

Received January 20, 2019, accepted March 12, 2019, date of current version April 17, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2907637

MRI Denoising Using Low Rank Prior and Sparse Gradient Prior

YUHAN ZHANG^{ID1}, ZHIPENG YANG^{ID2}, JINRONG HU¹, SHURONG ZOU¹, AND YING FU¹

¹School of Computer Science, Chengdu University of Information Technology, Chengdu 610225, China

²College of Electronic Engineering, Chengdu University of Information Technology, Chengdu 610225, China

Corresponding author: Ying Fu (fuying@cuit.edu.cn)

This work was supported in part by the Scientific Research Foundation of the Education Department of Sichuan Province under Grant 17ZA0063.

ABSTRACT Image priors have been successfully introduced to solve ill-posed problems, such as image denoising. In this paper, we propose a new denoising model for magnetic resonance images (MRIs) which employs the image low-rank and sparse gradient priors. First, we use a Gaussian mixture model (GMM) to guide the clustering of non-local self-similar patches by learning the structure of external noise-free MRI patches to help retain the low rank of the noisy MRI patch matrix. Second, we fit the heavy-tailed gradient of MRI with a hyper-Laplacian distribution to reduce ringing artifacts. Third, we adopt an alternating iterative algorithm. The experimental results show that our proposed algorithm outperforms many classical MRI denoising methods, such as unbiased nonlocal means (UNLM) filtering and block-matching and 3D (BM3D) filtering in both visual and numerical results.

INDEX TERMS MRI denoising, Rician noise, patch self-similarity, gradient sparsity.

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) is an effective and useful tool for medical diagnosis. However, the noise generated during image acquisition or transmission damages the MRI quality and seriously reducing the accuracy of these diagnoses. The noisy, low-quality MRI image affects the accuracy of automated computer analyses, such as classification, segmentation, and registration. Therefore, research into MRI denoising has become significant for obtaining high-quality MRI output.

Image denoising is an ill-posed problem that is solved by introducing image priors. Image patches contain abundant local structure information, providing sufficient priors for image denoising. Buades *et al.* first proposed the non-local mean (NLM) denoising algorithm [1]. This method used the weighted average value of pixels in one window to estimate a single pixel at one point. This approach made full use of the self-similarity of non-local image patches and preserved image edge detail while removing noise. Manjón *et al.* applied this algorithm successfully to denoise MRI images [2]. However, the high time complexity of NLM led Coupé *et al.* to propose a parallel NLM optimization

The associate editor coordinating the review of this manuscript and approving it for publication was Vishal Srivastava.

algorithm for 3D MRI [3]. Block-Matching and 3D filtering(BM3D) combined the non-local idea with the threshold shrinkage method of the transform domain and achieved better denoising performance with both natural and MRI images than traditional methods [4]–[6].

It is well known that the data matrix formed by image patches with similar structure has a low-rank property, so many researchers have presented denoising methods using the low-rank constraint of the image matrix [7]–[13]. However, denoising using low-rank priors may fail with increased noise levels. To solve the problem, others have proposed principal component dictionary [14] and Gaussian mixture model(GMM) [12] to learn external noise-free image priors to guide clustering of noisy image patches.

The low-rank denoising methods using image patches result in ringing artifacts due to aggregation of overlapped patches. Chang *et al.* proved that the ringing artifacts could be reduced by exploiting the spectral sparsity in multispectral images [15]. Liu *et al.* also found that the MRI image gradient was heavy-tailed and able to be fitted by a Hyper-Laplacian distribution [16]. Motivated by these efforts, we propose a denoising model for MRI that takes advantage of the self-similarity of the nonlocal patches prior and the image gradient sparsity prior. Meanwhile, we use a Gaussian mixture prior to guide clustering of noisy image patches, where the

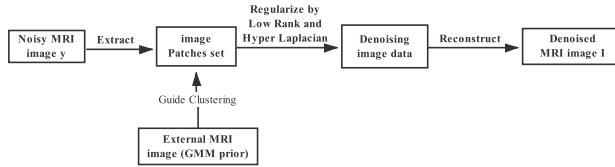


FIGURE 1. Framework of the proposed model.

Gaussian mixture prior is learned from external noise-free MRI patches. The structure of this paper is as follows. Section II describes our proposed MRI image denoising model and optimization. Section III shows the experimental results. Finally, Section IV gives our conclusions and a discussion of them.

II. PROPOSED METHOD

Figure 1 shows the framework of our proposed denoising model. We use a Gaussian mixture prior learned from an external noise-free MR image patch to guide noisy MR image patch clustering with the low-rank matrix well describing the nature of the clustered similar patches. At the same time, we use a hyper-Laplacian distribution to fit the MRI gradient sparsity. Thus, our proposed model preserves the MRI structure during denoising.

A. NONLOCAL SELF-SIMILARITY PRIOR

Generally speaking, image patches with high structural similarity can be clustered. We denote a set of extracted overlapped patches from a given clean MRI image \mathbf{I} as $R\mathbf{I} = (R_1\mathbf{I}, \dots, R_i\mathbf{I}, \dots, R_m\mathbf{I})$, where R_i is the operator that extracts the i -th patch from \mathbf{I} . If the latent structures of the set $R\mathbf{I}$ form K low-dimensional subspaces, the probability that image patch $R_i\mathbf{I}$ can be defined as a weighted sum of K Gaussians is

$$p(R_i\mathbf{I}|\Theta) = \sum_{k=1}^K (w_k p_k(R_i\mathbf{I}|\mu_k, \Sigma_k)) \quad (1)$$

where $\Theta = (\mu_1, \dots, \mu_k, \Sigma_1, \dots, \Sigma_k, w_1, \dots, w_k)$ is the parameter space of GMM. w_k is the k -th weight of the Gaussian distribution, and $\sum_{k=1}^K w_k = 1$. The gaussian density function of the k -th Gaussian distribution in GMM is

$$p_k(R_i\mathbf{I}|\mu_k, \Sigma_k) = c \cdot \exp(-\frac{1}{2}(R_i\mathbf{I} - \mu_k)^T \Sigma_k^{-1} (R_i\mathbf{I} - \mu_k)) \quad (2)$$

where c is the normalized constant and the negative exponent is Mahalanobis distance measuring the similarity between $R_i\mathbf{I}$ and μ_k . μ_k is the mean, and Σ_k is the covariance matrix. To simplify the expression, we use the class label $C = (c_1, \dots, c_m)$, $c_i \in \{1, \dots, K\}$ to denote the Gaussian class to which the image patch $R_i\mathbf{I}$ belongs. $p(R_i\mathbf{I}, c_i = k|\Theta)$ denotes the probability that $R_i\mathbf{I}$ belongs to the k -th Gaussian class under the Θ . For independent image patches, the probability of clustering $R\mathbf{I}$ into K classes is:

$$p(R\mathbf{I}, C|\Theta) = \prod_{i=1}^m p(R_i\mathbf{I}, c_i|\Theta) \quad (3)$$

Taking the logarithm on both sides of Equation (3), we obtain

$$\begin{aligned} \log p(R\mathbf{I}, C|\Theta) &= \sum_{i=1}^m \log p(R_i\mathbf{I}, c_i|\Theta) \\ &= \sum_{i=1}^m \log p(c_i)p(R_i\mathbf{I}|c_i) \end{aligned} \quad (4)$$

Combining with Equation (1), we obtain

$$\sum_{i=1}^m \log p(c_i)p(R_i\mathbf{I}|c_i) = \sum_{i=1}^m \log w_{c_i} p_{c_i}(R_i\mathbf{I}|\mu_{c_i}, \Sigma_{c_i}) \quad (5)$$

Given a noisy MRI image \mathbf{y} , we determine the set of extracted overlapped patches from \mathbf{y} , denoting them as $R\mathbf{y} = (R_1\mathbf{y}, \dots, R_i\mathbf{y}, \dots, R_m\mathbf{y})$. We learn the GMM parameters from the external noise-free MR image patches and then cluster $R\mathbf{y}$ into K classes under the GMM prior. We define $\bar{R}_k\mathbf{y} = [R_{k1}, \dots, R_{kd(k)}]$ to denote the matrix composed of all image patches in the k -th Gaussian class, where $d(k)$ represents the number of similar image patches belonging to the k -th class. Because image patches in the same Gaussian class are similar, $\bar{R}_k\mathbf{y}$ can be represented as:

$$\bar{R}_k\mathbf{y} = Z_k + n_k \quad (6)$$

where Z_k is a low rank matrix corresponding to the denoised MR image patches matrix and n_k is noise matrix. We obtain Z_k by minimizing the energy function $E(Z_k) : E(Z_k) = \tau \|Z_k\|_* + \frac{1}{\sigma^2} \|\bar{R}_k\mathbf{y} - Z_k\|_F^2$, where τ is a positive constant, $\|\bullet\|_*$ is the nuclear norm, and $\|\bullet\|_F$ is the Frobenius norm.

Therefore, given a noisy MRI image \mathbf{y} , we obtain the denoised MRI image \mathbf{I} according to the following preliminary model:

$$\begin{aligned} (\hat{\mathbf{I}}, \hat{C}, \{\hat{Z}_k\}) &= \arg \min_{\mathbf{I}, C, \{Z_k\}} \frac{\lambda}{\sigma^2} \|\mathbf{y} - \mathbf{I}\|_2^2 \\ &\quad - \log p(R\mathbf{I}, C|\Theta) + \sum_{k=1}^K (E(Z_k)) \end{aligned} \quad (7)$$

where $\lambda, \tau (> 0)$ is constant.

Figure 2 shows the clustering result. Figure 2(a) is noisy MRI image(15% Rician noise). Figures 2(b) is the noisy MRI image clustering result. Fig. 2(c), 2(d) and 2(e) are the denoised clustering results in iterations 1, 3 and 6. Figure 2(f) is the noise-free MRI clustering result for reference. It is clear that proposed model(7) exploits GMM prior to guide similar patch clustering, with the clustering results approaching Figure 2(f) in successive iterations.

B. GRADIENT SPARSITY PRIOR

Although the MRI image denoising algorithm using image patches nonlocal self-similarity prior obtains impressive denoising effects, ringing artifacts remain (Figure 3(d) and Figure 3(g)). To solve this problem and preserve more of the MRI image details, we analyze the hyper-Laplacian distribution using the MRI image gradient. As shown in Figure 3(b), the gradient empirical distribution curve (shown in blue)

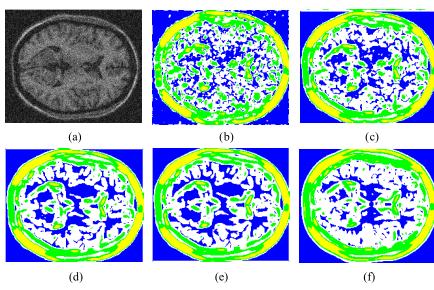


FIGURE 2. The clustering result. (a) Noisy MRI image. (b) The noisy MRI image clustering result. (c) ~ (e) The denoised clustering results in iterations 1,3 and 6. (f) The noise-free MRI clustering result.

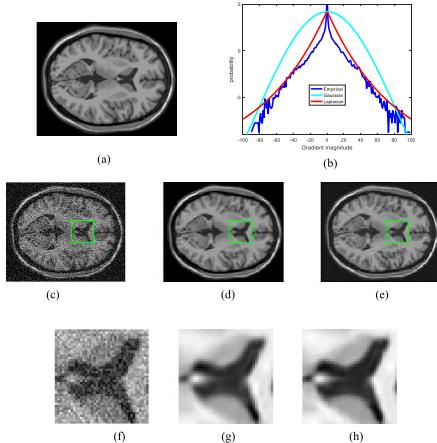


FIGURE 3. Effectiveness of the gradient sparsity prior. (a) Clean MRI image. (b) Fitting empirical distribution on gradient of left clean MRI image. (c) Noise MRI image (9%). (d) The de-noised result of proposed method without gradient sparsity prior (PSNR=24.02dB). (e) The de-noised result of proposed method with Sparsity prior (PSNR=24.71dB). (f)~(h): corresponding enlargement of (c)~(e).

of the noise-free MRI image is sparser than the Gaussian distribution (shown in cyan) and the Laplacian (shown in red) curves. A curve with this characteristic can be well-fitted by a hyper-Laplacian distribution. As shown in Figure 3(e), the denoising result with the hyper-Laplacian prior effectively reduces the ringing artifacts. Therefore, we consider the gradient sparsity prior fitted by a hyper-Laplacian distribution for MRI image denoising.

C. PROPOSED MODEL

For a noisy MRI image \mathbf{y} , the final denoising process for reconvering the latent MRI image \mathbf{I} from \mathbf{y} is

$$\begin{aligned} \{\hat{\mathbf{I}}, \hat{C}, \{\hat{Z}_k\}\} = & \arg \min_{\mathbf{I}, C, \{Z_k\}} \frac{\lambda}{\sigma^2} \|\mathbf{y} - \mathbf{I}\|_2^2 \\ & + \xi(-\log p(R\mathbf{I}, C|\Theta) + \sum_{k=1}^K (E(Z_k))) + \psi \|\nabla \mathbf{I}\|_p \end{aligned} \quad (8)$$

where λ, ξ, ψ are regularization parameters; $\|\mathbf{y} - \mathbf{x}\|_2^2$ is a data fidelity term; and $\log p(R\mathbf{I}, C|\Theta)$ and $\psi \|\nabla \mathbf{I}\|_p$ are the clustering regularization and hyper-Laplacian regularization

terms, respectively. $\sum_{k=1}^K (E(Z_k)) = \sum_{k=1}^K (\tau \|Z_k\|_* + \frac{1}{\sigma^2} \|\bar{R}_k \mathbf{I} - Z_k\|_F^2)$ is the low rank regularization term; ∇ is gradient operator; and $p \in [0, 1]$ is the parameter controlling the sparsity of Hyper Laplacian distribution.

D. OPTIMIZATION

We determine the GMM parameters through training. Other variables including the class label \mathbf{C} , low-rank matrix Z_K , and the denoising MRI image \mathbf{I} are unknown, so we apply an alternating minimization algorithm to solve them.

1) ESTIMATION OF CLASS LABEL \mathbf{C}

In this sub-problem, we fix variable \mathbf{I} and optimize Class label \mathbf{C} with:

$$C = \arg \max_k p(k|R_i \mathbf{I}) = \frac{w_k p_k(R_i \mathbf{I} | \mu_k, \Sigma_k)}{\sum_{j=1}^K w_j p_j(R_j \mathbf{I} | \mu_j, \Sigma_j)} \quad (9)$$

where $i = 1, \dots, m, k = 1, \dots, K$

2) ESTIMATION OF THE LOW RANK MATRIX Z_K

In this subproblem, We fix the variable class label C and perform low rank decomposition on $\bar{R}_k \mathbf{I}$ ($k = 1, 2, \dots, K$), which denotes the matrix formed by the vectorized and stacked image patches in the k -th class.

$$\hat{Z}_k = \arg \min_{Z_k} \tau \|Z_k\|_* + \frac{1}{\sigma^2} \|\bar{R}_k \mathbf{I} - Z_k\|_F^2 \quad (10)$$

The above formula is a typical low-rank minimization problem that we solve using the weighted singular value threshold algorithm [8].

3) RECOVERY OF THE MRI IMAGE \mathbf{I}

To simplify the solution, we separate the variables unrelated to \mathbf{I} in the objective function

$$\hat{\mathbf{I}} = \arg \min_{\mathbf{I}} \frac{1}{\sigma^2} \sum_{k=1}^K \|\bar{R}_k \mathbf{I} - Z_k\|_F^2 + \frac{\lambda}{\sigma^2} \|\mathbf{y} - \mathbf{I}\|_2^2 + \psi \|\nabla \mathbf{I}\|_p \quad (11)$$

To solve this formula, we apply the alternative direction multiplier method (ADMM) [17]by introducing the auxiliary variable H to split the following equation into two simpler sub-problems:

$$\begin{aligned} \{\hat{\mathbf{I}}, \hat{H}\} = & \arg \min_{\mathbf{I}, H} \frac{\xi}{\sigma^2} \sum_{k=1}^K \|\bar{R}_k \mathbf{I} - Z_k\|_F^2 + \frac{\lambda}{\sigma^2} \|\mathbf{y} - \mathbf{I}\|_2^2 \\ & + \psi \|H\|_p + \frac{\alpha}{2} \|H - \nabla \mathbf{I} - \frac{\gamma}{\alpha}\|_F^2 \end{aligned} \quad (12)$$

where H is an auxiliary variable of the same size as \mathbf{I} , γ is a Lagrange multiplier, and ξ and α are positive constants.

1) We first estimate the hyper-Laplacian operator H by separating the variables unrelated to H to obtain

$$\hat{H} = \arg \min_H \psi \|H\|_p + \frac{\alpha}{2} \|H - \nabla \mathbf{I} - \frac{\gamma}{\alpha}\|_F^2 \quad (13)$$

Considering the convergence and accuracy of the algorithm, we apply the generalized iterated shrinkage algorithm [18] to solve the non-convex l_p -norm minimization problem of the preceding equation.

- 2) Similar to the estimation of H , we obtain the following sub-problems by separating variables that are not related to \mathbf{I} :

$$\hat{\mathbf{I}} = \arg \min_{\mathbf{I}} \frac{\xi}{\sigma^2} \sum_{k=1}^K \|\bar{R}_k \mathbf{I} - Z_k\|_F^2 + \frac{\lambda}{\sigma^2} \|\mathbf{y} - \mathbf{I}\|_2^2 + \frac{\alpha}{2} \|H - \nabla \mathbf{I} - \frac{\gamma}{\alpha}\|_F^2 \quad (14)$$

Here, we again use ADMM to separate \mathbf{I} from $\bar{R}_k \mathbf{I}$ by introducing the auxiliary variable M :

$$\begin{aligned} \{\hat{\mathbf{I}}, \hat{M}\} &= \arg \min_{\mathbf{I}, M} \frac{\lambda}{\sigma^2} \|\mathbf{y} - \mathbf{I}\|_2^2 + \frac{\alpha}{2} \|H - \nabla \mathbf{I} - \frac{\gamma}{\alpha}\|_F^2 \\ &\quad + \frac{\xi}{\sigma^2} \sum_{k=1}^K \|\bar{R}_k M - Z_k\|_F^2 + \frac{\beta}{2} \|M - \mathbf{I} - \frac{\gamma_1}{\beta}\|_F^2 \end{aligned} \quad (15)$$

where M is an auxiliary variable of the same size as \mathbf{I} , γ_1 is the Lagrange multiplier, and β is a positive constant. Considering the difference operator, we solve the resulting subproblems using the fast Fourier transform. The approximate iterative solution of the above equation consists of the following:

$$\begin{aligned} \mathbf{I}^{(l+1)} &= \arg \min_{\mathbf{I}} \frac{\lambda}{\sigma^2} \|\mathbf{y} - \mathbf{I}\|_2^2 + \frac{\alpha}{2} \|H - \nabla \mathbf{I} - \frac{\gamma}{\alpha}\|_F^2 \\ &\quad + \frac{\beta}{2} \|M^l - \nabla \mathbf{I} - \frac{\gamma_1^{(l)}}{\beta}\|_F^2 \end{aligned} \quad (16)$$

$$\begin{aligned} M^{(l+1)} &= \arg \min_M \frac{\xi}{\sigma^2} \sum_{k=1}^K \|\bar{R}_k M - Z_k\|_F^2 \\ &\quad + \frac{\beta}{2} \|M - \mathbf{I}^{(l+1)} - \frac{\gamma_1}{\beta}\|_F^2 \end{aligned} \quad (17)$$

$$\gamma_1^{(l+1)} = \gamma_1^{(l+1)} + \beta^l (\mathbf{I}^{(l+1)} - M^{(l+1)}) \quad (18)$$

$$\beta^{(l+1)} = \rho \beta^{(l+1)} \quad (19)$$

where $\rho > 1$. The variables I and M can be approximated with the following equations:

$$\begin{aligned} \mathbf{I}^{(l+1)} &= \mathcal{F}^{-1} \\ &\times \left(\frac{\mathcal{F}(\mathbf{y} + \nabla^T(\alpha^{(l)} H - \gamma) + (\beta^{(l)} M^{(l)} - \gamma_1^{(l)}))}{1 + \alpha^{(l)} (F(\nabla))^2 + \beta^{(l)}} \right) \end{aligned} \quad (20)$$

$$\begin{aligned} M^{(l+1)} &= (2\sigma^2 \sum_{k=1}^K \bar{R}_k^T \bar{R}_k + \beta^{(l)} E) \\ &\times (2\sigma^2 \sum_{k=1}^K \bar{R}_k z_k + \beta^{(l)} \mathbf{I}^{(l+1)} + \gamma_1^{(l)}) \end{aligned} \quad (21)$$

where \mathcal{F} is the 2-D fast Fourier transform, \mathcal{F}^{-1} is the inverse transform, and E is the corresponding unit matrix.

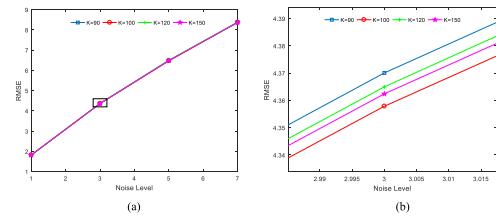


FIGURE 4. The RMSE curve of our proposed method for different values of clustering parameter. (a) The RMSE curve in different Rician noise level. (b) Enlarged curve in 3% Rician noise level of left RMSE curve.

III. EXPERIMENTAL RESULTS

We conducted all experiments using MATLAB running under the Windows 10 Professional operating system on an Intel Core i3-6100 CPU at 3.70 GHz with 8 GB of memory. According to experimental experience, we set the parameters λ , ξ , ψ and p in the proposed model are fixed to 0.19, 1, 0.05 and 0.8, respectively, and the patch size to 10. To validate the performance of our proposed model, we compared it with three classical MRI image denoising methods: unbiased nonlocal means (UNLM) [2], Anisotropic Diffusion Filter (ADF) [19], and BM3D [20]. BM3D designed for Gaussian noise can handle Rician noise by using a variance stabilization transformation (VST) [21]. Source code for the NULM and BM3D-VST methods were obtained from their respective authors' websites with ADF parameters taken from the recommendations of Mohan *et al.* [22] to achieve the best performance.

A. CLUSTER NUMBER K

The main purpose of learning the noise-free image patches is to obtain the mean μ , covariance matrix Σ , and weight w of each Gaussian class through the EM algorithm [23]. However, in addition to these parameters, the clustering number for the GMM is also vital for the denoising performance. Figure 4 shows the root mean square error (RMSE) curve of our proposed method using the Brainweb dataset (Simulated Brain Database, SBD) [24] with different values of K 90, 100, 120, and 150. Figure 4 shows that the lowest RMSE was obtained with $K = 100$ indicating that $K = 100$ achieved the best denoising performance. Thus, we set K to 100. We learned the GMM parameters with 100 mixture components from a set of 6.9×10^6 patches sampled from the 50 clean MR images from SBD. This process required around 20 hours with Chen's MATLAB code from <http://www.mathworks.com/matlabcentral/fileexchange/26184-em-algorithm-for-gaussian-mixture-model>.

B. SYNTHETIC MRI IMAGE

We conducted experiments with synthetic MR data (T1w, PDw, and T2w) with 1 mm³ voxels, obtained from the SBD. Their sizes are 181*217*181 and the 90th slice of them was used for experiment. The denoising results from our proposed model and the comparison algorithms are shown in Figures 5 through 8 and Tables 1 through 3.

TABLE 1. PSNR and SSIM results for different methods on the T1w image with different Rician noise levels.

Method	3%		7%		11%		15%	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Noise	30.121	0.7364	22.711	0.5244	18.731	0.3786	16.053	0.2820
ADF	29.708	0.8993	23.131	0.5475	19.117	0.3899	16.303	0.2876
UNLM	37.627	0.9510	31.924	0.8756	28.529	0.7937	26.005	0.7122
BM3D-VST	36.758	0.9783	32.079	0.9318	29.552	0.8826	27.704	0.8215
OUR	37.849	0.9796	32.506	0.9393	29.246	0.8928	26.746	0.8396

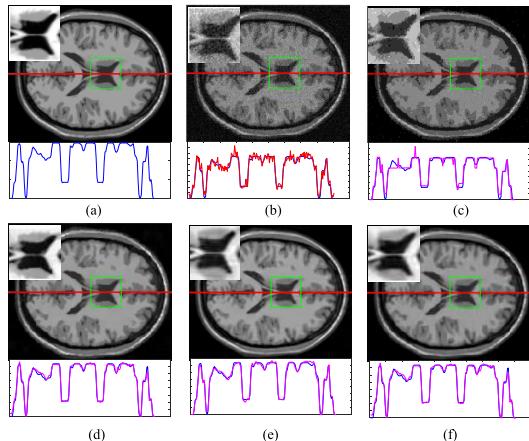
**FIGURE 5.** Denoising results with 5% Rician noise and associated 1D intensity value curves of different de-noising methods. (a) The clean MRI image. (b) The noisy MRI image. (c)~(f) The denoising results of ADF, UNLM, BM3D-VST, and our method respectively.

Figure 5 shows the denoising results (top) on a T1 weighted image and their associated one-dimensional intensity value curve (bottom) under the 5% Rician noise Level. Figure 5(a) is the clean MRI image for reference. Figure 5(b) is the noisy MRI image. Figures 5(c) through 5(f) show the denoising results of different algorithms. Figure 5(c) shows little residual noise in the ADF denoising result. UNLM, BM3D-VST, and our method effectively removed noise and achieved good visual denoising performance. We reach the same conclusion from the 1D pixel value curve shown at the bottom of Figure 5. Figure 5(c) shows a significant difference between the 1D curve (magenta) obtained from ADF and the clean MR image (blue). It is clear that intensity values from BM3D-VST and our proposed algorithm are more structurally similar to the original MRI image. However, the 1D curve from our algorithm is closer to that of the noise-free MRI image in Figures 5(e) and 5(f).

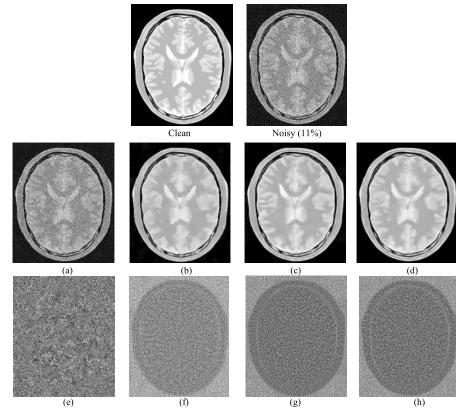
**FIGURE 6.** Qualitative comparison of different methods on the PDw image with 11% Rician noise. ((a)~(d): denoising result of different filters. (e)~(f): residuals of different filters. From left to right: ADF, UNLM, BM3D-VST and Our method).

Figure 6 presents a qualitative comparison of the different methods on a PD weighted image with 11% Rician noise. Figures 6(a) through 6(d) show the denoising results from ADF, UNLM, BM3D-VST, and our algorithm, and Figures 6(e) through 6(h) show the corresponding residual images (i.e., the difference between the denoised and noisy images). Figure 6(e) shows that ADF retains some structural details in the residual images. Both BM3D-VST and our proposed algorithm showed better performance than other algorithms. However, our algorithm yielded higher PSNR and SSIM values than the other methods with an 11% noise level as shown in Table 2.

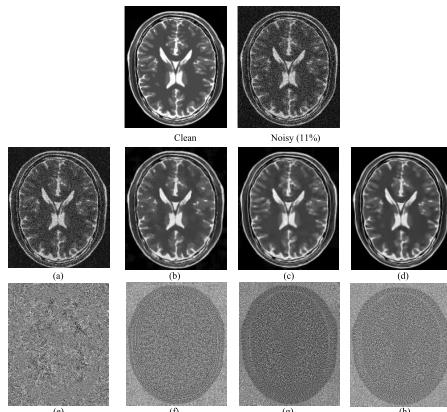
To demonstrate the robustness of proposed algorithm, we conducted an experiment on T2 weighted images with 11% Rician noise level. The qualitative denoising results of the different filters is shown in Figure 7. Compared to other denoised results, our algorithm produced a residual

TABLE 2. PSNR and SSIM results for different methods on the PDw image with different Rician noise levels.

Method	3%		7%		11%		15%	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Noise	29.884	0.7358	22.361	0.4956	18.389	0.3725	15.653	0.2911
ADF	30.068	0.8678	23.054	0.5312	18.856	0.3845	15.920	0.2969
UNLM	36.706	0.9572	31.345	0.8854	28.235	0.7991	26.272	0.7428
BM3D-VST	36.063	0.9522	31.013	0.9134	28.463	0.8658	26.681	0.8279
OUR	37.163	0.9520	32.096	0.9112	29.594	0.8806	27.448	0.8390

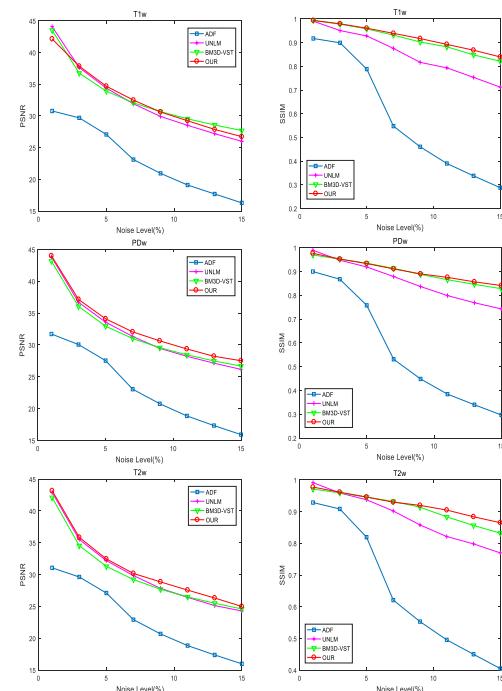
TABLE 3. PSNR and SSIM results for different methods on the T2w image with different Rician noise levels.

Method	3%		7%		11%		15%	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Noise	29.878	0.7722	22.382	0.5893	18.451	0.4828	15.760	0.4002
ADF	29.630	0.9086	22.930	0.6211	18.850	0.4955	15.989	0.4051
UNLM	35.518	0.9589	29.846	0.9021	26.502	0.8215	24.290	0.7700
BM3D-VST	34.555	0.9592	29.236	0.9322	27.504	0.8836	24.608	0.8322
OUR	35.860	0.9614	30.185	0.9299	27.566	0.9050	24.997	0.8652

**FIGURE 7.** Qualitative comparison of different methods on the T2w image with 11% Rician noise. ((a)~(d)): denoising result of different filters. (e)~(f): residuals of different filters. From left to right: ADF, UNLM, BM3D-VST and Our method).

(between denoised and clean images) that was less correlated. Table 3 provides the quantitative results associated with the different methods in Figure 7. Compared with other MRI image denoising methods, our algorithm better preserved the MRI image details and obtained the highest PSNR and SSIM scores on the T2 weighted images.

To make a quantitative assessment of the effectiveness of our algorithm, we show the PSNR and SSIM curves for the T1w, PDw, and T2w image under different Rician noise levels in Figure 8. Our proposed algorithm almost obtained the highest PSNR and SSIM among the denoising methods on the three different MR image types and achieved better denoising performance than the others. Tables 1 through 3 show more detailed quantitative results. The performance of our algorithm was significantly superior to ADF, UNLM, and BM3D-VST methods in terms of both PSNR and SSIM. For the T1w image, our proposed algorithm yielded better performance than all the other methods when noise levels were under 7%. When the noise level increased, the PSNR from BM3D-VST was slightly better than our algorithm, but only subtly so. Our algorithm still obtained the higher SSIM. For the PDw image, our proposed method outperformed BM3D-VST with PSNR improvements of 0.767 (15%) and 1.131dB (11%) and corresponding SSIM improvements of 0.0111 and 0.0148. Our method outperformed UNLM with PSNR improvements of 1.359dB (11%) and 1.176dB (15%) and corresponding SSIM improvements of 0.0815 and 0.0962. Compared to BM3D-VST, our proposed method on the T2w image improved the PSNR ranging from 0.389dB to 1.305dB and

**FIGURE 8.** Denoising results (PSNR/SSIM) on different image types and noise conditions.

the corresponding SSIM from 0.0022 to 0.033. Compared to UNLM, the improvement ranged from 0.342dB to 1.064dB and 0.0025 to 0.0952 in PSNR and SSIM, respectively.

C. CLINICAL MRI DATA

In this section, we further demonstrate the validity of our proposed algorithm on clinical MRI images from the IXI dataset (<http://brain-development.org/>) by comparing results with UNLM and BM3D-VST. Clinical MRI images have noise, so we first adopt the method proposed by Rajan *et al.* [25] to estimate the noise standard deviation. Figures 9 and 10 show the results. Qualitatively, UNLM, BM3D-VST, and our proposed algorithm effectively removed Rician noise from the real MRI image. From the corresponding enlarged parts of the rectangular regions and residual images, however, our algorithm produced a less correlated residual and better visual results (Figures 9 and 10).

D. CONVERGENCE AND RUNNING TIME

Because the denoising model we have proposed is an NP-hard problem, we now discuss the convergence of the model.

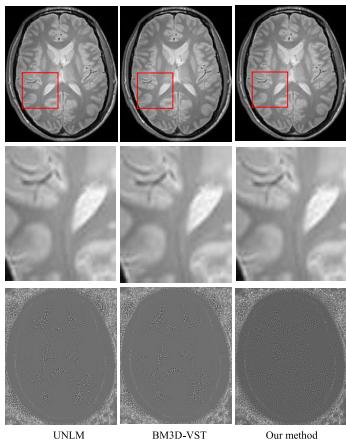


FIGURE 9. Qualitative comparison of different methods on a real PD weighted image (estimated $\sigma = 5$). From top to bottom: denoised results, enlarged images and the corresponding residuals. From left to right: UNLM, BM3D-VST, and our method.

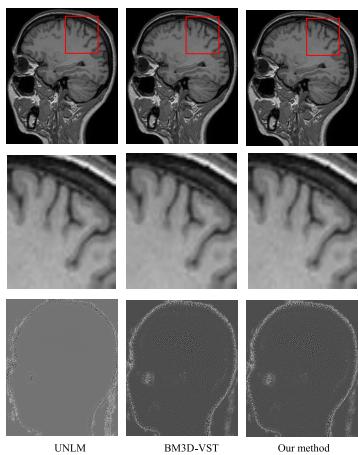


FIGURE 10. Qualitative comparison of different methods on a real T1 weighted image (estimated $\sigma = 1$). From top to bottom: denoised results, enlarged images, and the corresponding residuals. From left to right: UNLM, BM3D-VST, and our algorithm.

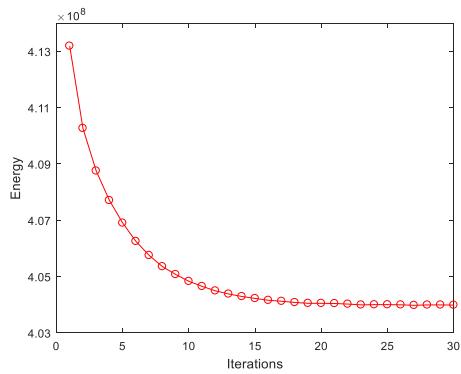


FIGURE 11. Convergence curve of the proposed method on SBD under 15% Rician noise.

The convergence curve of our proposed algorithm with a Rician noise level of 15% on SBD is shown in Figure 11. Our algorithm converges at about 25 iterations. We also record the denoising time of each method. In our experimental environment, the approximate running times of the

different method to denoise a 217×181 PDw image with 11% Rician noise level were: ADF 43.26 seconds; UNLM 4.32 seconds; BM3D-VST 0.37 seconds; and our method 7.48 seconds.

IV. CONCLUSION AND DISCUSSION

In this paper, we have proposed a new MRI image denoising algorithm using low-rank and sparse gradient priors. We train the GMM with noise-free MR image patches and use it to guide noisy MR patch clustering, helping to retain the low-rank of the image matrix and enhancing robustness to noise. We use a hyper-Laplacian distribution to fit the gradient sparse prior of MRI image to reduce ring artifacts caused by the aggregation of overlapped patches. The experimental results show that our MRI image denoising algorithm effectively suppresses Rician noise and retains MRI image details.

As shown in Section III.A, the clustering parameter is an empirical value, not an adaptive one. This shows that our proposed model suffers a limitation in the form of high time cost in training the GMM. In recent years, graphics processing units (GPUs) have been used to accelerate computing processes related to medical imaging [26], [27]. Therefore, we plan to develop a GPU parallel framework to accelerate the training process.

REFERENCES

- [1] A. Buades, B. Coll, and J.-M. Morel, “A non-local algorithm for image denoising,” in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Jun. 2005, pp. 60–65.
- [2] J. V. Manjón *et al.*, “MRI denoising using non-local means,” *Med. Image Anal.*, vol. 12, no. 4, pp. 514–523, Aug. 2008.
- [3] P. Coupé, P. Yger, and C. Barillot, “Fast non local means denoising for 3D MR images,” in *Proc. Int. Conf. Med. Image Comput. Comput-Assisted Intervent.*, pp. 33–40.
- [4] E. M. Eksioglu, “Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI,” *J. Math. Imag. Vis.*, vol. 56, no. 3, pp. 430–440, Nov. 2016.
- [5] X. B. Lin and T. S. Qiu, “Denoise MRI images using sparse 3D transformation domain collaborative filtering,” in *Proc. Int. Conf. Biomed. Eng. Informat.*, Oct. 2011, pp. 233–236.
- [6] P. Elahi, S. Beheshti, and M. Hashemi, “BM3D mri denoising equipped with noise invalidation technique,” in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, May 2014, pp. 6612–6616.
- [7] W. Dong, G. Shi, and X. Li, “Nonlocal image restoration with bilateral variance estimation: A low-rank approach,” *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 700–711, Feb. 2013.
- [8] S. Gu, L. Zhang, W. Zuo, and X. Feng, “Weighted nuclear norm minimization with application to image denoising,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2862–2869.
- [9] A. Rajwade, A. Rangarajan, and A. Banerjee, “Image denoising using the higher order singular value decomposition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 4, pp. 849–862, Apr. 2013.
- [10] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, “Compressive sensing via nonlocal low-rank regularization,” *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3618–3632, Aug. 2014.
- [11] D. M. Lyra-Leite, J. P. C. L. da Costa, and J. L. A. de Carvalho, “Improved MRI reconstruction and denoising using SVD-based low-rank approximation,” in *Workshop Eng. Appl.*, May 2012, pp. 1–6.
- [12] F. Chen, L. Zhang, and H. Yu, “External patch prior guided internal clustering for image denoising,” in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 603–611.
- [13] Y. Zhang, J. Liu, S. Yang, and Z. Guo, “Joint image denoising using self-similarity based low-rank approximations,” in *Proc. Vis. Communun. Image Process.*, Nov. 2014, pp. 1–6.

- [14] S. Yao, Y. Chang, X. Qin, Y. Zhang, and T. Zhang, "Principal component dictionary-based patch grouping for image denoising," *J. Vis. Commun. Image Represent.*, vol. 50, pp. 111–122, Jan. 2017.
- [15] Y. Chang, L. Yan, and S. Zhong, "Hyper-laplacian regularized unidirectional low-rank tensor recovery for multispectral image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 5901–5909.
- [16] R. W. Liu, L. Shi, W. Huang, J. Xu, S. C. Yu, and D. Wang, "Generalized total variation-based MRI Rician denoising model with spatially adaptive regularization parameters," *Magn. Reson. Imag.*, vol. 32, no. 6, pp. 702–720, Jul. 2014.
- [17] Z. Lin, R. Liu, and Z. Su, "Linearized alternating direction method with adaptive penalty for low-rank representation," in *Proc. Adv. Neural Inf. Process. Syst.*, pp. 612–620, 2011.
- [18] W. Zuo, D. Meng, L. Zhang, and X. Feng, "A generalized iterated shrinkage algorithm for non-convex sparse coding," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 217–224.
- [19] G. Gerig, O. Kubler, R. Kikinis, and F. A. Jolesz, "Nonlinear anisotropic filtering of MRI data," *IEEE Trans. Med. Imag.*, vol. 11, no. 2, pp. 221–232, Jun. 1992.
- [20] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [21] A. Foi, "Noise estimation and removal in MR imaging: The variance-stabilization approach," in *Proc. IEEE Int. Symp. Biomed. Imag., Nano Macro*, Mar./Apr. 2011, pp. 1809–1814.
- [22] J. Mohan, V. Krishnaveni, and Y. Guo, "MRI denoising using nonlocal neutrosophic set approach of Wiener filtering," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 779–791, Nov. 2013.
- [23] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc. Ser. B Methodological*, vol. 39, no. 1, pp. 1–22, Sep. 1977.
- [24] A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans, "BrainWeb: Online interface to a 3D MRI simulated brain database," *NeuroImage*, vol. 5, no. 4, p. S425, 1997.
- [25] J. Rajan, D. Poot, J. Juntu, and J. Sijbers, "Noise measurement from magnitude MRI using local estimates of variance and skewness," *Phys. Med. Biol.*, vol. 55, no. 16, pp. 441–449, 2010.
- [26] L. Shi, W. Liu, H. Zhang, Y. Xie, and D. Wang, "A survey of GPU-based medical image computing techniques," *Quant. Imag. Med. Surg.*, vol. 2, no. 3, pp. 188–206, Sep. 2012.
- [27] H. Wang, H. Peng, Y. Chang, and D. Liang, "A survey of GPU-based acceleration techniques in MRI reconstructions," *Quant. Imag. Med. Surg.*, vol. 8, no. 2, pp. 196–208, Mar. 2018.

Authors' photographs and biographies not available at the time of publication.

• • •