Proc. VIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney

Adaptive Magnetic Resonance Image Denoising Using Mixture Model and Wavelet Shrinkage

Lei Jiang, Wenhui Yang

jiangl@mail.iee.ac.cn
Biomedical Engineering Division, Institute of Electrical Engineering
Chinese Academy of Sciences, Beijing, P.R.China

Abstract. This paper proposes a new adaptive wavelet-based Magnetic Resonance images denoising algorithm. A Rician distribution for background-noise modelling is introduced and a Maximum-Likelihood method for the parameter estimation procedure is used. Further discrimination between edge- and noise-related coefficients is achieved by updating the shrinkage function along consecutive scales and applying spatial constraints. The efficacy of the algorithm is demonstrated on both simulated and real Magnetic Resonance images. The results is shown to be promising and outperform other denoising approaches.

1 Introduction

In Magnetic Resonance Imaging (MRI), there is a trade off between the signal-to-noise ratio (SNR), spatial resolution and acquisition time required by clinical application. The acquired complex dates is known to be corrupted by white noise. The noise contribution to each of real and imaginary parts of k-space date are additive Gaussian and to be uncorrelated, however the corresponding noise in magnitude MR image is Rician, which is signal-depending. Some of the noise originates in the acquisition hardware, others are of physiological origin. These noise artifacts affect the quality and interpretation of medical image date in varying degrees. Therefore, it is important to take advantage of useful date while simultaneously reduce noise artifacts.

Numerous approaches to recover Magnetic Resonance image from its noise have been proposed, which starting from the classic (e.g. Wiener filtering [11]) to the more modern, and usually non-linear, such as wavelets analysis. Application of wavelets for denosing of MR images has been pioneered by Weaver *et al*[1]. They applied their scheme on MR images of the human neck which reduced noise from 10% to 50% without reducing edge sharpness. A recent article by Wood and Johnson[13] employed wavelet packets to treat the real and imaginary parts of MR date separately, but this method would induce the distortions of both phase and amplitude. Some methods combining wavelets analysis with Wiener filtering algorithm have been developed. Alexander[8] developed a wavelet domain Wiener-type denoising algorithm that analyzes the real and

imaginary parts together as a complex signal. Nowak[6] used wavelet domain filtering on the square-magnitude image, and employed a threshold scheme based on a wavelet domain analog of the classical wiener filter.

In the above application of wavelet denoising of MRI date, the denosing was performed on the each component of the complex MR dates separately or together, and the noise was treated as Gaussian. In this paper, we propose a new adaptive wavelet-based Magnetic Resonance images denoising algorithm on the image themselves. Rather, Henkelman R.M [7] has demonstrated that the corresponding noise in magnitude MR images is Rician not Gaussian. This assumption applies more correctly to the magnitude images. A Maximum-likelihood (ML) method for the parameter estimation procedure is used. Further discrimination between edge- and noise-related coefficients is achieved by updating the shrinkage function along consecutive scales and applying spatial constraints. Finally, the efficacy of the algorithm is demonstrated on both simulated and real Magnetic Resonance images.

2 Methods

The most common method for acquiring MR image is to sample the object of interest in the k-space, and the reconstruction technique in MRI is to computer the inverse discrete Fourier transform (DFT) of the raw frequency-domain date:

$$\begin{aligned} y[m,n] = (s[m,n]\cos(\theta[m,n]) + n_{RE}[m,n]) + \\ i(s[m,n]\sin(\theta[m,n]) + n_{IM}[m,n]) \end{aligned} \tag{1}$$

Here s denotes the signal of interest, and n_{RE} and n_{IM} denote Gaussian white noises with variance σ^2 in real and imaginary channels, separately. $\theta[m,n]$ represents the phase error in the m,n-th pixel. From Eq.(1), we can obtain the magnitude of $y[m,n]$ through the square-root of the sum of two independent Gaussian random variables;

$$\begin{aligned} |y[m,n]| = [(s[m,n]\cos(\theta[m,n]) + n_{RE}[m,n])^2 + \\ (s[m,n]\sin(\theta[m,n]) + n_{IM}[m,n])^2]^{\frac{1}{2}} \end{aligned} \tag{2}$$

In this paper, we mainly focus on methods of denoising for the images obtained from magnitude reconstruction.

2.1 Wavelet Shrinkage

In this paper, we perform the image multiresolution decomposition by a redundant wavelet transform, and the 2-D wavelet transform decomposition uses only two detail images (horizontal and vertical details). We first use the term 2-D smoothing function $\theta(x,y)$, whose integral over x and y is equal to 1 and converges to 0 at infinity, to define two wavelet functions given by:

$$\psi^1 = \frac{\partial}{\partial x}\theta(x,y) \qquad \psi^2 = \frac{\partial}{\partial y}\theta(x,y) \tag{3}$$

The wavelet transform of $f(x,y) \in L^2(R^2)$ at the scale 2^j has two component given by:

$$W^i_{2^j} f(x,y) = f(x,y) * \psi^i_{2^j}(x,y) \tag{4}$$

Where $\psi^i_{2^j}(x,y) = (\frac{1}{2^j})^2 \psi^i(\frac{x}{2^j}, \frac{y}{2^j}) \quad j = 1,2$. Thus the multi-resolution wavelet coefficients are as following:

$$\nabla_{2^j} = [W^1_{2^j} f(x,y), W^2_{2^j} f(x,y)] \tag{5}$$

The edge magnitudes can be calculated from the image gradient as follows:

$$M_{2^j} f(x,y) = \sqrt{(W^1_{2^j} f)^2 + (W^2_{2^j} f)^2} \tag{6}$$

The edge orientation is given by the gradient direction, which is expressed by:

$$A_{2^j} f(x,y) = \arctan\left(\frac{W^2_{2^j} f(x,y)}{W^1_{2^j} f(x,y)}\right) \tag{7}$$

Wavelet shrinkage is a well known approach for noise reduction. It estimates the degree about the coefficients contaminated by noise and introduce some measurement, such as probability and degree of membership, to gain the shrinkage proportion. In this paper, we propose an updating shrinkage method based on wavelet multi-scale decomposition and its local adaptivity. For each scale, we aim to find a nonnegative and nondecreasing shrinkage function g_j $(0 \leq g \leq 1)$ in order to update the wavelet coefficients W^1 and W^2 according the following rule:

$$W^i_{2^j} f'(m,n) = W^i_{2^j} f(m,n) * g_j(m,n) \tag{8}$$

Where $i = 1,2$. $g_j(m,n)$ is the shrinkage factor. $g_j(m,n) = g_j[M_{2^j} f(m,n)]$. We will analyze how to obtain this function in the following sections.

2.2 Mixture Model

By analyzing the magnitude image M_{2^j}, we find that some of these coefficients are related to noise and others to edge. We assume that the real and imaginary parts of these complex components are independently Gaussian distribution with the noise variance σ_{noise}. The probability density functions (p.d.f) of the noise-related coefficients is modelled as random variable following a Rician distribution of the squared root of the squared-sum of Gaussian random variables:

$$f_j(i, \sigma_{noise}) = \frac{M_i}{\sigma^2_{noise}} e^{-(\frac{M_i^2 + A^2}{2\sigma^2_{noise}})} I_0(\frac{AM_i}{\sigma^2_{noise}}) u(M_i) \tag{9}$$

Here I_0 is the modified zeroth-order Bessel function of the first kind. M_i denote the ith date point of the magnitude image. The unit step function u is used to

indicate that the expression for the p.d.f of M_i is valid for only nonnegative values. Furthermore, Parameter A is given by $A = \sqrt{A_R^2 + A_I^2}$. From Eq.(9) we can find that in low intensity regions of the magnitude image the noise distribution tends to the Rayleigh distribution, while in regions of high intensity tends to a Gaussian distribution.

In practice, the noise-free images typically consist of homogeneous regions and not many edges. The wavelet coefficients $W_{2^j}^1$ and $W_{2^j}^2$ related exclusively to edge is only a small proportion of the intensity frequency histogram. Hence, for simplicity, we could approximate it by a simple Rayleigh model:

$$f_j(i, \sigma_{edge}) = \frac{M_i}{\sigma_{edge}^2} e^{-(\frac{M_i^2}{2\sigma_{edge}^2})} \tag{10}$$

The noise- and edge-related coefficients can be combined into a mixture model:

$$f_j(i) = w_{noise}^j f_j(i, \sigma_{noise}) + (1 - w_{noise}^j) f_j(i, \sigma_{edge}) \tag{11}$$

where w_{noise}^j is the *a priori* probability for the noise-related gradient magnitude distribution and $(1-w_{noise}^j)$ is the one related to edges. Based on Eq.(9)(10)(11), the shrinkage function $g_j(i)$ for each resolution 2^j is given by the posterior probability function:

$$g_j(i) = f_j(\sigma_{edge}, i) = \frac{(1 - w_{noise}^j) f_j(i, \sigma_{edge})}{f_j(i)} \tag{12}$$

2.3 Parameter Estimation

In all above equations, the unknown parameters are $\sigma_{edge}, \sigma_{noise}, w_{noise}^j$ and A. Conventional estimation methods that are optimal for Gaussian distributed date yield biased results when applied to Rician distributed date. In this paper, we perform a novel estimated technique for optimal estimation of signal as well as noise from Rician distribution based on Maximum Likelihood (ML) estimation.

Initialization can be performed using estimators derived via the method of moments. The second and fourth moment of the Rician variable M are $E[M^2] = A^2 + 2\sigma^2$ and $E[M^4] = A^2 + 8\sigma^2 A^2 + 8\sigma^4$. Solving for A and σ, both of which are non-negative, and replacing the expected values with sample averages yields

$$\sigma_{noise}^2 = \frac{1}{2} \left\{ \frac{1}{n} \sum_{i=1}^{n} M_i^2 - \left(2(\frac{1}{n} \sum_{i=1}^{n} M_i^2)^2 - \frac{1}{n} \sum_{i=1}^{n} M_i^4 \right)^{\frac{1}{4}} \right\} \tag{13}$$

The log-likelihood function of N independent Rician distributed date points is given by

$$\log(L_1) = \sum_{i=1}^{n} \log \frac{M_i}{\sigma_{noise}^2} I_0 \left(\frac{A M_i}{\sigma_{noise}^2} \right) - \frac{N A^2}{2\sigma_{noise}^2} - \sum_{i=1}^{n} \frac{M_i^2}{2\sigma_{noise}^2} \tag{14}$$

$$\log(L_2) = \sum_{i=1}^{n} \log \frac{M_i}{\sigma_{edge}^2} - \frac{NA^2}{2\sigma_{edge}^2} - \sum_{i=1}^{n} \frac{M_i^2}{2\sigma_{edge}^2} \tag{15}$$

Based on the Eq.(14)(15), we can obtain the outcome given by

$$A_{ML} = \arg\left\{\max_{A}(\log L_1)\right\} \tag{16}$$

$$\sigma_{edge}^2 = \arg\left\{\max_{A_{ML},\sigma_{edge}^2}(\log L_2)\right\} \tag{17}$$

$$w_{noise}^j = \arg\left\{\max_{A_{ML},\sigma_{edge}^2,\sigma_{noise}^2}(\log f_j(i))\right\} \tag{18}$$

Here $(0 \leq w_{noise}^j \leq 1)$. Once the parameters $\sigma_{edge}, \sigma_{noise}, w_{noise}^j$ and A are estimated, the shrinkage function $g_j(x)$ can be calculated using Bayes theorem by Eq.(12). In the following section, we will introduce a method to update it.

2.4 Function Updating

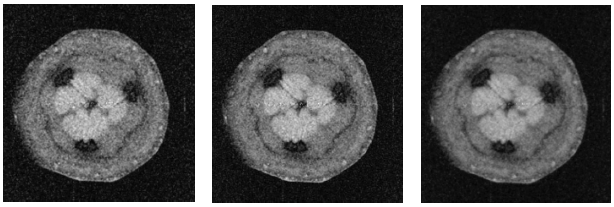
It is known that coefficients associated with noise tend to vanish as the level increases, while coefficients related to edges tend to be preserved. Likewise, if the value is close to unity for several consecutive levels 2^j, it is more likely that $M_{2^j} f$ is associated with an edge, otherwise belong to noise. Further discrimination between edge- and noise-related coefficients is achieved by updating the shrinkage function along consecutive scales and applying spatial constraints. *Xu. etal* [3] used the direct multiplication of the subband decompositions of an image to locate important edges and suppress noise, however he also noticed that involving more than two scales in direct multiplication had a negative impact on the result.

To solve this problem, we compress the dynamic range of the correlation image. Given the relative sizes of the dynamic range of the correlation coefficients between the strong and weak edges we chose to use a logarithmic transform in this prototype

$$I'(x,y) = k * \log(1 + |I(x,y)|) \tag{19}$$

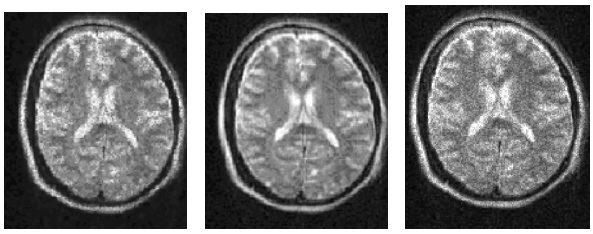
Where $I(x,y)$ is the pixel intensity at location (x,y) of correlation image, and $I'(x,y)$ is the compressed dynamic range correlation image. Then we combine the shrinkage factor $g_j(i)$ in consecutive scales, obtaining the updating shrinkage factor $g'_j(i)$ given by

$$g'_j(i) = \left(\frac{[g_j(i)]^\gamma + [g_{j+1}(i)]^\gamma + \ldots + [g_{j+k}(i)]^\gamma}{K+1}\right)^{\frac{1}{\gamma}} \tag{20}$$



(a) (b) (c)

Fig. 1. (a) Original iamge. (b) Filtering image through our method. (c) Filtered image, using Wiener filtering



(a) (b) (c)

Fig. 2. Effects of various denoising operations are shown on MR images. (a) Original noise image. Filtering operations are performed using (b) our proposed method and (c) the adaptive wiener technique

Where γ is an adjustable parameter, and $(K+1)$ is the number of consecutive scales under consideration. When $\gamma = 1$, the function above is exactly the average of the shrinkage functions. For $\gamma < 1$, smaller shrinkage factor carry more weight and tend to dominate the summation. This process is applied from coarser to finer resolution. The factor $g'_j(i)$ corresponding to the coarsest level 2^j is defined as $g_j(i)$. Other resolutions depend on Eq.(20) defined.

3 Results

Our algorithm was implemented in MATLAB on a PC with Intel Pentium 4 1.7GHz processor and 256M RAM. We applied our techniuqe to image with natural and artificial noise, and compared the result with those obtained by adaptive Wiener denosing method. Furthermore, in this section, we compare two images with low Signal-to-Noise Ratio (SNR) because at high SNRs, Rician noise is well-approximated as Gaussian, which can not show the advantages of our method.

We first apply the quantitative evaluation of performance by defining the signal-to-noise ratio, which is given by:

$$SNR = 10 \log_{10} \left(\frac{S^2_{mean}}{\sigma^2_N} \right) \tag{21}$$

In Fig.(1), we examine the performance of the filtering algorithms with a low SNR MRI date. Fig.(1) shows the original (SNR=9 dB) and filtered images. These parameters are estimated by Section(2.3). Both algorithms reduce the noise without loss of image details, while the SNR of filtered images is about 12.41 dB for our method and 10.26 dB for Wiener filtering.

We also evaluate the visual performance on MR data. A simulated brain Magnetic Resonance Image is show in Fig.(2) with the SNR$\approx$ 6.0 dB, and the denoised images corresponding to the proposed and Wiener filtering are shown, respectively in Fig.2(b) and Fig.2(c). It can be noticed that our algorithm can filter most noise, preserve more details and have a sharpener edges. Wiener algorithm, however, blurs some edges and makes many noise remained.

The above results demonstrate that the proposed algorithm can outperform the Wiener method and obtain more better result with low SNR images because of the noise nature of Rician distribution model.

4 Conclusions

This paper derives a novel wavelet-domain filter that adapt to variations in both the signal and the noise of MR images. Our denosing procedure consists basically of three steps. Initially, the image multiresolution decomposition is performed by a redundant wavelet transform and we obtain the mixture model of noise and edge coefficients. Next, the shrinkage function is calculated through the posterior probability function and all the parameters are Estimated by Maximum Likelihood method. Finally, the shrinkage function is updated by using scale consistency and sequentially the final filtered image is reconstructed by using the modified coefficients. The experimental results obtained are promising.

We acknowledge that the results is preliminary and more research are required in the future. Further work will concentrate on how to adopt more precise coefficient model and more efficient parameter estimation algorithm to reduce operation time and improve the performance.

References

1. Weaver JB, Xu Y, and *et al.*: Filtering noise from images with wavelet transforms. Magn Reson Med. **21**, (1991): 288-295
2. S.G.Mallat, S.Zhong: Characterization of signal from multiscale edges. IEEE Trans. PAMI, **14(7)**, (1992): 710-732.
3. Y.Xu, J.B.Weaver, D.M.Healy: Wavelet transform domain filters:a spatially selective noise filtrateion techniques. IEEE Trans. Image processing, **3(12)**, (1994): 747-758.
4. Cludio, R.Jung, JacobScharcanski: Adaptive image denoising and edge enhancement in scale-space using the wavelet transform. Pattern Recognition Letters, **24**, (2003): 965-971
5. Sijbers, A.J.denDekker, *etal*: Maximum likelihood estimation of Rician distribution parameters. IEEE Trans. Med Imaging, **17(3)**, (1998): 357-361
6. R.D.Nowak: Wavelet-based Rician noise removal for magnetic resonance imaging. IEEE Trans. Image Processing, **8**, (1999): 1408-1419.

7. Henkelman R.M.: Measurement of signal intensities in the presence of noise in MR images. Med. Phys. **12**, (1985): 232- 233; Erratum: Med. Phys. **13**, (1986): 544.
8. M.E.Alexander: A wavelet-based method for improving signal-to-noise ratio and comntrast in MR images. Magnetic Resonance Imaging, **18**, (2000): 168-180.
9. Sleem, Gadi.Goelman: Complex denoising of MR date via wavelet analysis:Application for functional MRI. Magnetic Resonance Imaging, **18**, (2000): 59-68
10. Farshad Faghih, Michael Smith: Combining spatial and scale-space techniques for edge detection to provide a spatially adaptive wavelet-based noise filtering algorithm. IEEE Trans. PAMI, **11(9)**, (2002): 1062-1071
11. Zaroubi S, Hoffman Y, and *et al*: Wiener reconstruction of the large scale structure. Astrophysical Journal, **449**, (1995): 446-459
12. Canny: A computational approach to edge detection. IEEE Trans. PAMI, **8(6)**, (1986): 679-698
13. Wood.J.C, M.K.Johnson Wavelet packet denoising of magnetic resonance images: importance of Rician noise at low SNR. Magnetic Resonance in Medicine, **41(3)**, (1999): 631-635
14. Redpath.T.W: Signal-to-noise ratio in MRI. the British Journal of Radiology, **71**, (1998): 704-707
15. M.C.Veigh, Bronskill.M.J: Noise and filtration in MR imaging. Medical Physics, **18**, (2000): 168-180.
16. E.P.Simonce, lliandE.Adelson: Noise removal via Bayesia wavelet coring. Proc. IEEE International Conference on Image Processing Lausanne Switzerland, September (1996): 279-382.