A Wavelet Multiscale Denoising Algorithm for Magnetic Resonance (MR) Images

Xiaofeng Yang 1,2, Baowei Fei 2 *

1 Department of Biomedical Engineering, School of Life Science and Technology,
Xi’an Jiaotong University, Xi’an, Shaanxi China 710049

2 Department of Radiology, Emory University, Atlanta, GA 30329

* Corresponding author:
Baowei Fei, Ph.D., Eng.D.
Center for System Imaging
Department of Radiology and Imaging Sciences
Emory University School of Medicine
1841 Clifton Road NE
Atlanta, GA 30329
Telephone: 404-712-5649
Fax: 404-712-5689
E-mail: bfei@emory.edu

Running Head:
MR image denoising
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 the 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.
1. Introduction

There is a practical limit on the signal-to-noise (SNR) when acquiring MR image data [1]. post-processing methods on how to remove noise is 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 blurring edges due to averaging nonsimilar patterns. In order to avoid this problem, many edge preserving filters have been proposed. One example is Anisotropic Diffusion Filter (ADF) [3-6]. Such filter preserves edges by averaging pixels in the orthogonal direction of the local gradient. However, those denoising methods usually erase small features and change image statistics due to its edge enhancement effect.

Another approach for denoising relies on statistical inference of 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 MR image was squared and the square of Rician random variable was modeled by a scale noncentral chi-square distribution. 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 their 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 to denoising for MRI. The method uses Markov probability density function (PDF) to estimate 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 report, 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, translation invariant wavelet transform is employed to decompose 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 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 mean while 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 $\sigma$, the probability distribution function of the magnitude data will be a Rician distribution, as described by

$$p_{mag}(M) = \frac{M}{\sigma^2} e^{-(M^2 + A^2)/2\sigma^2} I_0(AM / \sigma^2)$$ \hspace{1cm} (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 $A = \sqrt{A_r^2 + A_i^2}$. Eq. (1) is plotted in Fig.1 for different values of the SNR, $A / \sigma$. Notice that the Rician distribution tends to be a Rayleigh distribution.
when the SNR goes to zero and approaches a Gaussian distribution 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 \( \frac{A}{\sigma} = 3 \) (\( SNR = 10 \log_{10} 3^2 \approx 10dB \)), it starts to approximate the Gaussian distribution. For small SNR (\( \frac{A}{\sigma} \leq 1 \)) the Rician distribution is far from a Gaussian distribution. Note that the mean of the distributions, \( \frac{M}{\sigma} \), which is shown by the vertical lines in Fig. 1, is not the same as \( \frac{A}{\sigma} \). This bias is due to the nonlinear transform of the noisy data.

Fig. 1. The Rician distribution of \( M \) for several SNRs (\( A/\sigma \)) and the corresponding means.

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

\[
p_{mag}(M) \approx \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(M^2-(A+\sigma^2))}{2\sigma^2}} \approx \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(M^2-A)}{2\sigma^2}}
\]

(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 \( \sigma^2 \) and mean \( A \). Hence, in high SNR regions the noise can be viewed as a Gaussian white noise with
variation $\sigma^2$ and zero mean. Based on such assumption separation of signal and noise is fairly straightforward in the wavelet domain. However, the Gaussian approximation may introduce error for regions with 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, 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 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:

$$ Rf(\alpha, s) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(s - x \cos \alpha - y \sin \alpha) dx dy \quad \alpha \in [0, \pi] \quad s \in \mathbb{R} \quad (3) $$

Where $s$ is perpendicular distance of a line from the origin; $\alpha$ 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 intensity, i.e., white matter (WM), gray matter (GM) and scalp have high intensity; and the noise in these regions is very close to Gaussian distribution. But other regions like skull, nasal sinuses and cerebrospinal fluid (CSF) have very low intensity [13], so the noise in these regions is close to Rayleigh distribution. We sum these regions along a line and thus sum the noise with different distributions. The sum between several Rician distributed noises has a symmetric distribution. Such sum operation, which is the Radon transform, makes the distribution of noise close to a Gaussian distribution. In Fig. 2, we compare the sum distribution with 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. Fig.2 is the direct comparison between the Gaussian PDF and the sum PDF.
Fig. 2. Gaussian PDF and PDFs of the sum of two or more Rician distributed sets. In (a), (b), (c) and (d) the sums in sequence are: \( \text{Sum}=p_{mag}(1)+10*p_{mag}(2) \), \( \text{Sum}=p_{mag}(2)+10*p_{mag}(3) \), \( \text{Sum}=p_{mag}(3)+10*p_{mag}(4) \), and \( \text{Sum}=p_{mag}(1)+5*p_{mag}(2)+5*p_{mag}(3)+10*p_{mag}(4) \).

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

\[
R(x, y) = \frac{Cov(x, y)}{[\text{StdDev}(x) \cdot \text{StdDev}(y)]}
\]  

(4)

here \( Cov \) is the correlation function and \( \text{StdDev} \) 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 Fig. 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 the larger \( A/\sigma \) is and the more similar to Gaussian the sum PDF is. In order to demonstrate this conclusion further I calculated additional 24 sets of
similarity degree \( R^2 \) between real Gaussian PDF and PDFs of the sum of two or more Rician distributed sets, and it is shown in Table 1. In the table we can see all \( R^2 \)s are over 92\%, and all \( R^2 \)s except \( P_{mag}(0)+5*P_{mag}(1) \) and \( P_{mag}(0)+10*P_{mag}(1) \) are over 95\%. Here \( P_{mag}(0) \) is PDF of the magnitude data in air part inside MR images not background, because we got rid of noise in background using brain mask before the Radon transform. So it is reasonable to use techniques based on the assumption of Gaussian distribution for sinogram images. In fact, for the heart and brain images, the bright regions in body images have SNR of over 20 dB (\( A/\sigma > 10 \)) [24]. Based on two facts above and Eq. (2) we demonstrated that the noise in sinogram images is Gaussian noise with mean \( \mu_i \) and variance \( \sigma_i^2 \). Here from the Radon transform we can get \( \mu_x = n_x \mu \) and \( \sigma_x^2 = n_x \sigma^2 \), and it varies with the number of pixels \( n_x \) along a line through the MR brain image.

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

<table>
<thead>
<tr>
<th>Sum</th>
<th>( R^2 )</th>
<th>Sum</th>
<th>( R^2 )</th>
<th>Sum</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P_{mag}(0)+5*P_{mag}(1) )</td>
<td>0.9232</td>
<td>( P_{mag}(0)+5<em>P_{mag}(1)+10</em>P_{mag}(2) )</td>
<td>0.9509</td>
<td>( P_{mag}(1)+10*P_{mag}(4) )</td>
<td>0.9902</td>
</tr>
<tr>
<td>( P_{mag}(0)+5*P_{mag}(2) )</td>
<td>0.9556</td>
<td>( P_{mag}(0)+5<em>P_{mag}(2)+10</em>P_{mag}(3) )</td>
<td>0.9898</td>
<td>( P_{mag}(1)+5<em>P_{mag}(2)+10</em>P_{mag}(3) )</td>
<td>0.9888</td>
</tr>
<tr>
<td>( P_{mag}(0)+5*P_{mag}(3) )</td>
<td>0.9952</td>
<td>( P_{mag}(0)+5<em>P_{mag}(3)+10</em>P_{mag}(4) )</td>
<td>0.9845</td>
<td>( P_{mag}(1)+5<em>P_{mag}(3)+10</em>P_{mag}(4) )</td>
<td>0.9836</td>
</tr>
<tr>
<td>( P_{mag}(0)+5*P_{mag}(4) )</td>
<td>0.9516</td>
<td>( P_{mag}(0)+5<em>P_{mag}(4)+5</em>P_{mag}(2)+10*P_{mag}(3) )</td>
<td>0.9590</td>
<td>( P_{mag}(2)+5*P_{mag}(3) )</td>
<td>0.9964</td>
</tr>
<tr>
<td>( P_{mag}(0)+10*P_{mag}(1) )</td>
<td>0.9235</td>
<td>( P_{mag}(1)+5*P_{mag}(2) )</td>
<td>0.9541</td>
<td>( P_{mag}(2)+5*P_{mag}(4) )</td>
<td>0.9750</td>
</tr>
<tr>
<td>( P_{mag}(0)+10*P_{mag}(2) )</td>
<td>0.9596</td>
<td>( P_{mag}(1)+5*P_{mag}(3) )</td>
<td>0.9953</td>
<td>( P_{mag}(2)+10*P_{mag}(4) )</td>
<td>0.9956</td>
</tr>
<tr>
<td>( P_{mag}(0)+10*P_{mag}(3) )</td>
<td>0.9971</td>
<td>( P_{mag}(1)+5*P_{mag}(4) )</td>
<td>0.9501</td>
<td>( P_{mag}(2)+5<em>P_{mag}(3)+10</em>P_{mag}(4) )</td>
<td>0.9907</td>
</tr>
<tr>
<td>( P_{mag}(0)+10*P_{mag}(4) )</td>
<td>0.9902</td>
<td>( P_{mag}(1)+10*P_{mag}(3) )</td>
<td>0.9975</td>
<td>( P_{mag}(3)+5*P_{mag}(4) )</td>
<td>0.9993</td>
</tr>
</tbody>
</table>
Fig. 3. Correlation plots between real Gaussian PDF and PDFs of the sum of two or more Rician distributed sets. (a), (b), (c) and (d) are the corresponding correlation plots of (a), (b), (c) and (d) in Fig. 2 respectively. The horizontal axis is real Gaussian distribution data and vertical axis is the distribution sum data.

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 its 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 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 time (frequency) domain. If the mother wavelet is denoted by \( \psi(t) \) other wavelets \( \psi_{a,b}(t) \) can be represented as

\[
\psi_{a,b}(t) = 1/\sqrt{|a|} \ast \psi((t - b)/a)
\]

(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 (DDWT)

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 two). Dyadic wavelet transform is a semi-discrete wavelet transform, makes scale factor binary discrete, while the translation factor to maintain continuous change. The discretized scale and translation parameters are given by \( a = 2^{-j} \) and \( b = k2^{-j} \), where \( j, k \in \mathbb{Z} \), the set of all integers. Thus, the family of wavelet functions is represented as:
\[ \psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \]  

(6)

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

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

(7)

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

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

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

(8)

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

(9)

where \( j_0 \) is the starting scale and is always 0. \( c_{\varphi} \) and \( d_{\psi} \) are the scale factor and the wavelet coefficients, respectively, and \( j \) is the decomposition levels. The development of Fast Wavelet Transform [25] concludes its identity to two channel subband decomposition. Thus, it reveals a remarkable relationship between DWT coefficients of adjacent scales. Using the coefficients \( d_{\psi} \) 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 domain, 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 discrete wavelet transform is due to the required decimation operation (the 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 the translation along $\alpha$ 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 $\alpha$) leads to suboptimal results compared with non-translation invariant wavelet transform. Although the circular shift along $\alpha$ 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 the following section 3.5 we will discuss some properties using 1D wavelet transform along $s$.

3.3.4. Threshold-based denoising in wavelet domain

Wavelet coefficients of signals after wavelet transform contain important information and wavelet coefficients of noise are corresponding 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 $\beta = \mu_s + \sigma_s, \mu_s + 2\sigma_s, \mu_s + 3\sigma_s, \cdots$ will suppress $68.26\%$, $95.44\%$, and $99.74\%$ of its values. Therefore, we choose $\beta = \mu_s + c\sigma_s$, and $c$ is a constant. By imposing $c$ between $3 \sim 4$, we can achieve good results. Based on the fact that the variance in each wavelet scale is also $\sigma_s^2$ in an orthogonal transformation we can get the final threshold

$$\beta = \mu_s + c\sigma_s = n_s\mu + cn_s\sigma$$  \hspace{1cm} (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 $\beta$ from each of the wavelet coefficient, and get the estimated signal. Although Donoho [29] proved the optimality of soft threshold in theory, Stein threshold has shown better results in SNR improvement. Thus

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

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

3.4. Noise Mean $\mu$ and Variance $\sigma^2$ Estimation
Our algorithms require the underlying noise variance $\sigma^2$, 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 Eq. (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 Eq. (1) reduces to

$$p_{mag}(M) = \frac{M}{\sigma^2} e^{-M^2/2\sigma^2}$$

(12)

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

$$M = \frac{\sigma \sqrt{\pi}}{2}$$

(13)

These relations can be used to estimate the "true" noise variance $\sigma^2$, from the magnitude image. So from Eq. (13) the pixel values in the region outside the patient provide us with a very reliable estimator

$$\sigma = \frac{\sqrt{2} M}{\sqrt{\pi}} \quad \text{and} \quad \mu = M$$

(14)

3.5. Properties of our method

1D wavelet transform of an image is equal to the projection of the image onto the 1D basis functions. If we refer to 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:
\[ Wf(\cdot, \cdot) = \int_0^\infty Rf(\alpha, s)\phi(s) ds \]

\[ = \int_0^\infty \left[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(s - x\cos\alpha - y\sin\alpha) dx dy \right] \phi(s) ds \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\delta(s - x\cos\alpha - y\sin\alpha) \phi(s) ds dx dy \]

\[ = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\phi(x\cos\alpha + y\sin\alpha) dx dy \quad (15) \]

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

\[ Wf(\cdot, \cdot) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)h(x, y) dx dy \quad (16) \]

where \( h(x, y) \) in 2D is defined from 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. Wavelet is a function with scale and point position, while \( h(x, y) \) is the function with scale and line position. Hence 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 inverse Radon transform in order to get original image. It is defined as

\[ f(x, y) = \int_{0}^{\pi} R(\alpha, x\cos\alpha + y\sin\alpha) d\alpha \quad (17) \]

where \( R \) are the filtered projections. Generally, three different inverse Radon transform methods are direct inverse Radon transform (DIRT), filtered back-projection (FBP) and
convolution filtered back-projection (CFBP) [30]. DIRT is computationally efficient, but it introduces some artifact. FBP based on 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 image domain is in the dual-spline space. We used SCFBP to propagate the denoised 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 SNR at a single scale. Meanwhile, a traditional multiscale wavelet method is also applied to these dataset 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 average $SNR$ is used as quality metric. It is given by

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

where $x[j,k]$ is original image and $\hat{x}[j,k]$ is the denoised image, and with results averaged over all images and reported as mean decibels ($dB$).
In order to validate our method Rician noise is added into the modified phantom and simulation brain MRI. But when denoising we can get a mask first through a threshold and eliminate the noise outside the phantom and brain. And the noise outside phantom and brain is only used to evaluate the noise variance. The mentioned SNR in following section only express the SNR in phantom or brain except outside air. Here for Wiener filter we used neighborhoods of size $3 \times 3$ to estimate the local image mean and standard deviation. 2D and 1D Db3 wavelet was used in wavelet filter and our method, respectively, and the images were decompounded four levels.

4.1. Brain phantom data

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

Fig. 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 SNR more than that Wiener and wavelet method. At the same time output SNR after Wiener filter appears almost linear growth, while other two methods can improve less with the increasing of input SNR in the original image, and approach the SNR curve of Wiener filter because the noise is closer to Gaussian noise.
Fig. 5. Input and corresponding output SNR plots for the three methods. The horizontal axis is the input SNR and the vertical axis is output SNR.

Fig. 6 exhibits residuals between original and denoised phantom. 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 better for bright regions than dark region, namely, wavelet denoising is better for Gaussian noise than Rician noise in 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 the dark and bright regions, and its denoised effectiveness in dark regions is better than the other two methods.
Fig. 6. Residuals between the original phantom and denoised phantoms. (a) original phantom; (b), (c) and (d) are the corresponding residuals between denoised and original phantom in sequences.

Fig. 7. Comparison of horizontal profiles between original phantom, noised phantom and phantoms denoised using Wiener, Wavelet and our method.

In order to compare the results after different methods quantitatively, we get profiles through the original phantom, the denoised phantom, and the phantom after Wiener filter, the wavelet method and our method as shown in Fig. 7. The Wiener filter can only smooth
denoised image in the whole image, and in the bright region, the wavelet method and our method have almost 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 the original phantom.

In Fig.8 we want to compare the difference of denoised effect in the Radon domain. Noised, denoised and no noise sinograms are showed from left to right in sequence. We can get the denoised sinogram that is close to the sinogram free noise visually and it shows our method has already decreased the noise in sinogram.

Fig.8. Comparison of Noised (left), denoised (middle) and noise-free (right) 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 mathematic simulation of the MR imaging physics [34].

23
As shown in Fig. 9, the whole volume is added the same Rician noise and three slices are showed 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 denoised images, it can be seen that our method enhances the image contract and makes the edges clearer than the other two methods. The Wiener method smoothes the images and makes the whole images blurred. The Wavelet method and our method both preserve image details but our method decrease noise more, and this point can be seen from quantitative SNR. The average SNR of original simulation brain with Ricain noise is 17.88 dB, the average SNR after Wiener is 19.33 dB, the average SNR after wavelet method is 23.12 dB, and the average SNR after our method is 26.84 dB.
Fig. 9. Denoised results of simulated brain data using different methods. The first column is the noised brain images of different slices. The second column is the results after Wiener filter. The third column is the results after Wavelet. The fourth column is the results after our method.

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

Fig. 10. Residuals between original image and denoised images. (a) original image; (b), (c) and (d) are the corresponding residuals between denoised and original image in sequence.
In Fig. 1, we quantitatively compare the profiles through the original image, noised image, and the image after Wiener filter, wavelet and our method. In this figure, the Wiener filter only smooths the noised image in the whole image and loses the image details. In bright regions, the result after wavelet method is almost same to 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 also was applied to real T1-weighted MR images of human brain. The MR images were acquired with a 4.0 Tesla MedSpec MRI scanner (Bruker BioSpin)
GmbH, Rheinstetten, Germany) on a Siemens Syngo platform (Siemens Medical Systems, Erlangen, Germany). 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 $256 \times 256 \times 176$ voxels covering the whole brain yielding 1.0 mm isotropic resolution.

Fig.12 illustrates qualitative comparison of denoised results on T1-weighted MR brain images and the comparison of different methods. We do not do any preprocessing for original MR brain image. In this figure our method makes the edges of MR image clearer and enhances the image contrast. This can be got 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 in the edge than other region. The traditional wavelet can decrease Rician noise better than Wiener, and it can decrease the noise in dark regions in some degree. In brighter regions with high SNR, our result has much less difference than the other two methods, and this means it does not affect the region with little noise. In dark regions that have low SNR, our result has 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 (Fig.13).
Fig. 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.
Fig. 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 down: 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, make Rician noise in MR images to take on the Gaussian distribution for MR sinogram images.

Our method combines the Radon transform and wavelet transform together 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. Wavelet is function with scale and point position, while our method is the function with scale and line position. Hence 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 noise nature we give how to get the accurate thresholds in different scales and evaluate original noise variance.

Our method can not only effectively decrease Rician noise in MRI but also can preserve the key image details and features. Brain phantom, simulation brain MR image and 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 multiscale traditional wavelet method. Meanwhile, it can enhance the image contract, and it performs better than the traditional wavelet and Wiener method in term 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.

Acknowledgment
The authors would like to thank referees and editor for some very helpful comments on the original version of the manuscript.

**Reference**


[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 *The 1St
International Conference of the IEEE on Bioinformatics and Biomedical Engineering,
ICBBE 2007 928-931


