A wavelet multiscale denoising algorithm for magnetic resonance (MR) images

Xiaofeng Yang[1,2] and Baowei Fei[2]

[1] Department of Biomedical Engineering, School of Life Science and Technology,
Xi'an Jiaotong University, Xi'an, Shaanxi 710049, People's Republic of China
[2] Department of Radiology, Emory University, Atlanta, GA 30329, USA

E-mail: xyang43@emory.edu

Abstract
Based on the Radon transform, a wavelet multiscale denoising method is proposed for MR images. The approach explicitly accounts for the Rician nature of MR data. Based on noise statistics we apply the Radon transform to the original MR images and use the Gaussian noise model to process the MR sinogram image. A translation invariant wavelet transform is employed to decompose the MR 'sinogram' into multiscales in order to effectively denoise the images. Based on the nature of Rician noise we estimate noise variance in different scales. For the final denoised sinogram we apply the inverse Radon transform in order to reconstruct the original MR images. Phantom, simulation brain MR images, and human brain MR images were used to validate our method. The experiment results show the superiority of the proposed scheme over the traditional methods. Our method can reduce Rician noise while preserving the key image details and features. The wavelet denoising method can have wide applications in MRI as well as other imaging modalities.

Keywords: magnetic resonance imaging (MRI), multiscale denoising, Rician distribution, Radon transform, wavelet transform, translation invariant

(Some figures in this article are in colour only in the electronic version)

1. Introduction

There is a practical limit on the signal-to-noise ratio (SNR) when acquiring MR image data [1]; post-processing methods to remove noise are important. Normally, denoising methods use the signal averaging principle which is based on the natural spatial pattern redundancy in the images. Gaussian filters have been widely used in many applications such as functional MR imaging (fMRI) [2]. However, they have the disadvantage of blurred edges due to averaging nonsimilar patterns. In order to avoid this problem, many edge-preserving filters have been proposed. One example is an anisotropic diffusion filter (ADF) [3–6]. Such a filter preserves edges by averaging pixels in the orthogonal direction of the local gradient. However, the denoising methods usually erase small features and change image statistics due to their edge enhancement effect.

Another approach for denoising relies on statistical inference of a multiscale representation of images. A prominent example includes methods based on wavelet transforms [7–9]. For denoising MR images, wavelet techniques based on soft thresholding were first applied by Healy [10]. In another approach, a wavelet-based Wiener-filter-like denoising method was used [11] where the magnitude of the MR image was squared and the square of a Rician random variable was modeled by a scaled noncentral chi-square distribution. The prior knowledge of the correlation of wavelet coefficients was used to represent significant features across scales [12]. A wavelet denoising method was compared with Gaussian smoothing methods [13]. A Wiener-like-filtering method was applied in the wavelet domain before the reconstruction of MR images [14]. However, typical wavelet-based methods can produce significant artifacts in the processed images because of the structure of the underlying wavelets.

Other denoising methods include a maximum posteriori estimation technique. Those methods account for Rician noise through a data likelihood term and a spatial smoothing prior [15]. Awate [16] used empirical Bayes for denoising in MRI. The method uses a Markov probability density function (PDF) to estimate the observed corrupted data and thus use it as a

prior in the Bayesian. In this way, the Bayesian denoising scheme bootstraps itself by estimating the prior through the optimization of an information theoretic metric using the expectation maximization (EM) algorithm. The parametric filter named non-local means (NLM) for random noise removal is analyzed and adapted to reduce the noise in MR images [17]. Jose [18] proposed a filter to reduce random noise in multi-component MRI by spatially averaging similar pixels using information from all available image components in order to perform the denoising process.

In this paper, a wavelet domain denoising procedure based on the Radon transform is proposed for MR images. The approach explicitly accounts for the Rician nature of the data. We employ the Radon transform to the original MRI data in order to transform it to the Radon domain. This approach is particularly useful to denoise dark regions because of the noise bias in low SNR (dark) regions. Furthermore, a translation invariant wavelet transform is employed to decompose the sinogram into multiscales for denoising step by step. Finally, we apply the inverse Radon transform to reconstruct the original MR images from the denoised sinogram. In the following sections, we describe our methods as well as our results from phantom, simulation brain MRI, and real patient MRI.

2. The distribution of noisy MRI data

One main source of noise in an MRI signal is the thermal noise [19]. The signal component of the measurements is present in both real and imaginary channels; each of the two orthogonal channels is affected by white Gaussian noise [20]. An MR image is usually reconstructed by computing the inverse discrete Fourier transform of the raw data. The noise in the reconstructed complex data is thus complex white Gaussian noise. The magnitude of the reconstructed MR image is used for visual inspection and for automatic computer analysis. Since the magnitude reconstruction is simply the square root of the sum of two independent Gaussian random variables, the magnitude image data are described by a Rician distribution. The term Rician noise is used to refer to the error between the underlying image intensities and the observed data [21]. Rician noise is not zero; meanwhile the mean depends on the local intensity in the image [22].

If the real and imaginary data, with mean values A_R and A_I, respectively, are corrupted by zero mean Gaussian, stationary noise with the standard deviation σ, the probability distribution function of the magnitude data will be a Rician distribution, as described by

$$P_{\mathrm{mag}}(M) = \frac{M}{\sigma^2} \mathrm{e}^{-(M^2+A^2)/2\sigma^2} I_0(AM/\sigma^2), \qquad (1)$$

where I_0 is the modified Bessel function of the first kind with order zero. The image pixel intensity in the absence of noise is denoted by A, and the measured pixel intensity by M. Here A is given by $A = \sqrt{A_R^2 + A_I^2}$. Equation (1) is plotted in figure 1 for different values of the SNR, A/σ. Note that the Rician distribution tends to be a Rayleigh distribution when the SNR goes to zero and approaches a Gaussian distribution

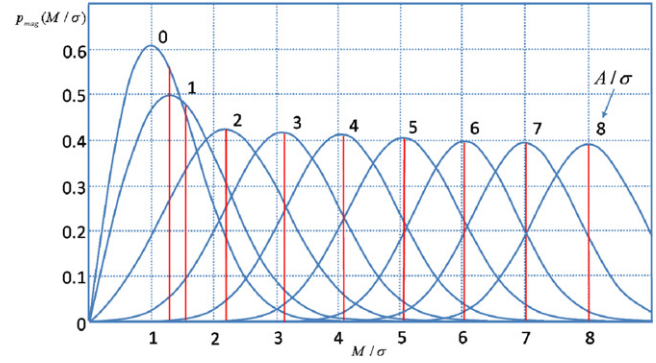**Figure 1.** The Rician distribution of M for several SNRs (A/σ) and the corresponding means.

at a high SNR. So in low-intensity (dark) regions on an MR image, the Rician distribution tends to be a Rayleigh distribution. In high-intensity (bright) regions, it tends to be a Gaussian distribution. For ratios as small as $A/\sigma = 3$ (SNR $= 10\log_{10} 3^2 \approx 10$ dB), it starts to approximate the Gaussian distribution. For a small SNR ($A/\sigma \leqslant 1$) the Rician distribution is far from a Gaussian distribution. Note that the mean of the distributions, $\bar{M}/\sigma$, which is shown by the vertical lines in figure 1, is not the same as A/σ. This bias is due to the nonlinear transform of the noisy data.

When the SNR is high ($A/\sigma \to \infty$), an interesting limit of equation (1) is described as

$$P_{\mathrm{mag}}(M) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(M^2-\sqrt{A^2+\sigma^2})^2/2\sigma^2}$$

$$\approx \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(M^2-A)^2/2\sigma^2}. \qquad (2)$$

This equation shows that for image regions with high signal intensities the noisy data distribution can be considered as a Gaussian distribution with variance σ^2 and mean A. Hence, in high SNR regions the noise can be viewed as a Gaussian white noise with variance σ^2 and zero mean. Based on such an assumption, separation of a signal and noise is fairly straightforward in the wavelet domain. However, the Gaussian approximation may introduce error for regions with a low SNR. As the mean of the magnitude image is not equal to the noise-free image, the magnitude image is biased.

3. Materials and methods

Our method includes four steps. First, we transform the MR image to the Radon domain using the Radon transform. Second, the wavelet transform is employed to decompose sinogram into multiscales. Third, noise variance is evaluated and a thresholding-based method is applied to denoise. Fourth, we reconstruct the denoised sinogram to get the original MRI using the inverse Radon transform.

3.1. Radon transform

The Radon transform of a 2D function is defined as

$$Rf(\alpha, s) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\delta(s - x\cos\alpha - y\sin\alpha)\,\mathrm{d}x\,\mathrm{d}y$$

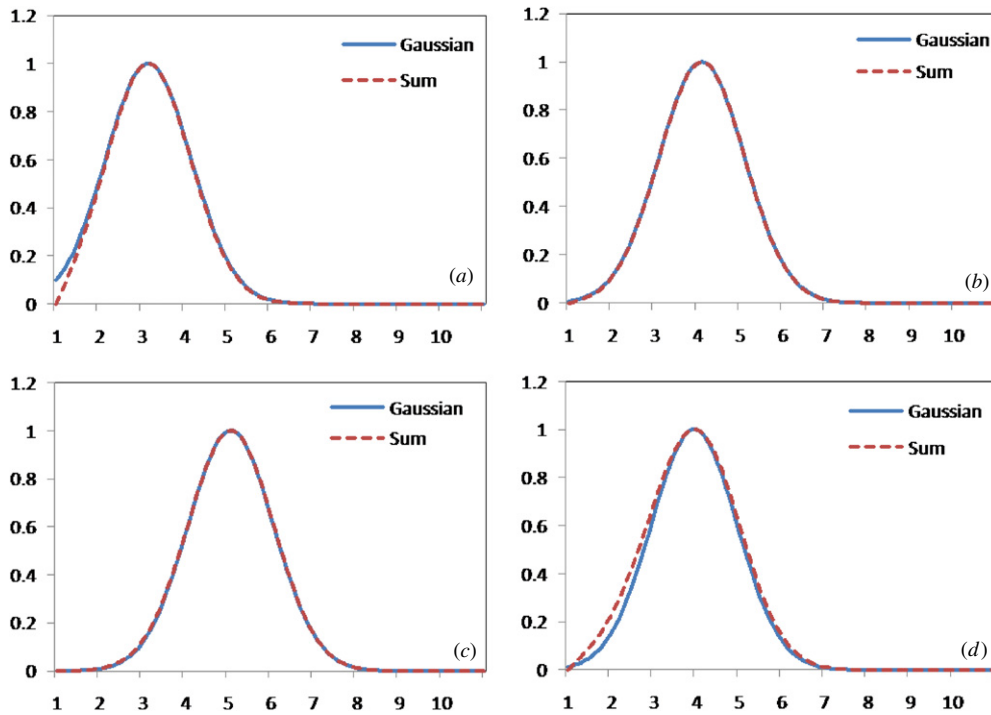$$\alpha \in [0, \pi] \quad s \in R, \qquad (3)$$

Figure 2. Gaussian PDF and the PDFs of the sum of two or more Rician distributed sets. In (a)–(d) the sums in sequence are: Sum $= P_{\text{mag}}(1) + 10{*}P_{\text{mag}}(2)$, Sum $= P_{\text{mag}}(2) + 10{*}P_{\text{mag}}(3)$, Sum $= P_{\text{mag}}(3) + 10{*}P_{\text{mag}}(4)$, and Sum $= P_{\text{mag}}(1) + 5{*}P_{\text{mag}}(2) + 5{*}P_{\text{mag}}(3) + 10{*}P_{\text{mag}}(4)$.

where s is the perpendicular distance of a line from the origin and α is the angle formed by the distance vector. According to the Fourier slice theorem, this transformation is invertible. The Fourier slice theorem states that for a 2D function $f(x, y)$, the 1D Fourier transforms of the Radon transform along s are the 1D radial samples of the 2D Fourier transform of $f(x, y)$ at the corresponding angles [23].

3.2. The distribution of noise in sinogram data

Rician noise differs from Gaussian noise in that it depends on the signal intensity, and the PDF of the noise is very asymmetric for low signal intensities. In brain MR images different regions have different intensities, i.e. white matter (WM), gray matter (GM) and scalp have high intensity, and the noise in these regions is very close to a Gaussian distribution. But other regions such as skull, nasal sinuses and cerebrospinal fluid (CSF) have very low intensity [13], so the noise in these regions is close to a Rayleigh distribution. We sum these regions along a line and thus sum the noise with different distributions. The sum of several Rician distributed noises has a symmetric distribution. Such a sum operation, which is the Radon transform, makes the distribution of noise close to a Gaussian distribution. In figure 2, we compare the sum distribution with the Gaussian distribution with $N(\mu, \sigma^2)$. On brain MR images, the low-intensity regions have fewer areas than those with high intensity. For example, the ratio of high-intensity regions to low-intensity regions can be more than 10 along a projection line during the Radon transform. Figure 2 shows the direct comparison between the Gaussian PDF and the sum PDF.

In order to measure the similarity degree (R^2), we use the following equation:

$$R(x, y) = \text{Cov}(x, y)/[\text{StdDev}(x) \cdot \text{StdDev}(y)], \qquad (4)$$

where Cov is the correlation function and StdDev is the standard variance. This equation is used to measure how well a regression line approximates the real data points statistically, e.g. R^2 of 1.0 (100%) indicates a perfect fit. From figure 3 we can see that the R^2 value is close to 1, indicating that the distribution of noisy data is close to a Gaussian function. At the same time we can see that the larger the A/σ the more similar to Gaussian the sum PDF is. In order to demonstrate this conclusion further we calculated an additional 24 sets of the similarity degree (R^2) between the real Gaussian PDF and the PDFs of the sum of two or more Rician distributed sets, and it is shown in table 1. In the table we can see that all R^2s are over 92%, and all R^2s except $P_{\text{mag}}(0) + 5{*}P_{\text{mag}}(1)$ and $P_{\text{mag}}(0) + 10{*}P_{\text{mag}}(1)$ are over 95%. Here $P_{\text{mag}}(0)$ is the PDF of the magnitude data in the air part inside MR images and not background, because we got rid of noise in the background using a brain mask before the Radon transform. So it is reasonable to use techniques based on the assumption of a Gaussian distribution for sinogram images. In fact, for the heart and brain images, the bright regions in body images have a SNR of over 20 dB ($A/\sigma > 10$) [24]. Based on the above-mentioned two facts and equation (2) we demonstrated that the noise in sinogram images is Gaussian noise with mean μ_s and variance σ_s^2. Here from the Radon transform we can obtain $\mu_s = n_s\mu$ and $\sigma_s^2 = n_s\sigma^2$, and it varies with the number of pixels (n_s) along a line through the MR brain image.
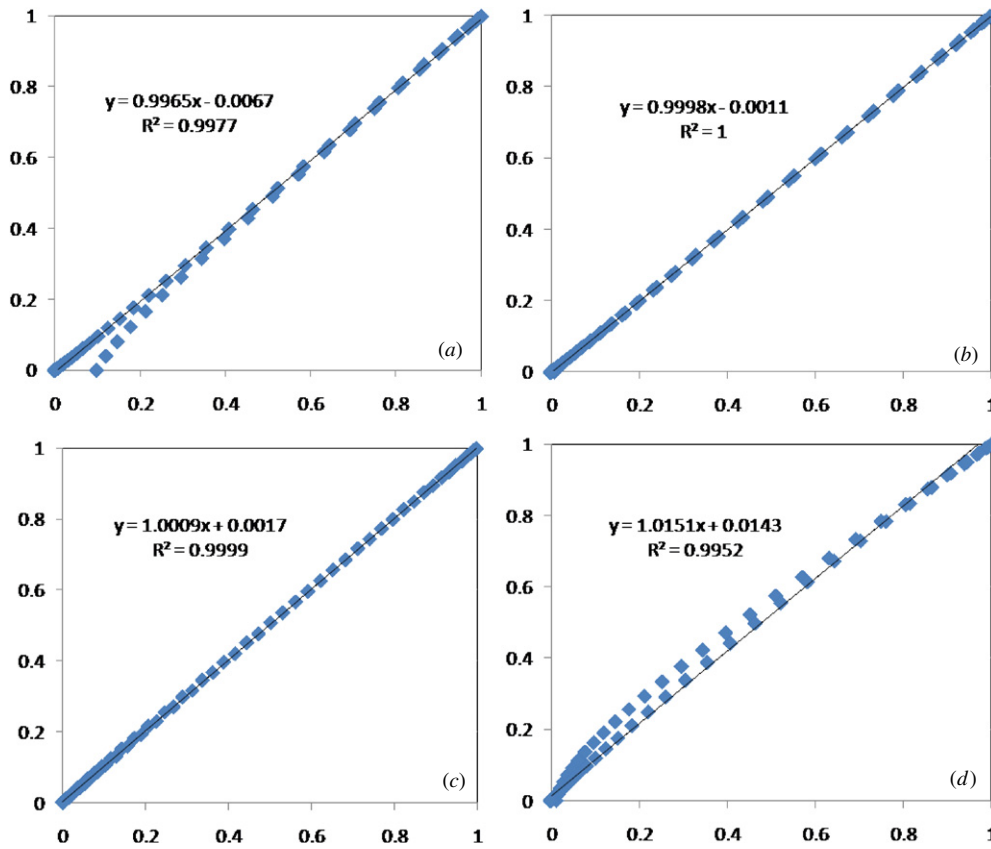
Figure 3. Correlation plots between the real Gaussian PDF and PDFs of the sum of two or more Rician distributed sets. (a)–(d) are the corresponding correlation plots of (a)–(d) in figure 2, respectively. The horizontal axis is the real Gaussian distribution data and the vertical axis is the distribution sum data.

Table 1. The similarity degrees (R^2s) between the real Gaussian PDF and PDFs of the sum of two or more Rician distributed sets.

Sum	R^2	Sum	R^2	Sum	R^2
$P_{mag}(0) + 5{*}P_{mag}(1)$	0.9232	$P_{mag}(0) + 5{*}P_{mag}(1) + 10{*}P_{mag}(2)$	0.9509	$P_{mag}(1) + 10{*}P_{mag}(4)$	0.9902
$P_{mag}(0) + 5{*}P_{mag}(2)$	0.9556	$P_{mag}(0) + 5{*}P_{mag}(2) + 10{*}P_{mag}(3)$	0.9898	$P_{mag}(1) + 5{*}P_{mag}(2) + 10{*}P_{mag}(3)$	0.9888
$P_{mag}(0) + 5{*}P_{mag}(3)$	0.9952	$P_{mag}(0) + 5{*}P_{mag}(3) + 10{*}P_{mag}(4)$	0.9845	$P_{mag}(1) + 5{*}P_{mag}(3) + 10{*}P_{mag}(4)$	0.9836
$P_{mag}(0) + 5{*}P_{mag}(4)$	0.9516	$P_{mag}(0) + 5{*}P_{mag}(1) + 5{*}P_{mag}(2) + 10{*}P_{mag}(3)$	0.9590	$P_{mag}(2) + 5{*}P_{mag}(3)$	0.9964
$P_{mag}(0) + 10{*}P_{mag}(1)$	0.9235	$P_{mag}(1) + 5{*}P_{mag}(2)$	0.9541	$P_{mag}(2) + 5{*}P_{mag}(4)$	0.9750
$P_{mag}(0) + 10{*}P_{mag}(2)$	0.9596	$P_{mag}(1) + 5{*}P_{mag}(3)$	0.9953	$P_{mag}(2) + 10{*}P_{mag}(4)$	0.9956
$P_{mag}(0) + 10{*}P_{mag}(3)$	0.9971	$P_{mag}(1) + 5{*}P_{mag}(4)$	0.9501	$P_{mag}(2) + 5{*}P_{mag}(3) + 10{*}P_{mag}(4)$	0.9907
$P_{mag}(0) + 10{*}P_{mag}(4)$	0.9902	$P_{mag}(1) + 10{*}P_{mag}(3)$	0.9975	$P_{mag}(3) + 5{*}P_{mag}(4)$	0.9993

3.3. Wavelet transforms

3.3.1. Wavelet definition. Wavelets are mathematical functions that decompose data into different frequency components that can be studied with a resolution matched to their scale. Wavelet transforms are multiresolution representations of signals and images. They decompose a signal into a hierarchy of scales ranging from the coarsest scale to the finest one. Wavelet coefficients of a signal are the projections of the signal onto the multiresolution subspaces. Wavelets are functions generated from one single function (basis function) called the prototype or mother wavelet by dilations (scalings) and translations (shifts) in the time (frequency) domain. If the mother wavelet is denoted by $\psi(t)$ other wavelets $\psi_{a,b}(t)$ can be represented as

$$\psi_{a,b}(t) = 1/\sqrt{|a|} * \psi((t-b)/a), \qquad (5)$$

where a and b are two arbitrary real numbers. The variables a and b represent the parameters for dilations and translations, respectively.

3.3.2. Dyadic discrete wavelet transform. The discrete wavelet transform (DWT) is an implementation of the wavelet transform using a discrete set of the wavelet scales and translation obeying some defined rules. For practical computations, it is necessary to discretize the wavelet transform. The scale parameter a is discretized on a logarithmic grid. The translation parameter b is then discretized with respect to the scale parameter, i.e. sampling is done on the dyadic sampling grid (as the base of the logarithm is usually chosen as 2). The dyadic wavelet transform makes the scale factor binary discrete, while the translation factor

maintains continuous change. The discretized scale and translation parameters are given by $a = 2^{-j}$ and $b = k2^{-j}$, respectively, where $j, k \in Z$, the set of all integers. Thus, the family of wavelet functions is represented as

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k). \quad (6)$$

From the multiresolution point of view, the wavelet decomposition of a discrete time signal $x[n]$ is given by

$$x[n] = \sum_k c_\varphi(j_0, k)\varphi_{j_0,k}(n) + \sum_{j=j_0}^{\infty} \sum_k d_\psi(j, k)\psi_{j,k}(n), \quad (7)$$

where $\varphi_{j_0,k}$, $\psi_{j,k}$ are the scaling functions and wavelet functions, respectively.

The scaling/approximation and the wavelet/detail coefficients are given respectively as

$$c_\varphi(j_0, k) = \sum_n x[n]2^{j_0/2}\varphi(2^{j_0}n - k), \quad (8)$$

$$d_\psi(j, k) = \sum_n x[n]2^{j/2}\psi(2^j n - k), \quad (9)$$

where j_0 is the starting scale and is always 0. c_φ and d_ψ are the scale factor and the wavelet coefficients, respectively, and j is the decomposition level. The development of the fast wavelet transform [25] concludes its identity to two-channel subband decomposition. Thus, it reveals a remarkable relationship between the DWT coefficients of adjacent scales. Using the coefficients d_ψ at a specific level $j + 1$ we can calculate the coefficients at level j using a filter bank. The wavelet decomposition of a 2D signal can be achieved by applying the 1D wavelet decomposition along the rows and columns of the image separately. This is equivalent to projecting the image onto separable 2D basis functions obtained from the products of 1D basis functions

3.3.3. Translation invariant wavelet. Since the DWT provides good localization in both spatial and spectral domains, low pass filtering is inherent to this transform. The DWT is computationally efficient. The only drawback is that it is not translation invariant, which can introduce artifacts during image reconstruction and exhibit Gibbs phenomena in the neighborhood of discontinuities because of the lack of translation invariance of the wavelet basis. The translation variance in the discrete wavelet transform is due to the required decimation operation (downsampling by two). This problem can be solved by applying additional discrete wavelet decomposition after shifting the sequence by one sample [26]. From the Radon transform we know that the translation along α in the Radon domain corresponds to the rotation of the input image. Although a translation invariant wavelet transform seems to be useful for this application, its application in both directions (s and α) leads to suboptimal results compared with a non-translation invariant wavelet transform. Although the circular shift along α corresponds to the rotation of the image, the circular shift along s does not correspond to a regular geometric distortion. The shift along s in the Radon domain corresponds to an image significantly different from the original image. To solve this problem, we only apply a 1D translation invariant wavelet transform along s [27, 28]. In section 3.5 we will discuss some properties using a 1D wavelet transform along s.

3.3.4. Threshold-based denoising in the wavelet domain. The wavelet coefficients of signals after a wavelet transform contain important information and the wavelet coefficients of noise correspond to smaller amplitude. In our method, a suitable threshold value is selected through different scales, and the wavelet coefficients less than the threshold are set to zero, while retaining the wavelet coefficients greater than the threshold. So the noise signal is effectively inhibited. Finally, the denoised signal is reconstructed using the wavelet inverse transform. It is well known that for independent and identically-distributed Gaussian noise $x \sim N(\mu_s, \sigma_s^2)$; a threshold $\beta = \mu_s + \sigma_s, \mu_s + 2\sigma_s, \mu_s + 3\sigma_s, \ldots$ will suppress 68.26%, 95.44%, 99.74% of its values. Therefore, we choose $\beta = \mu_s + c\sigma_s$, where c is a constant. By taking c between 3 and 4, we can achieve good results. Based on the fact that the variance in each wavelet scale is also σ_s^2 in an orthogonal transformation we can obtain the final threshold

$$\beta = \mu_s + c\sigma_s = n_s\mu + cn_s\sigma. \quad (10)$$

Under the Gaussian noise assumption, thresholding techniques successfully utilize the unitary transform property of the wavelet decomposition to distinguish statistically the signal components from those of the noise. Our desire, then, is to remove the estimated noise contribution β from each of the wavelet coefficient, and obtain the estimated signal. Although Donoho [29] proved the optimality of soft threshold in theory, Stein threshold has shown better results in SNR improvement. Thus

$$Wf(\hat{x}_i) = \max(1 - ((n_s\mu + cn_s\sigma)/|Wf(y_i)|)^2, 0)Wf(y_i), \quad (11)$$

where $\hat{x}_i$ is the denoised signal, y_i denotes the noisy observations and $Wf(y_i)$ each of the wavelet coefficients. Finally, we perform the wavelet inverse transform for $Wf(x_i)$ in order to obtain the denoised signal $\hat{x}_i$.

3.4. Noise mean μ and variance σ^2 estimation

Our algorithms require the underlying noise variance σ^2, which is usually unknown and has to be estimated from the data. Typical MR images include an empty region of air outside the patient. A simple estimator is based on the following argument.

From equation (1) we can obtain a special case of the Rician distribution in image regions where only noise is present, $A = 0$. This is better known as the Rayleigh distribution, and equation (1) reduces to

$$P_{\text{mag}}(M) = \frac{M}{\sigma^2}e^{-M^2/2\sigma^2} \quad (12)$$

This Rayleigh distribution governs noise in image regions without signals. The mean for this distribution can be evaluated analytically and is given by

$$\bar{M} = \sigma\sqrt{\pi/2}. \quad (13)$$

These relations can be used to estimate the 'true' noise variance σ^2 from the magnitude image. So from equation (13) the pixel values in the region outside the patient provide us with a very reliable estimator:

$$\sigma = \sqrt{2}\bar{M}/\sqrt{\pi} \quad \text{and} \quad \mu = \bar{M}. \quad (14)$$

3.5. Properties of our method

The 1D wavelet transform of an image is equal to the projection of the image onto the 1D basis functions. If we denote the Radon transform of the image by $Rf(\alpha, s)$, the wavelet transform coefficients by $Wf(\cdot, \cdot)$, and the corresponding 1D orthogonal wavelet basis functions by $\phi(s)$, then

$$
\begin{aligned}
Wf(\cdot, \cdot) &= \int_0^\infty Rf(\alpha, s)\phi(s)\mathrm{d}s \\
&= \int_0^\infty \left[\int_{-\infty}^\infty \int_{-\infty}^\infty f(x, y)\delta(s - x\cos\alpha \right. \\
&\quad \left. - y\sin\alpha)\,\mathrm{d}x\,\mathrm{d}y \right]\phi(s)\,\mathrm{d}s \\
&= \int_{-\infty}^\infty \int_{-\infty}^\infty f(x, y)\int_{-\infty}^\infty [\delta(s - x\cos\alpha \\
&\quad - y\sin\alpha)\phi(s)\,\mathrm{d}s]\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_{-\infty}^\infty \int_{-\infty}^\infty f(x, y)\phi(x\cos\alpha + y\sin\alpha)\,\mathrm{d}x\,\mathrm{d}y. \quad (15)
\end{aligned}
$$

By defining $h(x, y) = \phi(x\cos\alpha + y\sin\alpha)$

$$
Wf(\cdot, \cdot) = \int_{-\infty}^\infty \int_{-\infty}^\infty f(x, y)h(x, y)\,\mathrm{d}x\,\mathrm{d}y, \quad (16)
$$

where $h(x, y)$ in 2D is defined from the wavelet function $\phi(s)$ in 1D, and it is similar to 2D wavelet transforms except that the point parameter is replaced by a line parameter. The wavelet is a function with scale and point position, while $h(x, y)$ is a function with scale and line position. Hence the wavelet is effective in representing point singularities, and $h(x, y)$ is effective in representing singularities along the line. So our transform can capture singularities along lines and edges in an efficient way.

3.6. Reconstruction

After obtaining the denoised sinogram, we perform the inverse Radon transform in order to obtain the original image. It is defined as

$$
f(x, y) = \int_0^\pi R(\alpha, x\cos\alpha + y\sin\alpha)\,\mathrm{d}\alpha, \quad (17)
$$

where R are the filtered projections. Generally, three different inverse Radon transform methods are the direct inverse Radon transform (DIRT), filtered back-projection (FBP) and convolution filtered back-projection (CFBP) [30]. The DIRT is computationally efficient, but it introduces some artifacts. FBP based on the linear filtering model often exhibits degradation in recovering from noisy data [31]. Spline-convolution filtered back-projection (SCFBP) offers better approximation performance than the conventional lower-degree formulation (e.g. piecewise constant or piecewise linear models) [32]. For SCFBP the denoised sinogram in the Radon domain is approximated in the B-spline space, while the resulting image in the image domain is in the dual-spline space. We used SCFBP to propagate the deniosed sinogram back into the image space along the projection paths.

4. Evaluation and results

Our denoising method has been evaluated by modified brain phantom and simulation brain data. We also applied the method to denoise real brain MR images. We compare with the optimum linear filter, the Wiener filter [33], which is solely adapted to a SNR at a single scale. Meanwhile, a traditional multiscale wavelet method is also applied to these datasets in order to compare with our method. Here Wiener and wavelet methods are used to denoise images in the image domain directly. Finally, in order to prove the efficiency of the proposed method quantitatively an average SNR is used as a quality metric. It is given by

$$
\mathrm{SNR} = 10\log_{10}\frac{\sum_{j,k}|x[j, k]|^2}{\sum_{j,k}|x[j, k] - \hat{x}[j, k]|^2}, \quad (18)
$$

where $x[j, k]$ is the original image $\hat{x}[j, k]$ is the denoised image, with results averaged over all images and reported as mean decibels (dB).

In order to validate our method Rician noise is added to the modified phantom and simulation brain MRI. But while denoising we can get a mask first through a threshold and eliminate the noise outside the phantom and brain. The noise outside the phantom and brain is only used to evaluate the noise variance. The mentioned SNR in the following section only expresses the SNR in the phantom or brain except in outside air. Here for the Wiener filter we used neighborhoods of size 3×3 to estimate the local image mean and standard deviation. 2D and 1D db3 wavelets were used in the wavelet filter and our method, respectively, and the images were decompounded into four levels.

4.1. Brain phantom data

Figure 4 illustrates the visual assessment of denoised results on the brain phantom with different noises and the comparison of three methods. In order to compare the denoised effectiveness Rician noise of different degrees is added to this phantom. The images from top to bottom are the phantoms with different degrees of noise, and the images from left to right are the original noised phantom, the denoised result after Wiener, the traditional wavelet and our method in sequence. It can be seen that the Wiener filter makes images a little blurred, and both the traditional wavelet and our method can reduce Rician noise effectively while retaining phantom detail, but our method can suppress Rician noise more.

Figure 5 shows the SNR plots between input SNRs and output SNRs for the three methods. From this figure it can be seen that our method can improve the SNR more than Wiener and the wavelet method. At the same time the output SNR after the Wiener filter grows almost linearly, while the other two methods can improve less with the increase in the input SNR in the original image, and approach the SNR curve of the Wiener filter because the noise is closer to Gaussian noise.

Figure 6 exhibits residuals between the original and denoised phantoms. Comparing these figures we can see that the Wiener filter smoothes the whole image and is uniform for all regions in the image. And the wavelet method performs
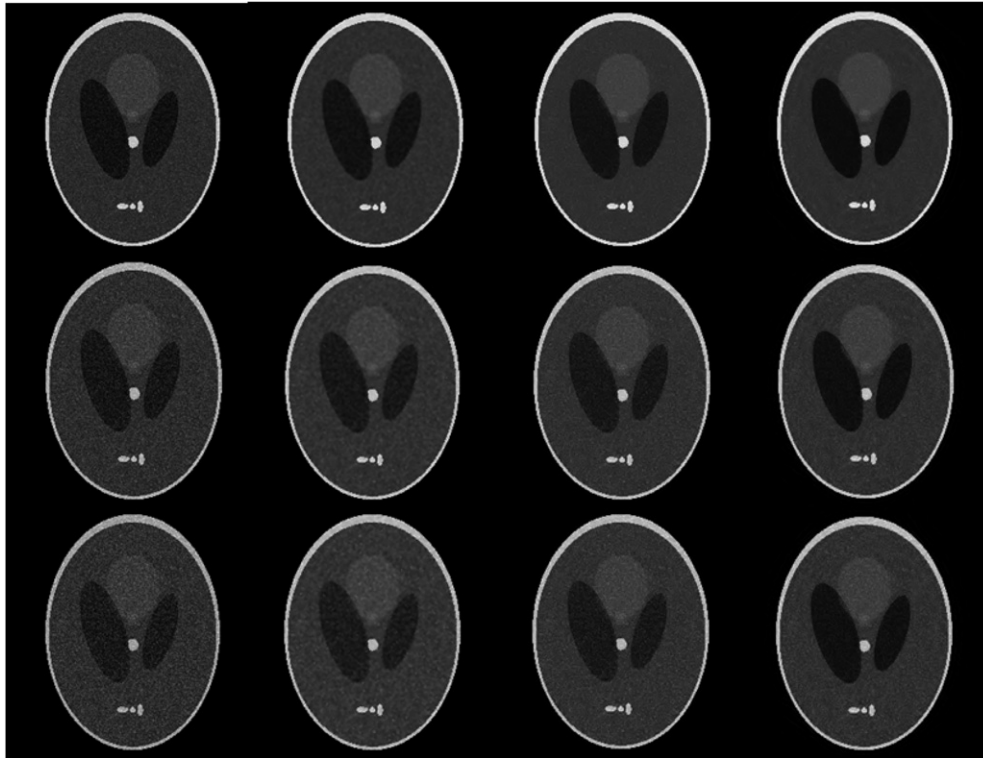
Figure 4. Denoised results of brain phantom with different noise using different methods. The first column is the phantom with different degrees of noise. The second column is the results after the Wiener filter. The third column is the results after wavelet. The fourth column is the results after our method.
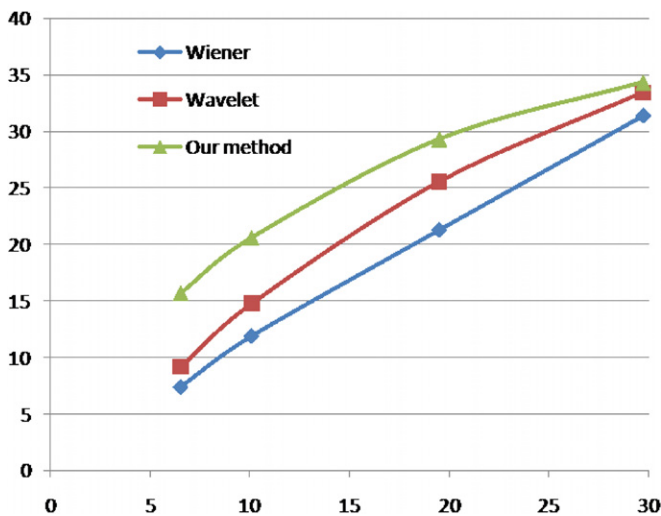

Figure 5. Input and corresponding output SNR plots for the three methods. The horizontal axis is the input SNR and the vertical axis is the output SNR.

better for bright regions than dark regions, namely wavelet denoising is better for Gaussian noise than Rician noise under the same condition. The dark part in the phantom is brighter than other parts in the wavelet residual figure. Finally, our method can decrease noise for both dark and bright regions, and its denoised effectiveness in dark regions is better than the other two methods.

In order to compare the results after different methods quantitatively, we obtain profiles through the original phantom,

the denoised phantom, and the phantom after the Wiener filter, the wavelet method and our method as shown in figure 7. The Wiener filter can only smooth a denoised image in the whole image, and in the bright region, the wavelet method and our method have almost the same effect because our method also applies wavelet to denoise. But for dark regions our method performs better than the wavelet method because of the noise distribution. The result after our method is closer to that of the original phantom.

In figure 8 we compare the difference of the denoised effect in the Radon domain. Noised, denoised and no-noise sinograms are shown from left to right in sequence. We obtain a denoised sinogram that is close to the noise-free sinogram visually, and it shows that our method has already decreased the noise in the sinogram.

4.2. Simulation brain data

We obtained the brain MR images from the McGill phantom brain database for comprehensive validation of the denoising methods. The MR volume contains $181 \times 217 \times 181$ voxels and covers the entire brain. Based on the realistic phantom, an MR simulator is provided to generate specified MR images. The MR tissue contrasts are produced by computing MR signal intensities from a mathematical simulation of the MR imaging physics [34].

As shown in figure 9, the same Rician noise is added to the whole volume and three slices are shown for comparison in this figure. By comparing the dark regions (e.g. skull and CSF) with the bright regions (e.g. WM, GM and scalp) in all
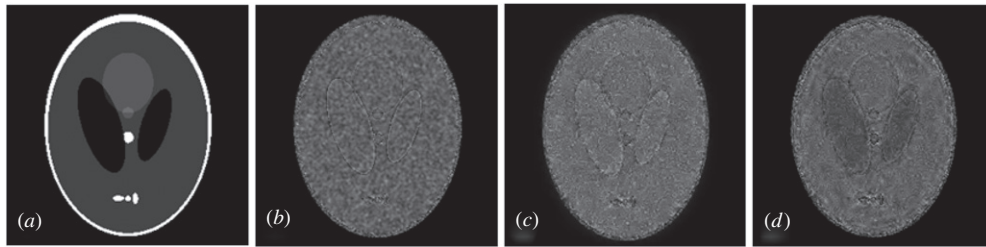
Figure 6. Residuals between the original phantom and denoised phantoms. (*a*) Original phantom; (*b*)–(*d*) the corresponding residuals between the denoised and original phantoms in sequence.
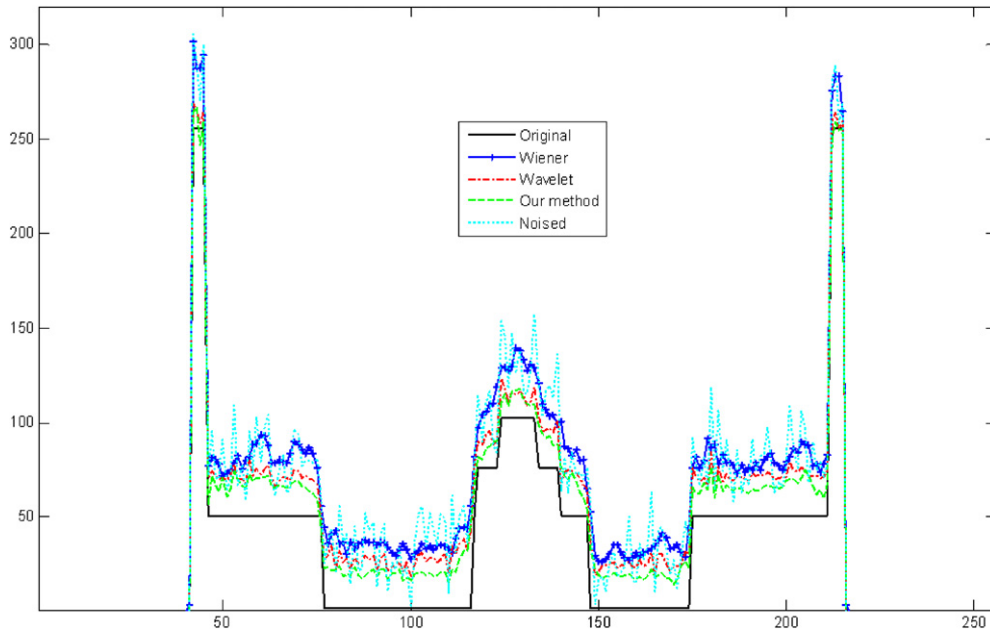

Figure 7. Comparison of horizontal profiles between the original phantom, noised phantom and phantoms denoised using Wiener, wavelet and our method.
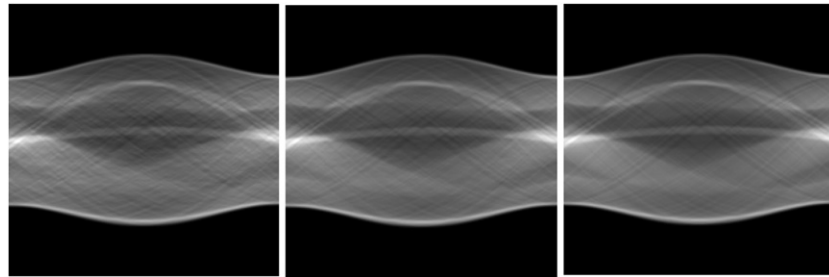

Figure 8. Comparison of the noised (left), denoised (middle) and noise-free (right) sinograms.

denoised images, it can be seen that our method enhances the image contrast and makes the edges clearer than the other two methods. The Wiener method smooths the images and makes whole images blurred. The wavelet method and our method both preserve image details but our method decreases noise more, and this point can be seen from the quantitative SNR. The average SNR of original simulation brain with Rician noise is 17.88 dB, the average SNR after Wiener is 19.33 dB, the average SNR after the wavelet method is 23.12 dB, and the average SNR after our method is 26.84 dB.

In the same way we obtain the residuals between the original image and denoised image after different methods as shown in figure 10. As discussed above, the Wiener filter smooths the whole image and its residuals are almost the same except for the edge parts, and the wavelet method is better for bright regions than dark regions, and our method can decrease noise for both dark and bright regions.

In figure 11, we quantitatively compare the profiles through the original image, noised image, and the image after the Wiener filter, wavelet and our method. In this figure, the Wiener filter only smooths the noised image in the whole
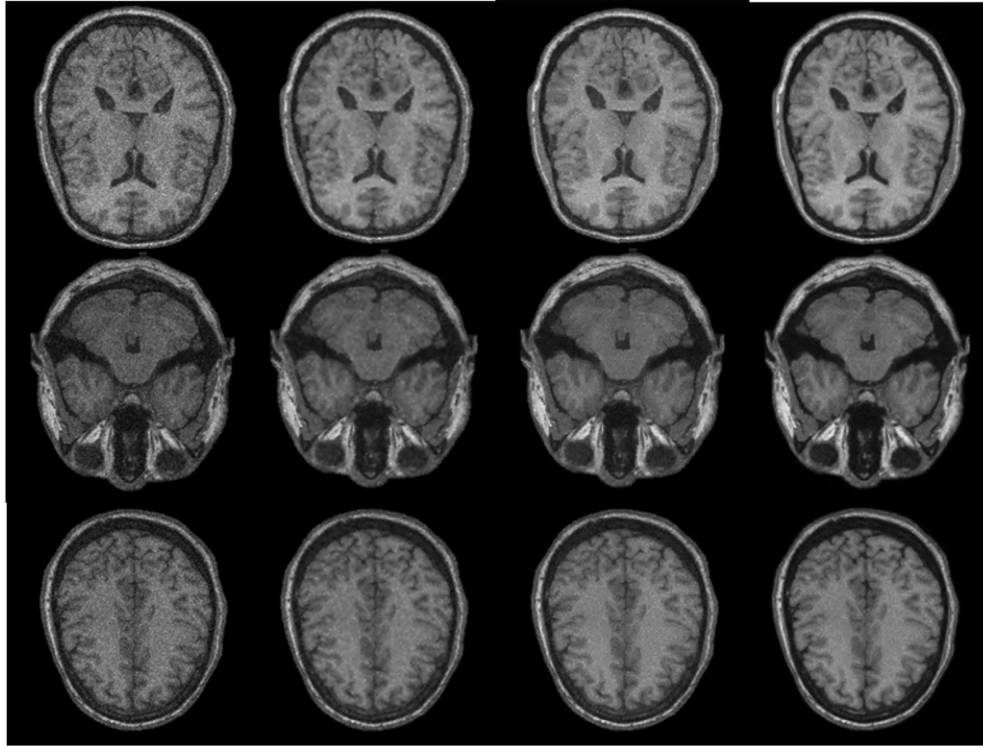
Figure 9. Denoised results of the simulated brain data using different methods. The first column is the noised brain images of different slices. The second column is the results after the Wiener filter. The third column is the results after wavelet. The fourth column is the results after our method.
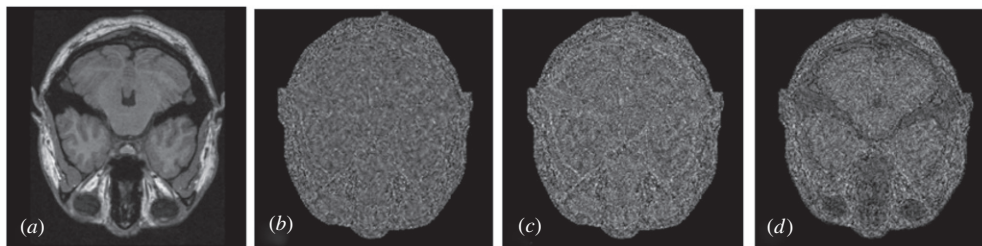

Figure 10. Residuals between the original image and denoised images. (*a*) Original image; (*b*)–(*d*) the corresponding residuals between the denoised and original images in sequence.

image and loses the image details. In bright regions, the result after the wavelet method is almost the same as the result after our method, but in dark regions our method excels the wavelet method. We can see the denoised image after our method is closest to the original image and it has the best denoised effect among the three methods.

4.3. Real brain data

The denoising method was also applied to real T1-weighted MR images of human brain. The MR images were acquired with a 4.0 T MedSpec MRI scanner (Bruker BioSpin GmbH, Rheinstetten, Germany) on a Siemens Syngo platform (Siemens Medical Systems, Erlangen, Germany). The T1-weighted magnetization prepared rapid gradient echo sequence (MPRAGE) (TR = 2500 ms and TE = 3.73 ms) was used for the image acquisition. The volume has $256 \times$ 256×176 voxels covering the whole brain yielding a 1.0 mm isotropic resolution.

Figure 12 illustrates qualitative comparison of denoised results on T1-weighted MR brain images and the comparison of different methods. We do not perform any preprocessing for the original MR brain image. In this figure our method makes the edges of the MR image clearer and enhances the image contrast. This can be obtained from the property of our method; it can capture singularities along lines and edges in an efficient way. Through comparing the three residuals we can see that the Wiener method makes the whole image smooth and that its residuals are bigger at the edge than in other regions. The traditional wavelet can decrease Rician noise better than Wiener, and it can decrease the noise in dark regions to some degree. In brighter regions with a high SNR, our result shows much less difference from the other two methods, and this means it does not affect the region with little noise. In dark
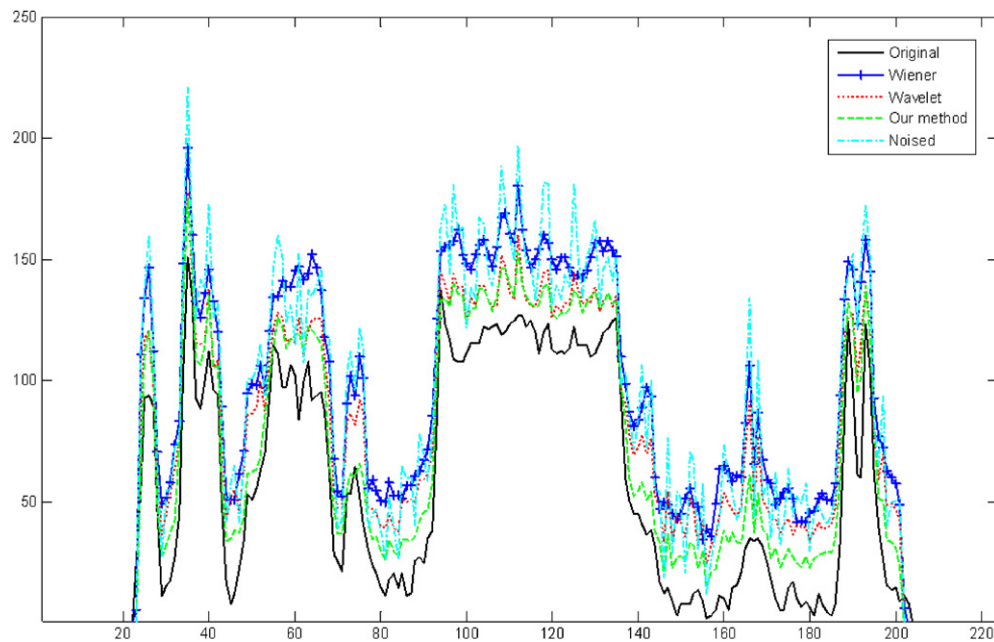
Figure 11. Comparison of horizontal profiles between the original image, noised image and images denoised using Wiener, wavelet and our method.
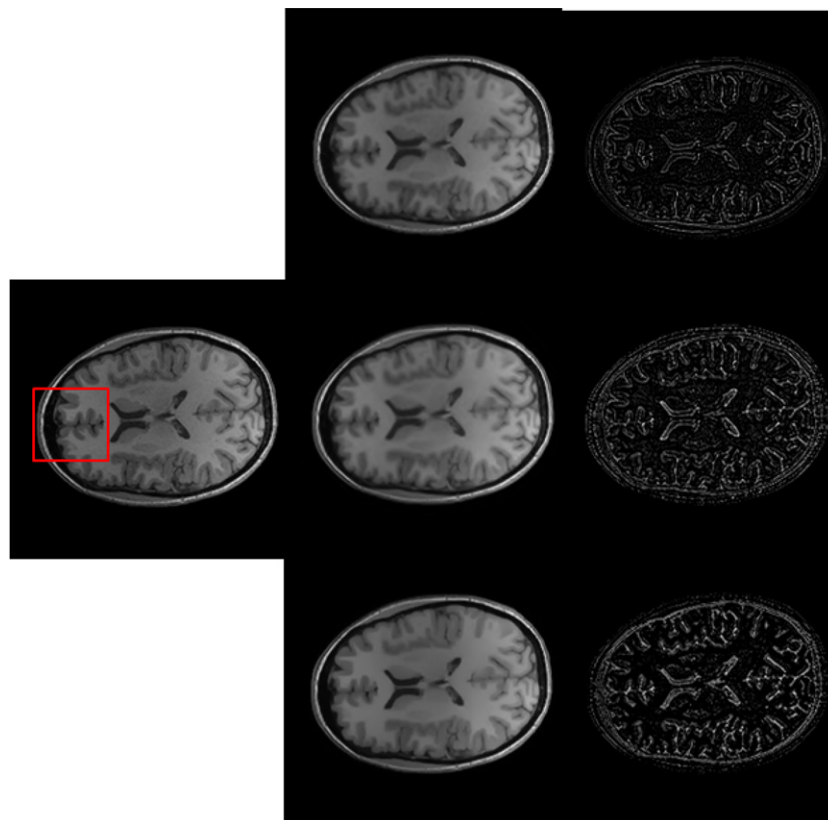

Figure 12. Qualitative comparison of denoising results obtained with different methods. The first column is the real MR brain image without any processing. From top to bottom, the second column is the denoised results by Wiener, wavelet, and our method, respectively. The third column is the corresponding residuals between the original image and denoised image.

regions that have a low SNR, our result shows bigger difference as compared to the original image than the other two methods, indicating that it greatly decreases the noise. This figure also shows that our method can decrease the noise more than the other two methods, especially for dark regions like skull and CSF (figure 13).
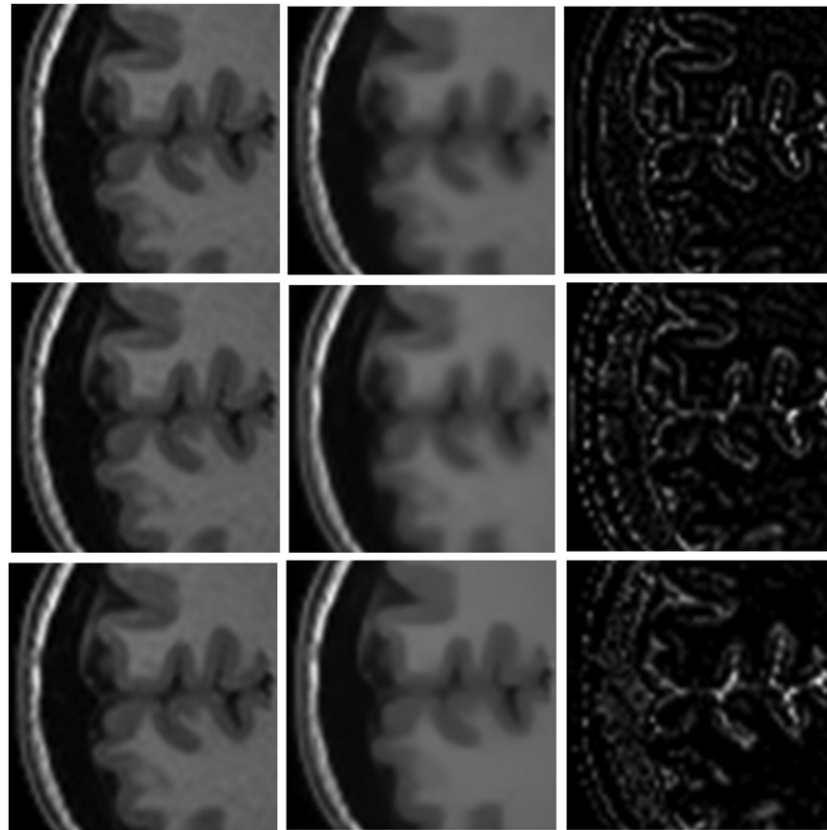
Figure 13. Qualitative detail comparisons of denoising results obtained with different methods. From left to right: original real image, denoised images and the corresponding residuals. From top to bottom: Weiner, wavelet and our method.

5. Discussion and conclusion

We developed and evaluated a wavelet domain denoising method based on the Radon transform for noise removal in MRI. The new approach explicitly accounts for the Rician nature of the MR data. In high-intensity (bright) regions of the MR images, the Rician distribution is well approximated as Gaussian. In low-intensity (dark) regions, the Gaussian approximation is no longer valid and the Rician distribution has two degrading effects: the random fluctuation of pixel values and the introduction of a signal-dependent bias. Based on the Rician noise encountered in MRI we apply the Radon transform to the original MR image, add Rician noise to MR images to consider the Gaussian distribution for MR sinogram images.

Our method combines the Radon transform and wavelet transform and it can be seen as a translation invariant and orthogonal wavelet transform. Based on the fact that the shifted Radon transform along s corresponds to an image significantly different from the original image, we apply a 1D translation invariant wavelet transform along s. Our method is similar to a 2D wavelet transform except that the point parameters are replaced by the line parameters. The wavelet is a function with scale and point position, while our method is a function with scale and line position. Hence the wavelet is effective in representing point singularities. And our method is effective in representing singularities along the line. So our transform can capture singularities along lines and edges in

an efficient way. Based on the nature of noise we show how to obtain accurate thresholds in different scales and evaluate original noise variance.

Our method can not only effectively decrease Rician noise in MRI but also preserve the key image details and features. Brain phantom, a simulation brain MR image and a real brain MR image are used to validate our method. Our method is compared with the optimum linear filter at a single scale, the Wiener filter, and the multiscale traditional wavelet method. Meanwhile, it can enhance the image contrast, and it performs better than the traditional wavelet and Wiener methods in terms of SNR. The experiment results show the superiority of the proposed scheme and outperform the traditional denoising methods. Our denoising method can have wide applications in MR imaging.

Acknowledgments

References

[1] Moseley M E, Liu C, Rodriguez S and Brosnan T 2009 Advances in magnetic resonance neuroimaging *Neurol. Clin.* **27** 271–19

[2] Petersson K M, Nichols T E, Poline J B and Holmes A P 1999 Statistical limitations in functional neuroimaging: I. Non-inferential methods and statistical models *Phil. Trans. R. Soc.* B **354** 1239–60

[3] Gerig G, Kubler O, Kikinis R and Jolesz F A 1992 Nonlinear anisotropic filtering of MRI data *IEEE Trans. Med. Imaging* **11** 221–32

[4] Alexei A S and Chris R J 2004 Noise-adaptive nonlinear diffusion filtering of MR images with spatially varying noise levels *Magn. Reson. Med.* **52** 798–806

[5] Murase K, Yamazaki Y, Shinohara M, Kawakami K, Kikuchi K, Miki H, Mochizuki T and Ikezoe J 2001 An anisotropic diffusion method for denoising dynamic susceptibility contrast-enhanced magnetic resonance images *Phys. Med. Biol.* **46** 2713–23

[6] McGraw T, Vemuri B C, Chen Y, Rao M and Mareci T 2004 DT-MRI denoising and neuronal fiber tracking *Med. Image Anal.* **8** 95–111

[7] Delakis I, Hammad O and Kitney R I 2007 Wavelet-based de-noising algorithm for images acquired with parallel magnetic resonance imaging (MRI) *Phys. Med. Biol.* **52** 3741–51

[8] Weaver J B, Xu Y S, Healy D M and Cromwell L D 1991 Filtering noise from images with wavelet transforms *Magn. Reson. Med.* **21** 288–95

[9] Yang X, Li P, Zhang X, Bian Z and Wang B 2007 De-noising of the Doppler fetal heart rate signal with wavelet threshold filtering based on spatial correlation *1st Int. Conf. IEEE on Bioinformatics and Biomedical Engineering, ICBBE 2007* pp 928–31

[10] Healy D M and Weaver J B 1992 Two applications of wavelet transforms in magnetic resonance imaging *IEEE Trans. Inf. Theory* **38** 840–60

[11] Nowak R D 1999 Wavelet-based Rician noise removal for magnetic resonance imaging *IEEE Trans. Image Process.* **8** 1408–19

[12] Pizurica A, Philips W, Lemahieu I and Acheroy M 2003 A versatile wavelet domain noise filtration technique for medical imaging *IEEE Trans. Med. Imaging* **22** 323–31

[13] Wink A M and Roerdink J B T M 2004 Denoising functional MR images: a comparison of wavelet denoising and Gaussian smoothing *IEEE Trans. Med. Imaging* **23** 374–87

[14] Wirestam R, Bibic A, Latt J, Brockstedt S and Stahlberg F 2006 Denoising of complex MRI data by wavelet-domain filtering: application to high-*b*-value diffusion-weighted imaging *Magn. Reson. Med.* **56** 1114–20

[15] Basu S, Fletcher T and Whitaker R 2006 Rician noise removal in diffusion tensor MRI *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2006* pp 117–25

[16] Awate S P and Whitaker R T 2007 Feature-preserving MRI denoising: a nonparametric empirical Bayes approach *IEEE Trans. Med. Imaging* **26** 1242–55

[17] Manjan J V, Carbonell-Caballero J, Lull J J, Garcia-Marti G, Marti-Bonmati L and Robles M 2008 MRI denoising using Non-Local Means *Med. Image Anal.* **12** 514–23

[18] Jose V M, Neil A Thacker, Juan J Lull and Gracian Garcia-Marti 2009 Multicomponent MR image denoising *Int. J. Biomed. Imaging* **2009** 897–907

[19] Macovski A 1996 Noise in MRI *Magn. Reson. Med.* **36** 494–7

[20] Anand C S and Sahambi J S 2008 MRI denoising using bilateral filter in redundant wavelet domain *TENCON 2008—IEEE Region 10 Conf.* pp 1–6

[21] Gudbjartsson H and Patz S 1995 The Rician distribution of noisy MRI data *Magn. Reson. Med.* **34** 910–4

[22] Sijbers J, Den Dekker A J, Van Audekerke J, Verhoye M and Van Dyck D 1998 Estimation of the noise in magnitude MR images *Magn. Reson. Imaging* **16** 87–90

[23] Jafari-Khouzani K and Soltanian-Zadeh H 2005 Rotation-invariant multiresolution texture analysis using Radon and wavelet transforms *IEEE Trans. Image Process.* **14** 783–95

[24] Ayres C F 2000 A variational approach to MR bias correction *MSc Thesis* University of Minnesota, MN, USA

[25] Mallat S G 1989 A theory for multiresolution signal decomposition—the wavelet representation *IEEE Trans. Pattern Anal. Mach. Intell.* **11** 674–93

[26] Pesquet J C, Krim H and Carfantan H 1996 Time-invariant orthonormal wavelet representations *IEEE Trans. Signal Process.* **44** 1964–70

[27] Jean-Luc S, Candes E J and Donoho D L 2002 The curvelet transform for image denoising *IEEE Trans. Image Process.* **11** 670–84

[28] Do M N and Vetterli M 2003 The finite ridgelet transform for image representation *IEEE Trans. Image Process.* **12** 16–28

[29] Donoho D L 1995 Denoising by soft-thresholding *IEEE Trans. Inf. Theory* **41** 613–27

[30] Yazgan B, Paker S and Kartal M 1992 Image reconstruction with diffraction tomography using different inverse Radon transform algorithms *Proc. 1992 Int. Biomedical Engineering Days* pp 170–3

[31] Nam-Yong L and Lucier B J 2001 Wavelet methods for inverting the Radon transform with noisy data *IEEE Trans. Image Process.* **10** 79–94

[32] Horbelt S, Liebling M and Unser M 2002 Discretization of the Radon transform and of its inverse by spline convolutions *IEEE Trans. Med. Imaging* **21** 363–76

[33] Skiadopoulos S, Karatrantou A, Korfiatis P, Costaridou L, Vassilakos P, Apostolopoulos D and Panayiotakis G 2009 Evaluating image denoising methods in myocardial perfusion single photon emission computed tomography (SPECT) imaging *Meas. Sci. Technol.* **20** 104–23

[34] Kwan R K S, Evans A C and Pike G B 1999 MRI simulation-based evaluation of image-processing and classification methods *IEEE Trans. Med. Imaging* **18** 1085–97

Yang X, Fei BW. A new denoising method for MR images, *Measurement Science and Technology* 2011;22:025803:1-12.