet-based filtering. We desire a wavelet with a large number of vanishing moments for good polynomial approximation, but this requires a wavelet with larger spatial support. To summarize we advocate using the Haar wavelet so that fine details are best preserved.

One final practical issue is the shift-variant nature of the wavelet-domain filter described above. Because the wavelet coefficient values depend on the alignment between the data and the wavelet basis functions, simple shifts of the data lead to complicated changes at the output of the filter. Moreover, misalignment between the wavelets and signal structure can lead to artifacts in the filtered image. To improve wavelet-domain filtering methods, several authors have proposed shift-invariant (undecimated) DWT methods [25, 26]. Shift-invariant wavelet transforms have been shown to provide better results in many cases since they are less sensitive to misalignments between the edges in images and the wavelet basis functions. Shift-invariant filtering can be achieved by applying the filter described above to every circular shift of the image, unshifting each result, and then averaging all the results together. If the image is $2^M \times 2^M$ pixels, then there are exactly $2^M \times 2^M$ circular shifts in total (2^M in both horizontal and vertical directions), regardless of the underlying wavelet used in the analysis. In fact, fast algorithms that exploit the redundancies in different shifts make shift-invariant filtering only modestly more computationally intensive than the shift-variant filter above [26, 25]. We make use of these ideas in the wavelet-domain filters for MRI described next.

IV. Wavelet-Domain Filters for MRI

In this section, two wavelet-domain filters for noise removal in MRI are described. Before describing the new wavelet-domain filter designed to handle Rician noise, we briefly review a filter based on a Gaussian approximation to the Rician noise in the magnitude MR image. Such an approximation is quite reasonable at high SNR. However at low SNR, in addition to being noisy, the wavelet and scaling coefficients of the magnitude image are also biased from the corresponding coefficients of the ideal noise-free signal s. The bias is signal-dependent and, hence, is quite difficult to remove. To overcome this problem, we introduce a second filter that operates on the *squared magnitude* MR image. The noise in the squared magnitude image is non-central chi-square distributed. Using the properties of the wavelet transform and the non-central chi-square distribution, we show that the wavelet coefficients of the squared magnitude image are unbiased estimators of the wavelet coefficients of the squared signal s^2. Furthermore, although the scaling coefficients of the squared magnitude image are biased estimators of the scaling coefficients of s^2, the bias is a simple constant, independent of the signal. Based on these observations we derive a wavelet-domain filter that removes noise and adjusts for the bias in the scaling coefficients. This filter may be applied in both high and low SNR regimes.

A. Algorithm I: Gaussian Noise Approximation

At high SNR, the Rician noise is well-approximated as Gaussian. For example, if the SNR in a pixel is greater than 10dB, then the mean pixel value is more than three standard deviations from the origin and the Rician distribution is approximately Gaussian. If the SNR is below 10dB, then the Rician noise begins to deviate from Gaussian as illustrated in Fig. 2. Hence, in high SNR regimes we may treat the MRI noise as Gaussian white noise with standard deviation σ. In effect, the noise is *signal-independent* in high SNR MRI, and consequently separation of signal and noise is fairly straightforward in the wavelet-domain.

To derive a wavelet-domain filter for high SNR MRI, let us assume that the noise in the MR magnitude image is white Gaussian with standard deviation σ. Based on this assumption and the fact that DWT is an orthogonal transformation, the variance in each wavelet coefficient is σ^2 in the spatial domain. This observation and equation (3.14) suggest a simple wavelet-domain filter for this high SNR situation.

$$\alpha_I = \left(\frac{d_I^2 - \tau\,\widehat{\sigma_I^2}}{d_I^2} \right)_+ \tag{4.1}$$

where $\widehat{\sigma_I^2} = \sigma^2$ for every wavelet coefficient and τ is chosen by the user. In practice, σ^2 may not be known precisely. In Section IV-C we suggest a simple procedure for estimating σ^2 directly from the noisy MR image. This filtering procedure, called Algoithm I, is outlined here.

Algorithm I

Wavelet-Domain Filtering the Magnitude Image

1. Compute the J-scale, $1 \leq J \leq M$, DWT of the $2^M \times 2^M$ MR magnitude image.

2. Apply wavelet-domain filter (4.1) to the wavelet coefficients to obtain the filtered coefficients $\widehat{\mathbf{d}}$.

3. Compute the inverse DWT using $\widehat{\mathbf{d}}$ to obtain the improved estimate of the image $\widehat{\mathbf{s}}$.

Algorithm I is similar to previously proposed filtering methods [12, 13] in that it treats the noise as though it were independent of the signal of interest and does not account for the bias in the image. This algorithm is quite effective in high SNR regimes.

However, the Gaussian approximation to the data has a major drawback — especially in low SNR applications. The mean of the magnitude image is not equal to the noise-free image s. That is, the magnitude image is biased. Consider a pixel $x[m] = s[m] + \eta[m]$. The pixel value is said to be unbiased if its expected value is equal to the signal component $s[m]$. This is not the case for Rician noise. The expected value of $x[m]$ is

$$\mathrm{E}\left(x[m]\right) = \sigma(\pi/2)^{1/2} e^{-s[m]^2/4\sigma^2} \left[(1 + s[m]^2/2\sigma^2) J_0(s[m]^2/4\sigma^2) + (A^2/2\sigma^2) J_1(s[m]^2/4\sigma^2) \right] \tag{4.2}$$

where $J_n(\cdot)$ denotes the modified Bessel function of the first kind of order n [16]. In general this expected value is not equal to the signal value $s[m]$. However, as $SNR = \frac{s[m]^2}{\sigma^2}$ tends to infinity $\mathrm{E}\left(x[m]\right)$ tends to $s[m]$. Hence, the magnitude image is asymptotically unbiased.

A problem remains at low SNRs. Fig. 3 illustrates the reduction in contrast in low SNR MRI. The SNR is computed as the squared background signal level relative to the underlying Gaussian noise power σ^2. The "contrast" is defined as

$$c(s) = \frac{\max(s) - \min(s)}{\max(s) + \min(s)}, \tag{4.3}$$

where s is the noise-free signal. The bias in the MR magnitude image can significantly reduce image contrast as shown in Fig. 4. In this case SNR $= 10\log_{10}(10^2/5^2) \approx 6\,\mathrm{dB}$, and the contrast of the noise-free signal is $c = 0.500$. Because of the reduction in contrast, it is of interest to not only reduce the random fluctuations due to the noise, but also to reduce the bias. The question is: How does the bias effect the wavelet and scaling coefficients, and can it easily be removed?

Recall that one of the key properties of wavelet-domain filtering is that the wavelet coefficients of polynomial signals are identically zero. If the noise-free s is polynomial, then the wavelet coefficients of s are all exactly zero. However, in general, the expected value of x given by (4.2) can not be described by a polynomial and hence the expected value of the wavelet coefficients of the noisy signal are not zero. This means that the wavelet coefficients of the noisy signal are biased. Therefore, the wavelet coefficients of the MR magnitude image are not only erroneous due to random fluctuations caused by the noise, but they are also biased from their true values. A similar analysis shows that the scaling coefficients of the magnitude image are also biased.

We see that the bias in the MR magnitude image manifests itself in both the wavelet and scaling coefficients. Algorithm I, based on a Gaussian noise approximation, does not account for this bias. Moreover, the bias is signal-dependent. In high intensity (high SNR) regions of the image the bias is negligible, whereas in low intensity (low SNR) regions the bias is considerable. Because of the signal-dependent nature of the bias, it is quite difficult to remove.

Recall, however, that the real and imaginary components contain simple additive Gaussian white noise. Hence, one might be tempted to apply existing wavelet-domain filtering techniques to each component separately. There are two problems with this approach. First, due to the unknown, spatially varying phase, the structure of the real and imaginary images may not be well suited for wavelet analysis at all. For example, if the phase oscillates quickly in space, then the real image may exhibit high frequency structure that is not naturally occuring in the physical object being imaged. The point is that a wavelet representation may not be as efficient for representing and analyzing the real and imaginary images as it is for magnitude image. The second problem is much more serious. Thresholding the real and imaginary images separately may introduce severe phase distortions in the magnitude reconstruction that can be very degrading. Simple experiments have been performed that show that such phase distortions can be very visible in resulting the magnitude image. To avoid phase distortions we must couple the thresholding performed on the real and imaginary images. The most natural way to do this is to base the threshold decisions on the combined magnitude of corresponding real and imaginary wavelet coefficients. This leads us right back to the Rician problem under consideration in this paper.

Therefore, to combat this difficult problem we propose a new wavelet-domain filter that operates on the squared magnitude, rather than magnitude, image. The squared magnitude data obeys a non-central chi-square distribution and can be more easily removed in the wavelet-domain.

B. Algorithm II: Non-Central Chi-Square Noise

Working with the squared magnitude image, instead of the magnitude image, offers two advantages which are stated below and proved later in this section.

1. The wavelet coefficients of the squared magnitude image are unbiased estimators of the wavelet coefficients of the squared signal s^2. are unbiased.
2. The scaling coefficients of the squared magnitude image are almost unbiased estimators of the scaling coefficients of s^2. The expected value of each scaling coefficient differs from the noise-free scaling coefficient by a constant offset that does not depend on the signal. Hence, the bias can be easily removed.

The square of a Rician random variable is described by a scaled non-central chi-square distribution [27]. It follows from (2.2) that the squared magnitude is

$$x^2[m] = (s[m]\cos(\theta) + n_{Re}[m])^2 + (s[m]\sin(\theta) + n_{Im}[m])^2 \qquad (4.4)$$

Equivalently, we can express the squared magnitude as

$$\begin{aligned} x^2[m] &= (s[m]\cos(\theta) + \sigma n_1[m])^2 + (s[m]\sin(\theta) + \sigma n_2[m])^2 \\ &= \sigma^2 \left[(s[m]\cos(\theta)/\sigma + n_1[m])^2 + (s[m]\sin(\theta)/\sigma + n_2[m])^2\right] \end{aligned} \qquad (4.5)$$

where $n_1 \sim N(0,1)$ and $n_2 \sim N(0,1)$ are independent standard Gaussian noises. The bracketed quantity is a non-central chi-square random variable with two degrees of freedom and non-centrality parameter $s^2[m]/\sigma^2$. Therefore, the squared magnitude $x^2[m]$ is simply a non-central chi-square random variable scaled by σ^2.

Using the moments of the non-central chi-square distribution [27], the mean and variance of the $x^2[m]$ are:

$$\mathrm{E}\left[x^2[m]\right] = s^2[m] + 2\,\sigma^2 \qquad (4.6)$$

$$\mathrm{Var}\left[x^2[m]\right] = 4\,\sigma^2(s^2[m] + \sigma^2) \tag{4.7}$$

Note that the mean is biased from the desired value $s^2[m]$.

Now consider a wavelet coefficient

$$\bar{d}_I = \sum_m \psi_I[m] x^2[m]. \tag{4.8}$$

Note that $\bar{d}_I$ is used to distinguish this wavelet coefficient from the wavelet coefficient d_I corresponding to the magnitude image $x[m]$. Using (4.6) we find the expected value of $\bar{d}_I$:

$$\mathrm{E}[\bar{d}_I] = \sum_m \psi_I[m](s^2[m] + 2\,\sigma^2) = \sum_m \psi_I[m] s^2[m] \tag{4.9}$$

where we have exploited the fact that the wavelet function has at least one vanishing moment, thereby annihilating the constant bias term $2\sigma^2$. That is, the wavelet basis functions are insensitive to the constant bias term in the squared magnitude data and, therefore, the wavelet coefficients of the squared magnitude image are unbiased estimators of the wavelet coefficients of the noise-free squared signal s^2. This shows that the wavelet coefficients of the squared magnitude image are unbiased estimators of the wavelet coefficients of the squared signal s^2.

Now we derive the correct wavelet-domain filter of the form (3.14). To construct the filter we need an estimator for the variance of the I-th wavelet coefficient of squared magnitude image.

$$\begin{aligned}
\mathrm{Var}\left[\bar{d}_I\right] &= \mathrm{E}\left[\left(\bar{d}_I - \mathrm{E}[\bar{d}_I]\right)^2\right] \\
&= \mathrm{E}\left[\bar{d}_I^2\right] - \left(\mathrm{E}[\bar{d}_I]\right)^2 \\
&= \sum_{k,l} \psi_I[k]\psi_I[l]\mathrm{E}\left[x^2[k]x^2[l]\right] - \left(\sum_m \psi_I[m]s^2[m]\right)^2 \\
&= 4\,\sigma^2 \sum_m \psi_I^2[m](s^2[m] + \sigma^2)
\end{aligned} \tag{4.10}$$

where we have used (4.6) and (4.7) in the last step.

Assuming σ^2 is known (a robust, data-based estimator for σ^2 are given next in Section IV-C), an unbiased estimator for σ_I^2 in terms of the noisy squared magnitude image itself is given by

$$\begin{aligned}
\widetilde{\sigma_I^2} &= 4\,\sigma^2 \sum_m \psi_I^2[m](x^2[m] - \sigma^2) \\
&= 4\,\sigma^2 \left[\left(\sum_m \psi_I^2[m]x^2[m]\right) - \sigma^2\right]
\end{aligned} \tag{4.11}$$

where we have used the fact that $\sum_m \psi_I^2[m] = 1$ since the wavelet functions are orthonormal.

Note that the minimum value of the true variance σ_I^2 is $4\sigma^4$. To insure that the estimated value does not fall below this minimum we replace (4.11) with the more robust

$$\widehat{\sigma_I^2} = 4\sigma^2 \max\left[\left(\sum_m \psi_I^2[m]x^2[m]\right) - \sigma^2 \,,\, \sigma^2\right] \tag{4.12}$$

To form this estimator we must compute the inner products of the image with the squared wavelet basis functions. We call this computation the discrete wavelet squared transform (DWST). The DWST also arises in other applications, and a fast filter bank algorithm exists for its computation [23]. The complexity of the DWST for a $2^M \times 2^M$ image with length-$2N$ filters is $O\left((2N)^3 \times 2^M \times 2^M\right)$, just slightly greater than that of the fast DWT.

Next consider the scaling coefficients. The I-th scaling coefficient is computed

$$\bar{c}_I = \sum_m \phi_I[m]x^2[m] \tag{4.13}$$

The $\bar{c}_I$ notation is used to avoid confusion with the scaling coefficients $c_I$corresponding to the magnitude image. Using (4.6) we find the expected value of $\bar{c}_I$:

$$\mathrm{E}[\bar{c}_I] = \sum_m \phi_I[m](s^2[m] + 2\,\sigma^2) = \sum_m \phi_I[m]s^2[m] + 2\,\sigma^2\left(\sum_m \phi_I[m]\right) \tag{4.14}$$

Hence, each scaling coefficient is biased from the noise-free scaling coefficient $\sum_m \phi_I[m]s^2[m]$ by a simple shift in the mean equal to the constant

$$C = 2\,\sigma^2\left(\sum_m \phi_I[m]\right) \tag{4.15}$$

For a J-scale 2-d DWT

$$\sum_m \phi_I[m] = 2^J \tag{4.16}$$

Hence, $C = 2^{(J+1)}\sigma^2$ does not depend on the signal or the position of the scaling function. The bias C only depends on the noise level $2\sigma^2$ and the number of scales in the DWT, and can easily be subtracted from the scaling coefficients. Note that this procedure for removing the bias is similar to the spatial-domain bias correction described in [5, 6]. The overall algorithm for estimating the underlying signal s using the squared magnitude image x^2 is given by the following algorithm.

Algorithm II

Wavelet-domain Filtering the Squared Magnitude Image

1. Compute the J-scale DWT and DWST of the squared magnitude image.
2. Form estimates of the variances σ_I^2 of each wavelet coefficient $\bar{d}_I$ according to (4.12).
3. Filter the wavelet coefficients according to (3.14).
4. Remove the bias from the scaling coefficients $\bar{c}_I$ by subtracting $C = 2^{(J+1)}\,\sigma^2$ from each.
5. Compute the inverse DWT of the filtered wavelet and scaling coefficients to obtain an estimate of s^2.
6. Take the pixel-by-pixel square-root of the result to obtain an estimate of s.

The main advantage of the Algorithm II, compared to Algorithm I, is that it accounts for the bias in the data. This can provide improvements in image quality and contrast as shown in Section V.

C. Estimating σ^2

Both Algorithm I and Algorithm II require the underlying noise power σ^2. In a real-world application, we typically do not know σ^2 *a priori*, and it must be estimated from the data itself. A simple estimator is based on the following argument. Typical MR images include an empty region of air outside the patient. If we consider the squared magnitude MR image, then in such empty regions $s^2[m] = 0$ and the average pixel value is $2\,\sigma^2$, according to (4.6). Hence, one-half of the mean of the squared pixel values in the region outside that patient provides us with a very reliable estimator of σ^2.

D. Shift-Invariant Filtering

As noted in Section III, wavelet-domain filtering based on the standard DWT is shift-variant. Shift-invariant methods can provide better performance [25, 26]. Although fast algorithms exist for shift-invariant DWTs, these methods are not directly applicable to the DWST used in Algorithm II. Therefore, in this paper we take a different approach. Recall that shift-invariant filtering can be achieved by applying the filter to all possible shifts of the image, unshifting each result, and then averaging all the results together. A shift-invariant version of Algorithm II is possible using this simple idea. However, this approach is very computationally intensive for large images (we have to repeat Algorithm II N times, where N is the number of pixels in the image). Therefore we propose an *approximately* shift-invariant scheme. We apply Algorithm II over just a small range of shifts, unshifting each result, and then averaging them together. For example, we can shift the image left and right, up and down, up to K pixel locations in each direction for a total of $(2K+1)^2$ shifted images. Again, all shifts are carried out in a circular fashion. While this does not guarantee shift-invariance, it does reduce the dependence of the filter output on the alignment between the data and wavelets. In our experiments in Section V, we used $K = 2$. Little improvement in perceptual quality or squared error was observed with larger values of K, including the full shift-invariant DWT.

V. Examples

In this section we compare the two wavelet-domain filtering algorithms' performances on real MRI and synthetic data. Attention is focused at low SNRs where the two filters produce significantly different results. At high SNRs, where the Rician noise is well-approximated as Gaussian, experiments have shown that the two algorithms perform equally well. Therefore, we have not included high SNR examples here.

In our experiments we have used the Haar wavelet, although the algorithms we derived may be used in conjunction with any wavelet basis. We have found that the Haar wavelet does a better job at preserving fine image details such as small vascular structure compared to other wavelets (recall the discussion in Section III-B). In all experiments, the threshold parameter $\tau = 2$ (see Section III-B), and we use the approximately shift-invariant filtering scheme with $K = 2$ (see Section IV-D) based on the $J = 4$ scale Haar DWT. Furthermore, in all image comparisons, the window and level of the image intensity display is constant and is chosen to provide the best visual quality for the original noisy image in each case.

A. Example 1: Contrast and Noise

In this example we assess the performance of both wavelet-domain filtering algorithms in terms of contrast and noise reduction. The simulated image shown in Fig. 4(a) is used in this example. Complex Gaussian white noise is added to this image. The magnitude of the noisy image is then Rician distributed and is shown in Fig. 4(b). Algorithms I and II are applied to the magnitude of the noisy image. In this example, the noise power σ^2 was assumed to be known in both algorithms. The SNR is computed as the squared background signal level relative to the underlying Gaussian noise power σ^2, and in this case SNR ≈ 6 dB. The filtered images resulting from Algorithm I and

II are shown Fig. 4 (c) and (d), respectively. The visual quality of the two filtered images is much better than the noisy image in (b). The contrast in the noise images is computed by

$$c = \frac{m_1 - m_2}{m_1 + m_2}, \tag{5.1}$$

where m_1 and m_2 are the mean pixel values in the disc and the background, respectively. In this case, the contrast is roughly 10% better in the result of Algorithm II compared to Algorithm I. Similarly, the sum of squared pixel errors is more than 20% lower for Algorithm II compared to Algorithm I.

To obtain a more complete understanding of the performances of the two algorithms, the experiment above was repeated over a range of SNRs. At each SNR, the 10 independent trials were conducted. Fig. 5 (a) depicts the mean square error (MSE) as a function of the SNR of the noisy image. Here, the MSE is computed by taking the average of the sum of squared pixel errors over the 10 trials at each SNR. This shows that both Algorithms I and II significantly reduce the noise over a wide range of SNRs. The MSE performance of Algorithm II is slightly better than that of Algorithm I, especially at low SNR. Fig. 5 (b) shows that the contrast is much higher in the images processed by Algorithm II at low SNR.

B. Example 2: MRI Data

In this example we examine the performance of the filtering algorithms with real MRI data. To simulate a low SNR MR image with a known "truth" we have taken a high SNR image, added complex Gaussian white noise ($\sigma^2 = 256$) to it, and computed its magnitude. The resulting image is a low SNR MRI image with a Rician distribution. The noise level was chosen so that the resulting image was similar to low SNR levels encountered in practice. In this case, the SNR in the low intensity (gray matter) region of the brain is approximately 6dB. The high and low SNR images are shown in Fig. 6 (a) and (b), respectively.

The low SNR image was filtered using both algorithms, and the results are shown in Fig. 6 (c) and (d). The noise power σ^2, required by both algorithms, was very accurately estimated directly from average squared pixel value in the lower-left-hand corner of the squared magnitude image (the region outside the patient). Recall, the motivation for this estimate given in Section IV-C. Note that both algorithms reduce the noise significantly, and the contrast of the image produced by Algorithm II is higher than that of Algorithm I. This is especially noticeable when the differences between the high SNR image and the filtered images are compared in Fig. 7.

The experiment was repeated 100 times, with independent noises in each case. The normalized average sum of squared pixels errors were 1.00 for the noisy magnitude image, 0.45 for the filtered image produced by Algorithm I, and 0.25 for the filtered image produced by Algorithm II. The true noise power was $\sigma^2 = 256.00$. The mean value of the esimated power over the 100 trials was $\widehat{\sigma}^2 = 263.94$ and the standard deviation was 12.35. Hence, the estimator proposed in Section IV-C is very reliable.

C. Example 3: Low SNR MRI Data

In this example we examine the performance of the filtering algorithms with actual low SNR MRI data. The data were acquired with a 1.5 Tesla Horizon System (General Electric Medical Systems, Waukesha, WI) equipped with EchoSpeed gradients. Fig. 8 shows the original and filtered images. The noise power σ^2 was estimated from the mean pixel value in the regions of the squared magnitude image outside the patients head. The SNR in the low intensity regions of the brain is approximately 9dB. Both algorithms again significantly reduce random fluctuations due to noise without loss of image detail. Furthermore, the contrast of the image produced by Algorithm II appears superior to that of Algorithm I, as expected.

The region of interest (ROI) indicated by the highlighted rectangle in Fig. 8 (b) is used to assess the contrast of the two filtered images. The contrast in the noise images is computed by

$$c = \frac{m_1 - m_2}{m_1 + m_2}, \tag{5.2}$$

where m_1 is the mean pixel value in the bright region running down the middle of the ROI, and m_2 is the mean pixel value in the background of the ROI. In this case, the contrast in the ROI is roughly 9% better in the result of Algorithm II compared to Algorithm I.

VI. Discussion and Conclusions

We have presented two wavelet-domain filtering methods for noise removal in MRI. In high SNR situations, the Rician noise inherent in MR magnitude images is well-approximated as Gaussian white noise. Algorithm I provides a simple and effective noise removal method for such cases. At low SNR, the Gaussian approximation is no longer valid and the Rician noise has two degrading effects; the random fluctuation of pixel values and the introduction of a signal-dependent bias. The bias can significantly reduce image contrast. To deal with this more complicated situation, we have derived a wavelet-domain filter that operates on the squared magnitude image. Algorithm II exploits the unique properties of the wavelet transform and the non-central chi square distribution to remove noise and bias from the squared magnitude image.

The computational complexity of both algorithms is minimal. Algorithm I may be implemented with the fast DWT requiring $O(2N \times 2^M \times 2^M)$ operations for a $2^M \times 2^M$ image, where $2N$ is the length of the filters in the DWT filter bank. Algorithm II may be implemented with the fast DWT and fast DWST. The fast DWST [28] can be computed in $O\left((2N)^3 \times 2^M \times 2^M\right)$ operations, slightly more than the DWT, and is the main computational burden of Algorithm II.

We have demonstrated the superior performance of Algorithm II, in comparison to that of Algorithm I, in low SNR MRI with simulations and real data. However, as a rule of thumb, we have found that the two algorithms produce roughly similar results if the SNR in the image (or region of interest) is greater than 15dB. This can be inferred from the Fig. 5(b), for example, where we see very little loss in contrast at SNRs above 15dB. Today, most clinical applications of MRI operate at relatively high SNRs. However, as discussed in the introduction, high resolution MRI may benefit from noise reduction methods such as Algorithm II. Hence, the advantage of Algorithm II is that it can be used in *both* high and low SNR imaging situations. The only advantage of Algorithm I in the high SNR case is that its complexity is slightly lower than that of Algorithm II.

Investigations are under way to assess the performance of the proposed algorithms in several MRI applications, including functional MRI, myocardial perfusion studies, and magnetic resonance angiography (MRA). MRA provides an image of the vascular structure within the brain [29]. The results of this paper have been applied to MRA to improve the visibility and contrast of small veins and arteries [30].

Acknowledgments

The author thanks Jim Siebert and Tom Cooper for their help in acquiring the data and for many stimulating conversations; and Dr. James Potchen for his cooperation in this investigation. The author also thanks the anonymous reviewers for their very helpful comments and suggestions.

References

[1] W. L. Wood, M. J. Bronskill, R. V. Mulkern, and G. E. Santyr, "Physical MR desktop data," *J. Magn. Reson. Imag.*, vol. 3, pp. 19–26, 1994.

[2] S. Kim, W. Richter, and K. Ugurbil, "Limitations of temporal resolution in functional MRI," *Magn. Reson. Med.*, vol. 37, pp. 631–636, 1997.

[3] W. A. Edelstein, P. A. Bottomley, and L. M. Pfeifer, "A signal-to-noise calibration procedure for NMR imaging systems," *Med. Phys.*, vol. 11, pp. 180–185, 1984.

[4] W. A. Edelstein, G. Glover, C. Hardy, and R. Redington, "The intrinsic signal-to-noise ratio in NMR imaging," *Magn. Reson. Med.*, vol. 3, pp. 604–618, 1986.

[5] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magn. Reson. Med.*, vol. 34, pp. 910–914, 1995.

[6] R. M. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," *Med. Phys.*, vol. 12, pp. 232–233, 1985.

[7] A. Macovski, "Noise in MRI," *Magn. Reson. Med.*, vol. 36, no. 3, pp. 494–497, 1996.

[8] E. R. McVeigh, R. M. Henkelman, and M. J. Bronskill, "Noise and filtration in magnetic resonance imaging," *Med. Phys.*, vol. 12, pp. 586–591, 1985.

[9] G. McGibney and M. R. Smith, "An unbiased signal-to-noise ratio measure for magnetic resonance images," *Med. Phys.*, vol. 20, pp. 1077–1078, 1993.

[10] A. J. Miller and P. M. Joseph, "The use of power images to perform quantitative analysis on low SNR images," *Magn. Reson. Imaging*, vol. 11, pp. 1051–1056, 1993.

[11] M. A. Berstein, D. M. Thomasson, and W. H. Perman, "Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction," *Med. Phys.*, vol. 16, no. 5, pp. 813–817, 1989.

[12] J. Weaver, Y. Xu, D. Healy, and J. Driscoll, "Filtering MR images in the wavelet transform domain," *Magn. Reson. Med.*, vol. 21, pp. 288–295, 1991.

[13] M. Hilton, T. Ogden, D. Hattery, G. Eden, and B. Jawerth, "Wavelet denoising of functional MRI data," in *Wavelets in Medicine and Biology*, vol. CRC Press, 1996.

[14] R. T. Constable and R. M. Henkelman, "Why MEM does not work in MR image reconstruction," *Magn. Reson. Med.*, vol. 14, pp. 12–25, 1990.

[15] D. Noll, D. Nishimura, and A. Macovski, "Homodyne detection in magnetic resonance imaging," *IEEE Tran. Med. Imaging*, vol. 10, no. 2, pp. 154–163, 1991.

[16] S. O. Rice, "Statistical properties of a sine-wave plus random noise," *Bell Syst. Tech. J.*, vol. 27, pp. 109–157, 1948.

[17] D. L. Donoho, "Unconditional bases are optimal bases for data compression and for statistical estimation," *App. and Comp. Harmonic Analysis*, vol. 1, pp. 100–115, Dec. 1993.

[18] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding.* Englewood Cliffs, NJ: Prentice-Hall, 1995.

[19] G. Strang and T. Nguyen, *Wavelets and Filter Banks.* Wellsley, MA: Wellsley-Cambridge, 1996.

[20] I. Daubechies, *Ten Lectures on Wavelets.* Philidelphia: SIAM CBMS-NSF Series in Applied Mathematics, no. 61, 1992.

[21] L. L. Scharf, *Statistical Signal Processing. Detection, Estimation, an Time Series Analysis.* Reading, MA: Addison-Wesley, 1991.

[22] R. D. Nowak, "Optimal signal estimation using cross-validation," *IEEE Signal Processing Letters*, vol. 4, no. 1, pp. 23–25, 1997.

[23] R. D. Nowak and R. G. Baraniuk, "Wavelet-domain filtering for photon imaging systems," *Proc. SPIE, Wavelet Applications in Signal and Image Processing V*, vol. 3169, pp. pp. 55–66, August 1997.

[24] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Statist. Assoc.*, vol. 90, pp. 1200–1224, Dec. 1995.

[25] R. Coifman and D. Donoho, "Translation invariant de-noising," in *Lecture Notes in Statistics: Wavelets and Statistics*, vol. New York: Springer-Verlag, pp. 125–150, 1995.

[26] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Letters*, vol. 3, no. 1, pp. 10–12, 1996.

[27] N. L. Johnson, S. Kotz, and A. W. Kemp, *Univariate Discrete Distributions.* New York: John Wiley and Sons, 1992.

[28] R. D. Nowak and R. G. Baraniuk, "Wavelet-based filtering for photon imaging systems," *IEEE Trans. Image Processing*, submitted April 1997.

[29] E. J. Potchen, E. M. Haacke, J. E. Siebert, and A. Gottschalk, *Magnetic Resonance Angiography: Concepts and Applications.* St. Louis: Mosby-Year Book Inc., 1993.

[30] R. Gregg and R. Nowak, "Noise removal methods for high resolution MRI," in *Proc. IEEE Medical Imaging Conf.*, 1997.

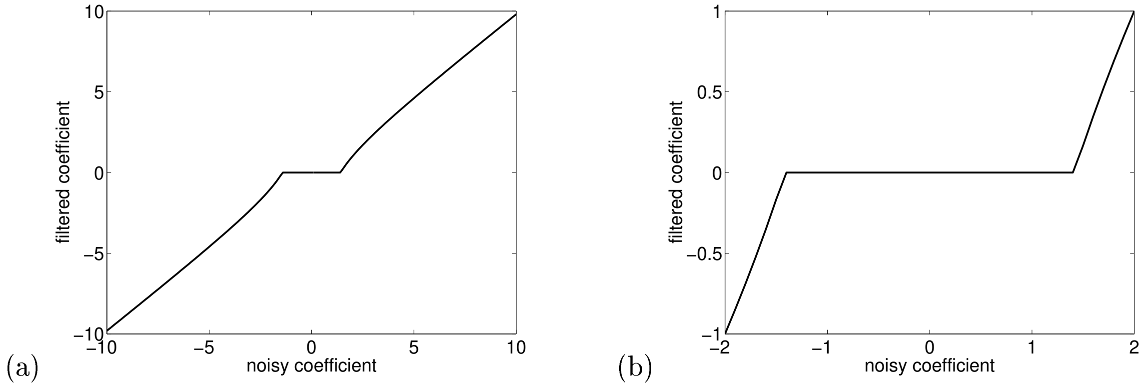


Fig. 1. *Wavelet-domain filtering nonlinearity for MRI. (a) Filtering operation depicted by plotting filtered coefficient value $\widehat{d}_I$ as a function of the original noisy coefficient value d_I. (b) Close-up of filtering operation near the threshold. The noise power depicted in this case is $\sigma_I^2 = 2$. The threshold level occurs at σ_I.*

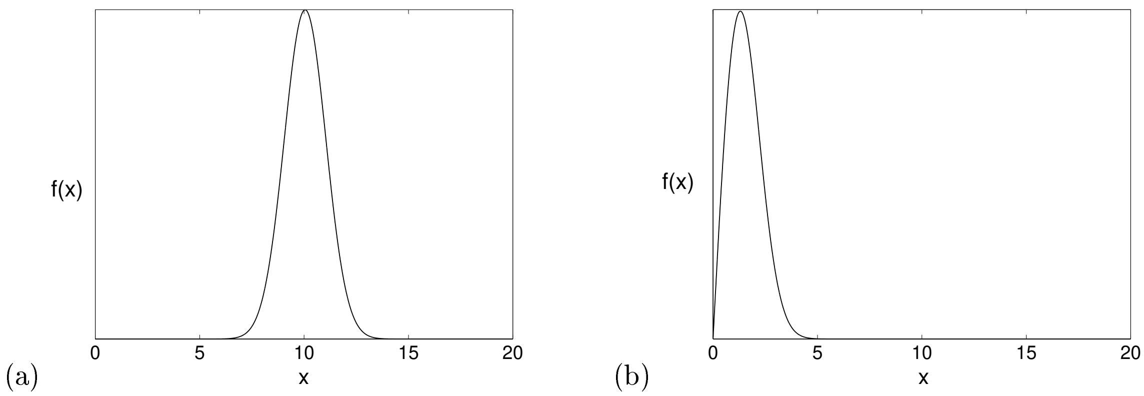
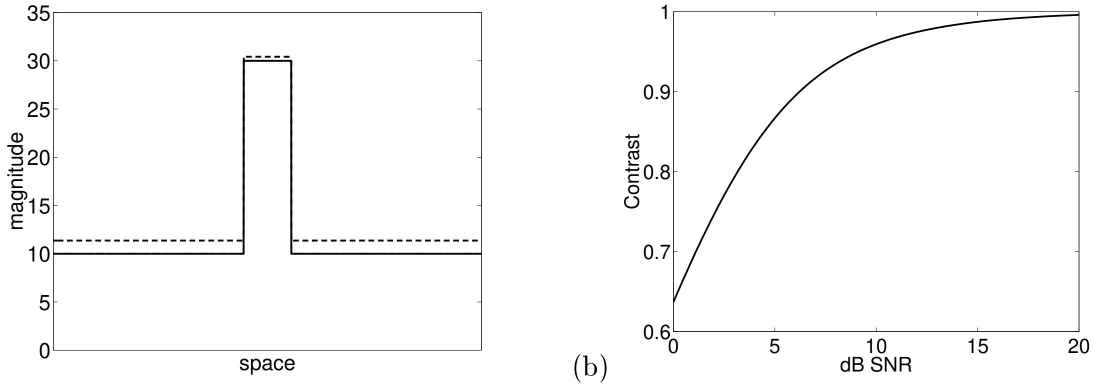


Fig. 2. *Signal dependence of Rician noise. (a) Depicted here is the Rician noise density $f(x)$ with signal $s[m] = 10$ and underlying Gaussian noise power $\sigma^2 = 1$. The SNR is 20dB. In this high SNR situation the density is nearly identical to the Gaussian density with mean $s[m]$ and standard deviation σ. (b) At low SNRs the situation is quite distinct from Gaussian. The Rician density depicted here corresponds to $s[m] = 1$ and $\sigma = 1$ (SNR = 0dB).*



(a) (b)

Fig. 3. *Reduction of contrast in MR magnitude image. (a) The solid pulse waveform denotes an ideal 1-d noise-free signal s. The dashed waveform depicts the mean value of an observation of this signal in Rician noise ($\sigma = 5$). In this example, the SNR ≈ 6 dB, and the contrast of the noise-free signal is $c = 0.500$. The contrast of the mean of the observed signal is $c = 0.455$. Hence, the contrast of the pulse is reduced by 9% due to the bias introduced by Rician noise. (b) The curve represents the contrast of the mean of the observed signal relative to the contrast of the noise-free signal. As the SNR increases the degradation of contrast becomes negligible. At low SNRs the contrast is significantly lower in the noisy observation.*

(a)

(b)

(c)

(d)

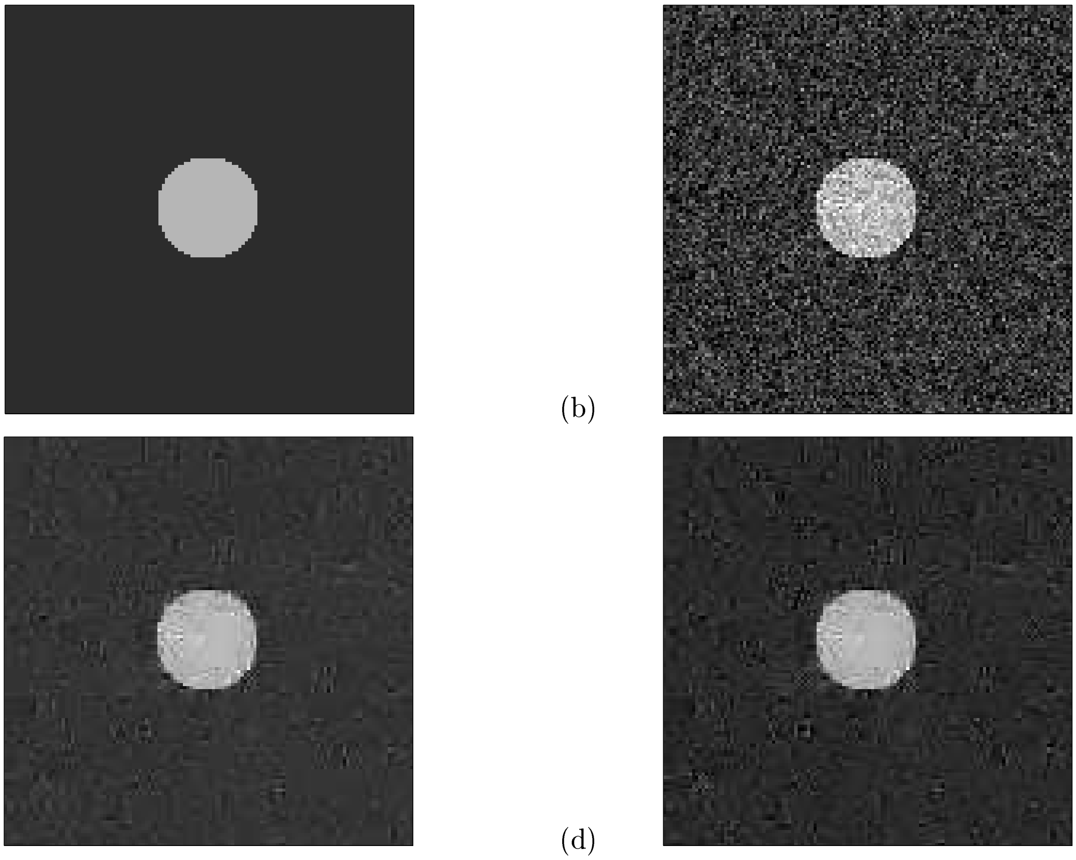
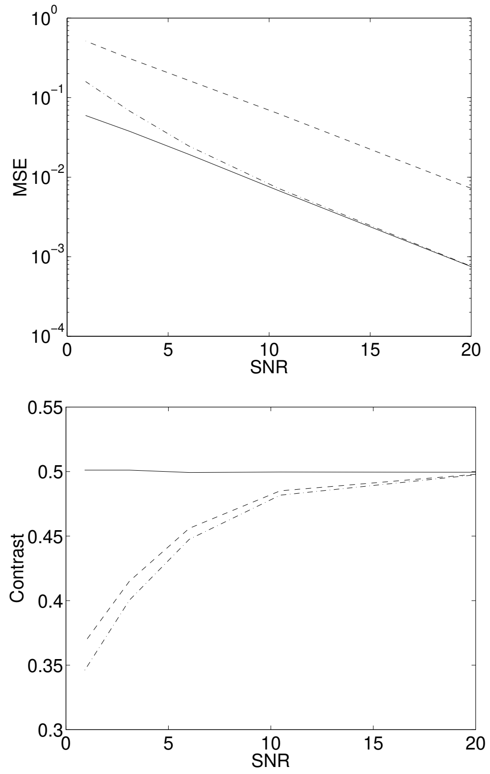


Fig. 4. *Reduction of contrast in MR magnitude image. (a) Noise-free signal (bright disc against dark background), contrast* $c = 0.50$. *(b) Image with Rician noise, SNR = 6dB, normalized sum of squared pixel errors* $= 1.00$. *(c) Filtered image using Algorithm I, contrast* $c = 0.44$ *(computed using* (5.1)*), squared error* $= 0.16$. *(d) Filtered image using Algorithm II, contrast* $c = 0.49$*, squared error* $= 0.13$. *The noise is considerably reduced in both (c) and (d) compared to (b). Note also that the contrast is much higher in (d) compared to (c).*



(a)

(b)

Fig. 5. *Contrast and noise reduction performance of wavelet-domain filtering algorithms. (a) Mean-square error (MSE) vs. SNR for noisy image (dash-dash), Algorithm I filtered image (dash-dot), Algorithm II filtered image (solid). Note that MSE of both filtered images is much lower than the MSE of the original noisy image. (b) Contrast vs. SNR for noisy image (dash-dash), Algorithm I filtered image (dash-dot), Algorithm II filtered image (solid). The contrast of the Algorithm II filtered image is significantly higher than that of the noisy image and the Algorithm I filtered image.*

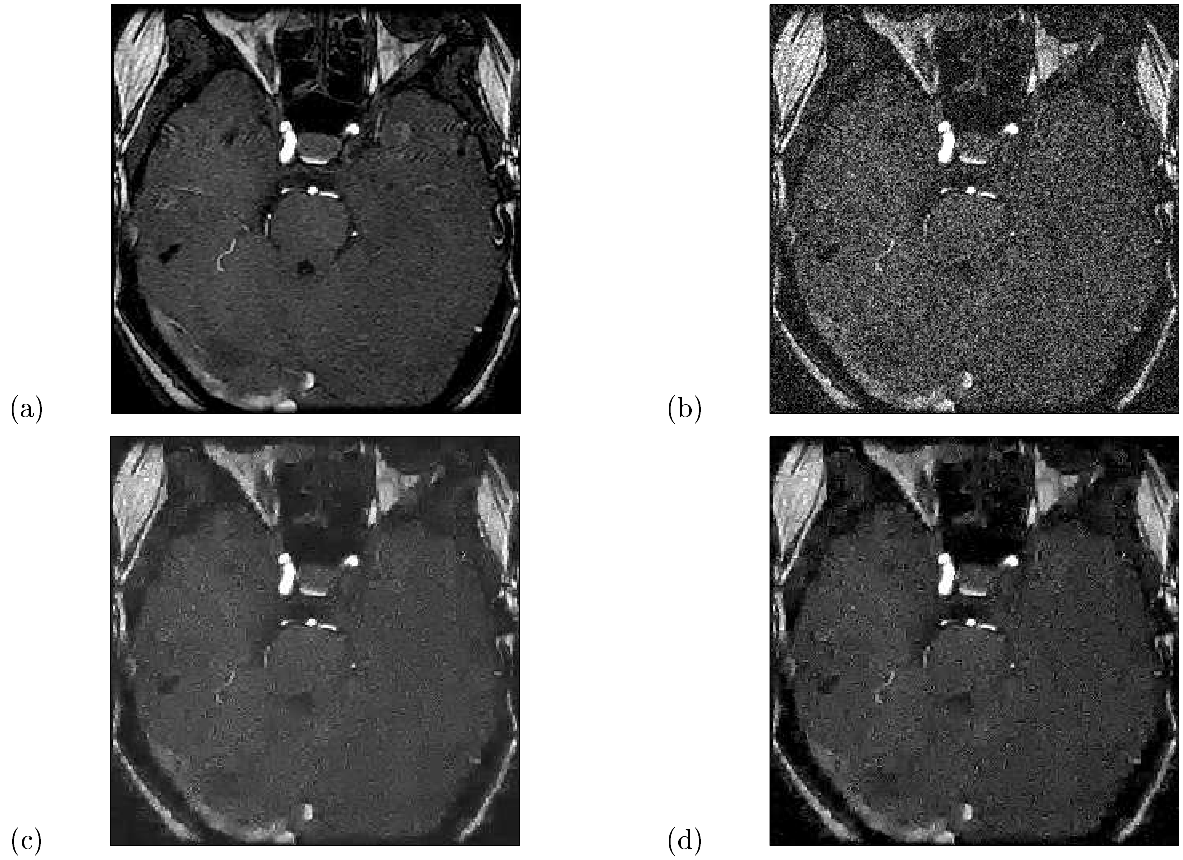


Fig. 6. *Wavelet-domain filtering to improve low SNR MRI data. (a) High SNR MR brain image. (b) Simulated low SNR image with Rician noise. The normalized sum of squared pixel errors between (a) and (b) is* 1.00*. (c) Filtered image resulting from Algorithm I, squared error 0.46. (d) Filtered image resulting from Algorithm II, squared error 0.25. Both Algorithm I and II remove much of the noise in low SNR image. Because Algorithm II accounts for the bias, its squared error performance is superior to that of Algorithm I, and the contrast in (d) appears to be higher than in (c). To assess the contrast improvement gained using Algorithm II, consider the bright blood vessel just down and slightly left of center. The contrast is computed by comparing the mean vessel pixel value to the mean background pixel level in a* 32×32 *subimage containing the vessel. The vessel contrast is* $c = 0.194$ *in (c), and* $c = 0.247$ *in (d). Hence in comparison to Algorithm I, Algorithm II improves the contrast of this vessel by over 20%.*

(a) 

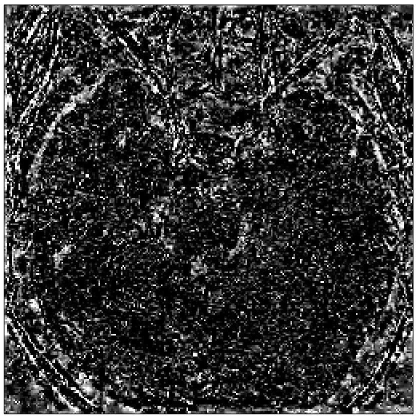
(b) 

Fig. 7. *Errors in filtered MRI data. (a) Difference between high SNR MR image and filtered image produced by Algorithm I (Fig. 7 (a) subtracted from Fig. 7 (c)). (b) Difference between high SNR MR image and filtered image produced by Algorithm II (Fig. 7 (a) subtracted from Fig. 7 (d)). The bias error in Algorithm I is clearly evident by the brighter appearance of (a) compared to (b) above.*

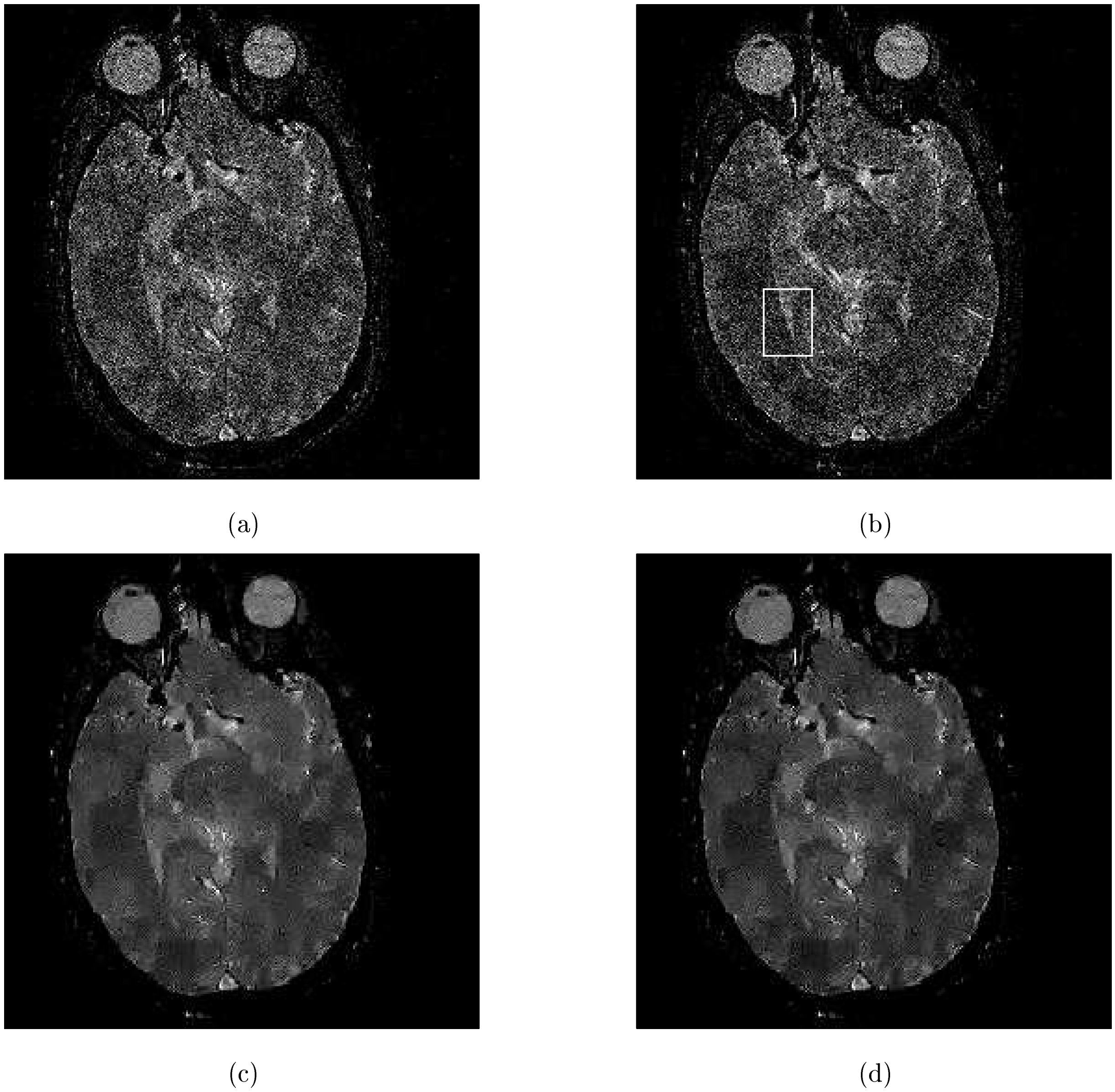


Fig. 8. *Filtering low SNR MRI data. (a) Low SNR MR brain image. SNR in the low intensity region of the brain is approximately 9dB. (b) Region of interest (ROI) highlighted for contrast comparisons. (c) Filtered image using Algorithm I. (d) Filtered image using Algorithm II. Both algorithms significantly reduce the noise in the image. The contrast in (d) appears better than that of (c) because Algorithm II accounts for the bias introduced by the noise in the magnitude image. Using the ROI indicated in (b), the contrast of the ROI in (c) is 0.117. The contrast of the ROI in (d) is 0.129. The contrasts are computed according to* (5.2).