Denoising Magnetic Resonance Images Using Collaborative Non-Local Means

Geng Chen[a,b], **Pei Zhang**[b], **Yafeng Wu**[a], **Dinggang Shen**[b], and **Pew-Thian Yap**[b,*]

[a]Data Processing Center, Northwestern Polytechnical University, Xi'an, China

[b]Department of Radiology and Biomedical Research Imaging Center (BRIC) University of North Carolina at Chapel Hill, NC, U.S.A

Abstract

Noise artifacts in magnetic resonance (MR) images increase the complexity of image processing workflows and decrease the reliability of inferences drawn from the images. It is thus often desirable to remove such artifacts beforehand for more robust and effective quantitative analysis. It is important to preserve the integrity of relevant image information while removing noise in MR images. A variety of approaches have been developed for this purpose, and the non-local means (NLM) filter has been shown to be able to achieve state-of-the-art denoising performance. For effective denoising, NLM relies heavily on the existence of repeating structural patterns, which however might not always be present within a single image. This is especially true when one considers the fact that the human brain is complex and contains a lot of unique structures. In this paper we propose to leverage the repeating structures from *multiple* images to *collaboratively* denoise an image. The underlying assumption is that it is more likely to find repeating structures from multiple scans than from a single scan. Specifically, to denoise a target image, multiple images, which may be acquired from different subjects, are spatially aligned to the target image, and an NLM-like block matching is performed on these aligned images with the target image as the reference. This will significantly increase the number of matching structures and thus boost the denoising performance. Experiments on both synthetic and real data show that the proposed approach, collaborative non-local means (CNLM), outperforms the classic NLM and yields results with markedly improved structural details.

Keywords

Denoising; Non-Local Means; Block Matching; Non-Parametric Regression

1. Introduction

Due to thermal noise, magnetic resonance (MR) images are susceptible to noise artifacts resulting from random fluctuation of the MR signal. Such artifacts cause uncertainty in

[*]Corresponding author: ptyap@med.unc.edu (Pew-Thian Yap).

signal measurements and unreliability in quantitative analyses performed using these images. It is thus critical to denoise these images to improve the robustness and reliability of subsequent analysis.

There are in general two kinds of approaches to noise removal in images. One is the hardware approach [1], which involves scanning an object of interest multiple times and averaging the resulting signals to increase signal-to-noise ratio (SNR). This approach is not always practical due to the long acquisition time. The other is the software approach [2, 3, 4, 5, 6, 7, 8, 9], which uses computer algorithms to extract the true signals from noisy measurements. In this work, we focus on the second approach because it can be applied to existing data without requiring expensive equipment upgrades.

Among the large number of algorithms developed for noise removal, a frequently used approach is to attempt to recover the true intensity value of a voxel by averaging the intensity values of neighboring voxels [10]. A popular example is the Gaussian smoothing filter. However, this kind of local averaging technique will remove not only noise but also structural details such as anatomical boundaries. The loss of such details is undesirable due to their potential clinical diagnostic value, such as in characterizing small pathological changes in the brain. To deal with this issue, patch-based approaches have been shown to obtain considerable improvements. Especially notable patch-based methods are the non-local means (NLM) algorithm [11] and the block matching and 3D filtering (BM3D) algorithm [12]. Instead of relying on voxels that are spatially close to each other, the NLM filter averages across (potentially distant) voxels that capture similar structures and thus avoids blurring structural details. The assumption is that real images often have many self-similar structures that are not necessarily spatially close to each other and that these repeating patterns may be used for effective noise removal. Similar to the NLM filter, the BM3D filter utilizes redundant information distributed throughout the whole image for effective denoising. It arranges similar patches into groups and then carries out denoising by shrinkage of the transform coefficients of the group of patches. We mainly focus on the NLM filter in this work because of its simplicity.

Although the NLM filter has been successfully applied to MR image denoising [2, 3, 4], it fails when self-repeating structures cannot be located. A natural solution to this problem is to extend the spatial search range to increase the chance of finding similar structures. However, this will increase computation time dramatically and there is no guarantee of success in finding similar structures. In fact, small weights given to a large number of dissimilar structures will often overwhelm the weights of a few true matching structures. To increase the number of matching structures, Prima and Commowick [3] proposed to capitalize on the bilateral symmetry of the human brain to double the chance of finding matching structures. This is achieved by using information from both ipsilateral and contralateral hemispheres. Despite the promising results, this approach only increases moderately the chance of finding matching structures because only information from one single image is used.

An alternative approach is to borrow information across multiple images for denoising [13]. For example, VBM3D [14] utilizes redundant information found within a frame as well as

other frames to carry out efficient video denoising. Foi [15] combines VBM3D and a variance-stabilization approach for multi-image denoising [16]. Note however that all of the above methods require repeatedly acquired images of the same object and are hence not applicable to MRI denoising. Repeated acquisition in MRI increases scan times and is hence clinically prohibitive.

To solve the above problem we propose to harness repeating structures from MR images of different individuals to boost image denoising performance. This is a generalization of the classic NLM filter. The underlying idea is that although human brains differ from each other, they all have many common structures that may be used for effective denoising. To increase the probability of finding matching structures, we spatially align a group of images, called co-denoising images, to a target noisy image and use them to improve denoising. NLM-like block matching is performed to locate matching blocks, not only in the target image itself, but also in the co-denoising images, significantly increasing the number of matching blocks. Such technique has been applied in multi-atlas segmentation [17], but its application in image denoising has not been investigated. Extensive experiments on both simulated and real datasets show that the proposed approach, called collaborative non-local means (CNLM), yields results with markedly improved structural details when compared with the classic NLM filter.

The rest of the paper is organized as follows. In Section 2, we will describe the proposed method. In Section 3, we will then describe the datasets used for evaluation. In Section 4, we will demonstrate the effectiveness of the proposed algorithm for both synthetic and real data. In Section 5, we will provide additional discussion and conclude the paper.

2. Method

2.1. Non-Local Means Filter

We first introduce the classic NLM filter. Let $\mathrm{NL}(u)(x_i)$ be a restored value of a given voxel at location $x_i \in \mathbb{R}^3$. It can be computed as the weighted average of all voxels within a search volume $V(x_i)$, i.e.

$$\mathrm{NL}(u)(x_i) = \sum_{x_j \in V(x_i)} w(x_i, x_j) u(x_j),$$

where $V(x_i)$ is a cubic volume centered at x_i, $u(x_j)$ is the intensity value of the voxel at x_j and $w(x_i, x_j)$ is the weight. The size of $V(x_i)$ is $(2M + 1)^3$, where M is a search radius. For structural matching, we define a smaller local cubic neighborhood $N(x_i)$ around x_i. The size of $N(x_i)$ is $(2d + 1)^3$, where d is a neighborhood radius. Let $u(N(x_i))$ be a vector which represents the intensity values of all voxels within $N(x_i)$, then $w(x_i, x_j)$ may be defined as a Gaussian function of the Euclidean distance between vectors $u(N(x_i))$ and $u(N(x_j))$ by

$$w(x_i, x_j) = \frac{1}{Z_i} \exp \left\{ -\frac{\|u(N(x_i)) - u(N(x_j))\|_2^2}{h_i^2} \right\}, \quad (1)$$

where h_i controls the attenuation of the exponential function and Z_i is a normalization constant to ensure that $w(x_i, x_j)$ sums up to one.

If h_i is too large, all voxels tend to have a same weight, leading to a strong smoothing effect. If h_i is too small, only a few very similar voxels will be involved in denoising, and the difference between the denoised image and the original image will be subtle. Coupé et al. [2] suggested to set $h_i = \sqrt{2\beta \hat{\sigma}_i^2 |N(x_i)|}$, where $\hat{\sigma_i}$ is an estimate of the standard deviation of the noise at voxel x_i, β is a constant and is set to 1 [2] and $|N(x_i)|$ is the size of $N(x_i)$. The weight $w(x_i, x_j)$ is required to satisfy $0 \quad w(x_i, x_j) \quad 1$ and $\Sigma_{x_i \in V_{(x_i)}} w(x_i, x_j) = 1$. If x_i and x_j are the same, the weight is too large. Hence, $w(x_i, x_i)$ is set according to $w(x_i, x_i) = \max(w(x_i, x_j))$, $\forall i \quad j$.

2.2. Collaborative NLM Denoising

NLM relies on recurring image information. But when the number of similar structures is small, particularly in regions that contain a corner or an edge, one encounters the *rare patch effect* [18, 19, 20]. This phenomenon leads to degradation of fine details and often manifests as halos around object boundaries. In the following, we will reformulate NLM to work with images scanned from different subjects to overcome the problem of insufficient structural recurrence. Unlike [2, 5, 4, 3, 21], which are focused on locating similar structures within an image, our approach, called collaborative non-local means (CNLM), will allow leveraging of common structures in different scans to improve denoising performance.

Suppose we have a target noisy image and a group of co-denoising images with indices denoted as set S. Let $V_k(x_i)$ be the search volume centered at x_i in image $k \in S$ and $\hat{w}_k(x_i, x_j)$ be an unnormalized weight, then the CNLM compute the restored value of the voxel at x_i as

$$\mathrm{NL}(u)(x_i) = \frac{\sum_{k \in S} \sum_{x_j \in V_k(x_i)} \hat{w}_k(x_i, x_j) u(x_j)}{\sum_{k \in S} \sum_{x_j \in V_k(x_i)} \hat{w}_k(x_i, x_j)}.$$

If we let

$$\mathrm{NL}_k(u)(x_i) = \frac{\sum_{x_j \in V_k(x_i)} \hat{w}_k(x_i, x_j) u(x_j)}{\sum_{x_j \in V_k(x_i)} \hat{w}_k(x_i, x_j)},$$

and

$$Z_{i_k} = \sum_{x_j \in V_k(x_i)} \hat{w}_k(x_i, x_j),$$

then we have

$$\mathrm{NL}(u)(x_i) = \frac{\sum_{k \in S} Z_{i_k} \mathrm{NL}_k(u)(x_i)}{\sum_{k \in S} Z_{i_k}}.$$

Hence, the restored value given by CNLM is just a weighted average of the restored values given by NLM across all images. An overview of CNLM is given in Fig. 1. Let $|S|$ be the cardinality of the set S, then every voxel will have $|S|$ search volumes, including one in the target noisy image itself and $|S| - 1$ in the co-denoising images. The sample size of the CNLM is thus $|S|$ times larger than the classic NLM and hence similar blocks can be found with greater probability for improving denoising.

The NLM filter can be viewed as non-parametric kernel regression in an image block space [22, 8]. Based on the theory of kernel regression, it can be proven that increasing sample size alone will not reduce estimation bias; the bandwidth of the kernel, which in our case is the Gaussian function defined in (1), has to be decreased accordingly [23]. Hence, following the proof given in [23], we set $h_i = \sqrt{2\beta \left(\hat{\sigma}_i |S|^{-1/5} \right)^2 |N(x_i)|}$, where h_i now depends on the number of images $|S|$ used for denoising[1]. This makes intuitive sense, because, as the number of co-denoising images increases, more matching structures become available, but at the same time more spurious structures are introduced. Reducing the bandwidth as the number of images increases will help ensure that only truly matching structures are used for denoising.

2.3. Block Preselection

A huge number of weight calculations between blocks need to be performed in CNLM, causing considerable computational burden. Decreasing unnecessary weight calculations is essential for denoising in a feasible amount of time. Mahmoudi et al. [25] proposed filters that eliminate unrelated blocks from the weighted average, reducing the original quadratic complexity involved in NLM to a linear one. The basic idea is to preclassify image blocks according to characteristics such as their average gray values and gradient orientation and only perform weight computation for blocks with similar characteristics. This method has been proven to not only decrease significantly the computational burden, but also enhance denoising efficacy. Manjón et al. [4] introduced an improved version of this method and applied it to denoise MR images. The improved preselection method is based on the mean values, the variance values and the inverted mean values of the query patch N_i and the search patch N_j. The involvement of the inverted mean value $\mathrm{inv}(u(N_i)) = \max(u) - u(N_i)$ in

---

[1] See [24] for a multivariate treatment on how the bandwidth should be set. Here, we have simplified the problem by leveraging the fact that image voxels are highly correlated.

the filtering process is to avoid treating high and low intensities differently [4]. To discard unrelated blocks, we set

$$w(x_i, x_j) = \frac{1}{Z_i} \exp\left\{-\frac{\|u(N(x_i)) - u(N(x_j))\|_2^2}{h_i^2}\right\}, \quad (2)$$

only if

$$\left(\mu_1 < \frac{\overline{u(N_i)}}{\overline{u(N_j)}} < \frac{1}{\mu_1} \ \lor \ \mu_1 < \frac{\text{inv}(\overline{u(N_i)})}{\text{inv}(\overline{u(N_j)})} < \frac{1}{\mu_1}\right) \land \sigma_1^2 < \frac{\text{var}(u(N_i))}{\text{var}(u(N_j))} < \frac{1}{\sigma_1^2}, \quad (3)$$

else we set $w(x_i, x_j) = 0$. Here, notation $\bar{\cdot}$ and var($\cdot$) denote respectively the mean and the variance of the respective neighborhood blocks. The parameters $0 < \mu_1 < 1$ and $0 < \sigma_1 < 1$ were chosen according to [2, 4].

2.4. Adaptation to Rician Noise

The noise in the MR magnitude signal is Rician-distributed [26]. The classic NLM algorithm is designed to remove Gaussian noise and needs to be modified to tackle Rician noise. A Rician-distributed variable X satisfies $E(X^2) = \mu^2 + \sigma^2$, where μ is the true value and σ is a scale parameter that determines the level of noise. Using this fact, it is suggested in [21] that an unbiased estimate of the intensity can be obtained as

$$\text{NL}(u)(x_i) = \sqrt{\sum_{x_j \in V(x_i)} w(x_i, x_j) u^2(x_j) - 2\hat{\sigma}_i^2}.$$

A similar adjustment is adapted for CNLM to deal with Rician noise, i.e.,

$$\text{NL}(u)(x_i) = \sqrt{\frac{\sum_{k \in S} \sum_{x_j \in V_k(x_i)} \hat{w}_k(x_i, x_j) u^2(x_j)}{\sum_{k \in S} \sum_{x_j \in V_k(x_i)} \hat{w}_k(x_i, x_j)} - 2\hat{\sigma}_i^2}.$$

3. Datasets

3.1. Rician Noise Simulation

In order to evaluate quantitatively the proposed method, the Simulated Brain Database of BrainWeb[2] was used. We used a noise-free T_1-weighted image from the database, simulated using a spoiled FLASH sequence with repetition time (TR) = 18 ms, echo time (TE) = 10 ms, and flip angle = 30°. The image size is $181 \times 217 \times 181$ and the slice thickness is 1 mm.

---

[2]http://brainweb.bic.mni.mcgill.ca/brainweb/

To evaluate our algorithm we added noise to the image. The noise was assumed to follow a Rician distribution [26]

$$p(m) = \frac{m}{\sigma^2}\exp\left(-\frac{m^2+A^2}{2\sigma^2}\right)I_0\left(\frac{Am}{\sigma^2}\right), \quad (4)$$

where σ is a scale parameter that is equivalent to the standard deviation of a Gaussian distribution when the SNR is high, A is the noise free signal, m is noisy signal actually observed, I_0 is the zeroth order modified Bessel functions of the first kind. When signal to noise ratio (SNR) is low, i.e. $A/\sigma \to 0$, the noise follows a Rayleigh distribution

$$p(m) = \frac{m}{\sigma^2}\exp\left(-\frac{m^2}{2\sigma^2}\right). \quad (5)$$

When SNR is high, i.e. $A/\sigma \to \infty$, the noise obeys a Gaussian distribution

$$p(m) \approx \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\left(m-\sqrt{A^2+\sigma^2}\right)^2}{2\sigma^2}\right). \quad (6)$$

To generate Rician noise we first created two intermediate images I_r and I_i using the following two equations

$$I_r(x_i) = I_0(x_i) + \eta_1(x_i), \quad \eta_1(x_i) \sim \mathcal{N}(0,\sigma), \quad (7)$$

and

$$I_i(x_i) = \eta_2(x_i), \quad \eta_2(x_i) \sim \mathcal{N}(0,\sigma), \quad (8)$$

where I_0 is the noise-free image. We then computed the noisy image I_N as

$$I_N(x_i) = \sqrt{I_r(x_i)^2 + I_i(x_i)^2}. \quad (9)$$

In this paper, the noise standard deviation is specified in terms of percentage. That is, $P\%$ noise implies $\sigma = v(P/100)$, where v is the brightest intensity value in the image (255 in our case).

3.2. Dataset 1: Baseline Synthetic Data

As the CNLM algorithm is capable of using multiple images for denoising, 11 noisy replicates of the T_1-weighted image were generated for each noise level (i.e., 3%, 5%, 7%, and 9%). One image was used as the target image for denoising and ten others as co-denoising images for providing additional information for denoising. Evaluation on this dataset provides a reference set of results on how the different denoising algorithms perform when there is no structural differences between images and when the images are perfectly

aligned. This reference dataset is also useful to validate that the flexibility afforded by the non-local block matching mechanism used in CNLM does not falsely deteriorate the results significantly when there is in fact no structural variation.

3.3. Dataset 2: Transformed Synthetic Data

To evaluate the effects of structural variation on denoising, we applied 10 rigid transformations to the noise-free co-denoising images mentioned above. The transformations include translations ([−2mm, 2 mm]) and rotations ([−2°, 2°]) along each axis. Noise was added to the 10 transformed co-denoising images and the target image. This dataset hence consists of 4 groups of images; each group has 11 images perturbed with the same level of noise.

3.4. Dataset 3: Real Data

This dataset consists of 11 T_1-weighted MR images acquired from different individuals using a Siemens 3T TIM Trio MR scanner with a common imaging protocol. One image was used as the target image and the rest were used as co-denoising images. The images are of size $256 \times 256 \times 160$ with isotropic 1mm resolution. All co-denoising images were warped to the target space using a large deformation diffeomorphic registration algorithm [27, 28].

4. Results

In all experiments, we set $\beta = 1$, $M = 2$, $d = 2$, $\mu_1 = 0.95$ and $\sigma_1^2 = 0.5$. $|S| = 11$, including one target noisy image and ten co-denoising images. $\hat{\sigma_i}$ was estimated based on a cubic volume (radius = 2) centered at x_i, similar to the methods described in [2, 29].

4.1. Baseline Synthetic Dataset

The peak signal to noise ratio (PSNR) was used to evaluate denoising accuracy. For 8-bit encoded images, the PSNR is defined as

$$PSNR = 20 \log_{10} \frac{255}{RMSE},$$

where RMSE is the root mean square error computed between the ground truth and the denoised image.

We ran both the NLM and the CNLM algorithms to denoise the synthetic data and then computed the PSNR values of the resulting images. The results are shown in Fig. 2. Perhaps unsurprisingly, we can see that our method outperforms the classic NLM at all noise levels. For moderate noise (i.e. 3%), the PSNR improvement given by CNLM over NLM peaks at 4.65 dB. Moreover, the performance of our method can be increasingly improved by using more co-denoising images (Fig. 3). Such improvement is more significant when the target image is less noisy because it is more challenging to identify matching structures if there is too much noise.

Simple averaging is used here as the comparison baseline. In this case, there is no structural variation and simple averaging is expected to produce optimal results. From Fig. 3, we can observe that CNLM degrades the results slightly due to its flexibility in including more information during denoising and hence introducing the possibility of including spurious information during the averaging process. However, the benefits of CNLM outweighs its imperfection especially when true structural variations exist. When structural variations occur, the performance of simple averaging degrades dramatically. This is due to its inability to correct for structural misalignment, unlike CNLM.

To better visualize the benefits of CNLM over NLM, we computed the absolute difference between the noise-free image and the denoised image given by each method. We then computed for each voxel a performance contrast value by calculating the difference of the absolute difference values given CNLM and NLM: Contrast = Absolute Difference$_{NLM}$ − Absolute Difference$_{CNLM}$. If the absolute difference value given by CNLM is lower than NLM, then the contrast value at the voxel is positive; otherwise it is negative. Repeating this process for each noise level leads to the images shown in Fig. 4. We can see that voxels with warm colors (positive values) dominate the whole brain for all noise levels, indicating that our method works significantly better than NLM.

For a more fine-grained analysis, we report the results for different regions of the brain, determined based on the Anatomical Automatic Labeling (AAL) template [30]. The AAL template was warped to the space of the target image using a large deformation diffeomorphic registration algorithm [27, 28]. Fig. 5 shows, for each region, the number of the voxels in which CNLM or NLM performs better than one another in terms of PSNR. The number at each bar indicates the ratio between the length of the CNLM (red) bar to that of the NLM (blue) bar. It can be seen that, in every brain region, CNLM yields better performance than NLM. The number of voxels where CNLM performs better is approximately twice the number of voxels where NLM performs better. The maximum ratio 2.96 occurs for the left rectus region.

To show that CNLM is indeed averaging over similar structures, we computed the sum of the top 10 un-normalized weights associated with each voxel. Greater weight sums indicate smaller structural differences between blocks that are deemed similar. The results, displayed in Fig. 6, confirm that CNLM is averaging over structures that are more similar than NLM. This is indicative that CNLM is able to preserve edges better than NLM, which is confirmed in our evaluation using real data, as reported in Section 4.3.

4.2. Transformed Synthetic Dataset

Fig. 2 shows that CNLM performs quite comparably on both baseline and transformed synthetic datasets. This implies that CNLM is insensitive to structural variations. For the baseline synthetic dataset where there is no structural misalignment, simple averaging yields optimal results. However, when structural misalignment occurs, the outcome of simple averaging degrades significantly. The block matching mechanism employed in CNLM helps to offset the effect of structural misalignment. This hence enables CNLM to borrow information from scans of different individuals for more effective denoising. Compared with NLM, which does not borrow information across individuals, CNLM yields higher

denoising performance, as confirmed by the results shown in Figs. 7 8, 9, and 10. The conclusions that can be drawn from these figures are very similar to their corresponding figures (Figs. 3, 4, 5, and 6) discussed in the previous section.

4.3. Real Data

Representative results for the real data consisting of scans from different individuals are shown in Fig. 11. It might not be immediately apparent from the images that CNLM gives better performance than NLM. But close-up views of some cortical structures (see Fig. 12), which are typically small and difficult to denoise without sacrificing structural details, indicate that CNLM is more effective in preserving edges. Preservation of the fine details in cortical regions is critical for applications such as cortical surface extraction [31] and cortical thickness measurement [32]. The sum-of-weights analysis (Fig. 13) again confirms that CNLM, compared with NLM, benefits from being able to leverage information from structures that have greater similarity. The mean values of the sums of weights are reported for the different AAL regions in Fig. 14. In some regions such as the left and right part of the pallidum, the mean values given by CNLM are more than 4 times greater than NLM. This is indicative of the fact that even though the images are acquired for different individuals, they share common structural information that can be harnessed for improving denoising performance.

4.4. Computational Time

Our implementation takes advantage of the multithreading capability of ITK. The algorithm took approximately 45 seconds to process one imaging slice on a machine with an Intel Core i5 processor (3.1GHz).

5. Discussion and Conclusions

We have demonstrated that information common across images of different individuals can be utilized for effective structure-preserving denoising. This is achieved by extending NLM to not only employ *self-similar* but also *mutually-similar* information occurring in other images for increasing the sample size for improving the estimation of the noise-free value of each voxel. This helps avoid the *rare patch effect*, discussed in [18, 19, 20], which affects the classic NLM when there is a lack of pattern reoccurrence.

Our work is greatly inspired by the label fusion work of [17]. The goal of our work however is not to transfer label information but to harness T_1-weighted images from different individuals for effective denoising. Furthermore, in our framework, we deal with varying bandwidth associated with the change in sample size. It is well known in the theory of nonparametric regression [23] that estimation bias does not improve with sample size if the bandwidth is not adjusted. In [17], the bandwidth is at most spatially adaptive and is invariant to sample size. Our work suggests that the statistical bias in [17] can be improved by better bandwidth selection.

Experiments on synthetic data have shown that the proposed method works significantly better than the classical NLM algorithm. Ideally, one should use a set of noise-free images acquired from different subjects in these experiments. However, to the best of our

knowledge, no existing database provides such images. The BrainWeb image (Section 3.2) that we have used for our evaluation is the only noise-free image with sufficiently structural complexity for realistic evaluation. Various levels of Rician noise was added to the image so that we could study the affect of noise on denoising performance. This however only satisfies partly the requirements of an ideal synthetic evaluation dataset because all the noisy image realizations are perfectly aligned. Hence, we further introduced spatial misalignment, via random transformations, to the synthetic dataset to create a more challenging dataset for evaluation. In the future, as more anatomical models become available, we will subject our method to more stringent evaluations.

Both qualitative and quantitative results of our evaluations using the real data support the notion that images from different individuals contain common structural information that can be used for mutual denoising. The results given by the proposed method clearly showed less structural blurring and greater detail preservation. The effectiveness of the proposed method can be attributed to the fact that estimation accuracy is improved by increasing the number of samples. A well-behaving estimator should asymptotically converge to the real value as the number of samples increases to infinity. The proposed method significantly increases the number of available samples by taking advantage of information redundancy between images.

In conclusion, we have proposed an MRI denoising framework that makes use of inter-subject structural correlations for effective estimation of the noiseless signal. This obviates the need for time-consuming multiple acquisitions. Future work will be directed to extend the proposed collaborative denoising framework to work with diffusion-weighted images [33]. This will help improve the reliability of studies investigating the integrity white matter tracts in relation to development, aging, and disorders [34, 35, 36, 37, 38, 39, 40, 41].

Acknowledgments

References

1. Gerig G, Kubler O, Kikinis R, Jolesz FA. Nonlinear anisotropic filtering of MRI data. IEEE Transactions on Medical Imaging. 1992; 11(2):221–232. [PubMed: 18218376]

2. Coupé P, Yger P, Prima S, Hellier P, Kervrann C, Barillot C. An optimized blockwise nonlocal means denoising filter for 3-D magnetic resonance images. IEEE Transactions on Medical Imaging. 2008; 27(4):425–441. [PubMed: 18390341]

3. Prima, S.; Commowick, O. Using bilateral symmetry to improve non-local means denoising of MR brain images. International Symposium on Biomedical Imaging (ISBI); IEEE; 2013. p. 1231-1234.

4. Manjón JV, Coupé P, Mart -Bonmat L, Collins DL, Robles M. Adaptive non-local means denoising of MR images with spatially varying noise levels. Journal of Magnetic Resonance Imaging. 2010; 31(1):192–203. [PubMed: 20027588]

5. Awate S, Whitaker R. Unsupervised, information-theoretic, adaptive image filtering for image restoration. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2006; 28(3):364–376. [PubMed: 16526423]

6. Wood JC, Johnson KM. Wavelet packet denoising of magnetic resonance images: Importance of Rician noise at low SNR. Magnetic Resonance in Medicine. 1999; 41(3):631–635. [PubMed: 10204890]

7. Chen, G.; Zhang, P.; Wu, Y.; Shen, D.; Yap, P-T. Collaborative non-local means denoising of magnetic resonance images. International Symposium on Biomedical Imaging (ISBI); IEEE; 2015. p. 564-567.

8. Yap P-T, An H, Chen Y, Shen D. Uncertainty estimation in diffusion MRI using the nonlocal bootstrap. IEEE Transactions on Medical Imaging. 2014; 33(8):1627–1640. [PubMed: 24801775]

9. Fang R, Zhang S, Chen T, Sanelli P. Robust low-dose CT perfusion deconvolution via tensor total-variation regularization. IEEE Transactions on Medical Imaging. 2015; 34(7):1533–1548.

10. McDonnell M. Box-filtering techniques. Computer Graphics and Image Processing. 1981; 17(1): 65–70.

11. Buades A, Coll B, Morel J-M. A review of image denoising algorithms, with a new one. Multiscale Modeling & Simulation. 2005; 4(2):490–530.

12. Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Transactions on Image Processing. 2007; 16(8):2080–2095. [PubMed: 17688213]

13. Buades, T.; Lou, Y.; Morel, JM.; Tang, Z. A note on multi-image denoising. International Workshop on Local and Non-Local Approximation in Image Processing (LNLA); 2009. p. 1-15.

14. Dabov, K.; Foi, A.; Egiazarian, K. Video denoising by sparse 3D transform-domain collaborative filtering. Proc. 15th European Signal Processing Conference; 2007. p. 7

15. Foi, A. Noise estimation and removal in MR imaging: The variance-stabilization approach. International Symposium on Biomedical Imaging (ISBI); 2011. p. 1809-1814.

16. Boracchi, G.; Foi, A. Multiframe raw-data denoising based on block-matching and 3-D filtering for low-light imaging and stabilization. International Workshop on Local and Non-Local Approximation in Image Processing (LNLA); 2008.

17. Asman AJ, Landman BA. Non-local statistical label fusion for multi-atlas segmentation. Medical image analysis. 2013; 17(2):194–208. [PubMed: 23265798]

18. Duval V, Aujol J-F, Gousseau Y. A bias-variance approach for the nonlocal means. SIAM Journal on Imaging Sciences. 2011; 4(2):760–788.

19. Deledalle C-A, Duval V, Salmon J. Non-local methods with shape-adaptive patches (NLM-SAP). Journal of Mathematical Imaging and Vision. 2012; 43(2):103–120.

20. Salmon J, Strozecki Y. Patch reprojections for non-local methods. Signal Processing. 2012; 92(2): 477–489.

21. Wiest-Daesslé N, Prima S, Coupé P, Morrissey SP, Barillot C. Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI: Applications to DT-MRI. Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2008; 5242:171–179. [PubMed: 18982603]

22. Takeda H, Farsiu S, Milanfar P. Kernel regression for image processing and reconstruction. IEEE Transactions on Image Processing. 2007; 16(2):349–366. [PubMed: 17269630]

23. Silverman, B. Monographs on Statistics and Applied Probablilty. Chapman and Hall; 1998. Density Estimation for Statistics and Data Analysis.

24. Scott, DW. Multivariate Density Estimation. 2. Wiley; 2015.

25. Mahmoudi M, Sapiro G. Fast image and video denoising via nonlocal means of similar neighborhoods. IEEE Signal Processing Letters. 2005; 12(12):839–842.

26. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magnetic Resonance in Medicine. 1995; 34(6):910–914. [PubMed: 8598820]

27. Zhang P, Niethammer M, Shen D, Yap P-T. Large deformation diffeomorphic registration of diffusion-weighted images with explicit orientation optimization. Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2013; 8150:27–34. [PubMed: 24579120]

28. Zhang P, Niethammer M, Shen D, Yap P-T. Large deformation diffeomorphic registration of diffusion-weighted imaging data. Medical Image Analysis. 2014; 18(8):1290–1298. [PubMed: 25106710]

29. Commowick, O.; Stamm, A. Non-local robust detection of DTI white matter differences with small databases. In: Ayache, N.; Delingette, H.; Golland, P.; Mori, K., editors. Medical Image Computing and Computer-Assisted Intervention (MICCAI). Vol. 7512. 2012. p. 476-484.

30. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage. 2002; 15(1):273–289. [PubMed: 11771995]

31. Eskildsen SF, Østergaard LR. Quantitative comparison of two cortical surface extraction methods using MRI phantoms. Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2007; 4791:409–416. [PubMed: 18051085]

32. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proceedings of the National Academy of Sciences. 2000; 97(20):11050–11055.

33. Yap, P.-T.; Shen, D. DWI denoising using spatial, angular, and radiometric filtering. MICCAI Workshop on Multimodal Brain Image Analysis (MBIA), Vol. 7509 of Lecture Notes in Computer Science; 2012. p. 194-202.

34. Yap P-T, Wu G, Shen D. Human brain connectomics: Networks, techniques, and applications. IEEE Signal Processing Magazine. 2010; 27(4):131–134.

35. Yap P-T, Fan Y, Chen Y, Gilmore J, Lin W, Shen D. Development trends of white matter connectivity in the first years of life. PLoS ONE. 2011; 6(9):e24678. [PubMed: 21966364]

36. Wee C-Y, Yap P-T, Li W, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D. Enriched white matter connectivity networks for accurate identification of MCI patients. NeuroImage. 2011; 54(3):1812–1822. [PubMed: 20970508]

37. Wee C-Y, Yap P-T, Zhang D, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D. Identification of MCI individuals using structural and functional connectivity networks. NeuroImage. 2012; 59(3):2045–2056. [PubMed: 22019883]

38. Shi F, Yap P-T, Gao W, Lin W, Gilmore JH, Shen D. Altered structural connectivity in neonates at genetic risk for schizophrenia: A combined study using morphological and white-matter networks. NeuroImage. 2012; 62(3):1622–1633. [PubMed: 22613620]

39. Jin Y, Shi Y, Zhan L, Gutman BA, de Zubicaray GI, McMahon KL, Wright MJ, Toga AW, Thompson PM. Automatic clustering of white matter fibers in brain diffusion MRI with an application to genetics. NeuroImage. 2014; 100:75–90. [PubMed: 24821529]

40. Dennis EL, Jin Y, Villalon-Reina JE, Zhan L, Kernan CL, Babikian T, Mink RB, Babbitt CJ, Johnson JL, Giza CC, Thompson PM, Asarnow RF. White matter disruption in moderate/severe pediatric traumatic brain injury: Advanced tract-based analyses. NeuroImage: Clinical. 2015; 7:493–505. [PubMed: 25737958]

41. Jin, Y.; Shi, Y.; Zhan, L.; Thompson, PM. Automated multi-atlas labeling of the fornix and its integrity in Alzheimer's disease. International Symposium on Biomedical Imaging (ISBI); IEEE; 2015. p. 140-143.
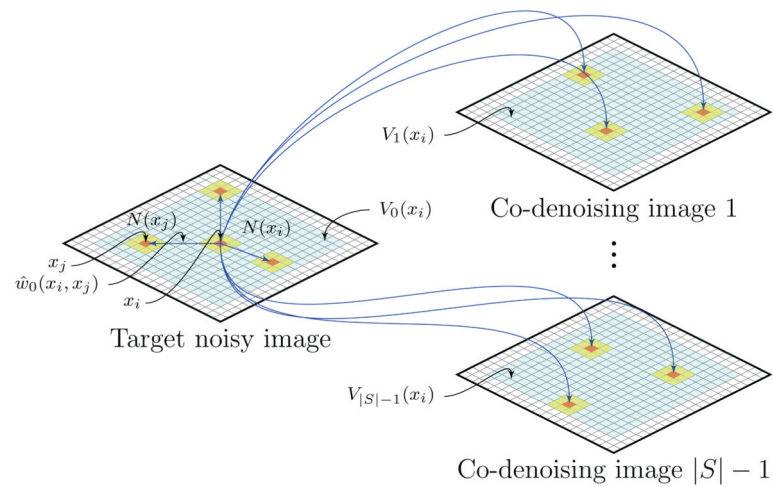
Figure 1.
An overview of CNLM. The value of the voxel at location x_i is computed based on the weighted average of voxels within the associated search volumes $V_k(x_i)$ (cyan squares) in the target image and the co-denoising images. Each weight $\hat{w}_k(x_i, x_j)$ is determined based on the similarly between voxel neighborhoods $N(x_i)$ and $N(x_j)$ (yellow squares).
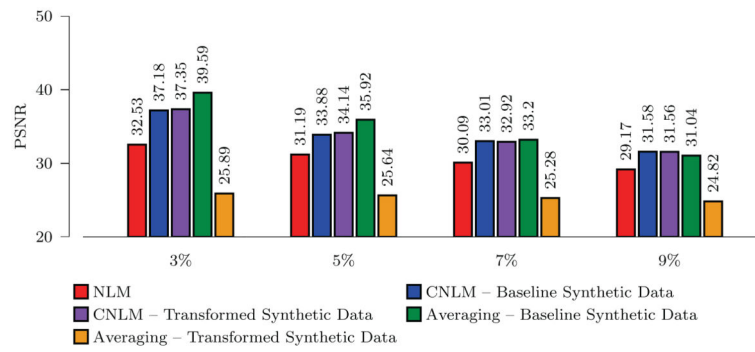
Figure 2. PSNR Comparison
Comparison of denoising performance between NLM and CNLM for both baseline and transformed synthetic datasets.
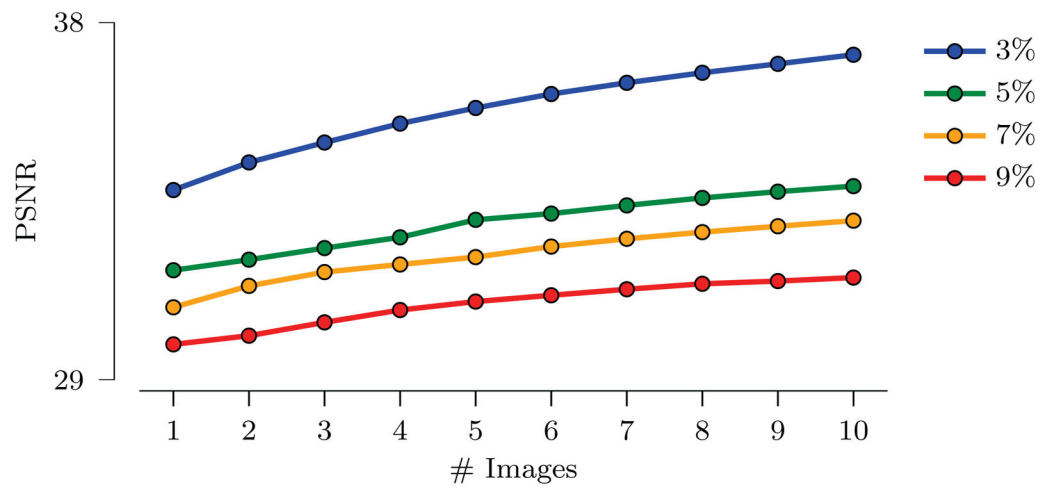
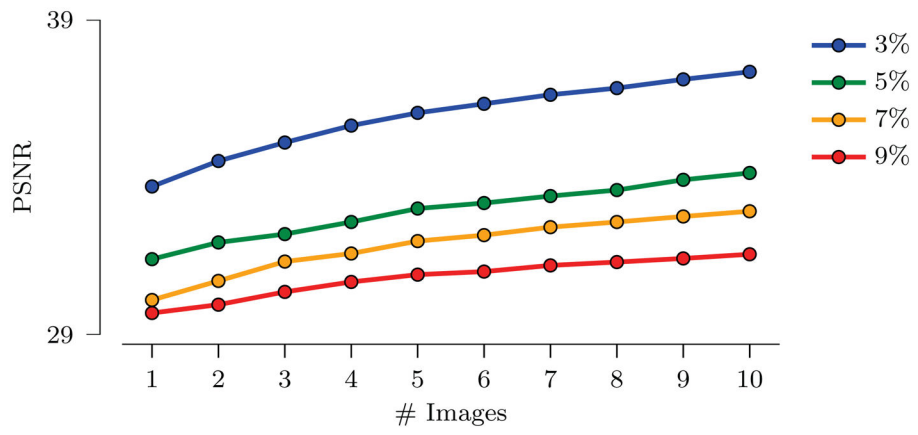**Figure 3. Influence of the Number of Co-Denoising Images**
PSNR trend in relation to the number of co-denoising images for 3%, 5%, 7%, 9% noise evaluated based on the baseline synthetic dataset.
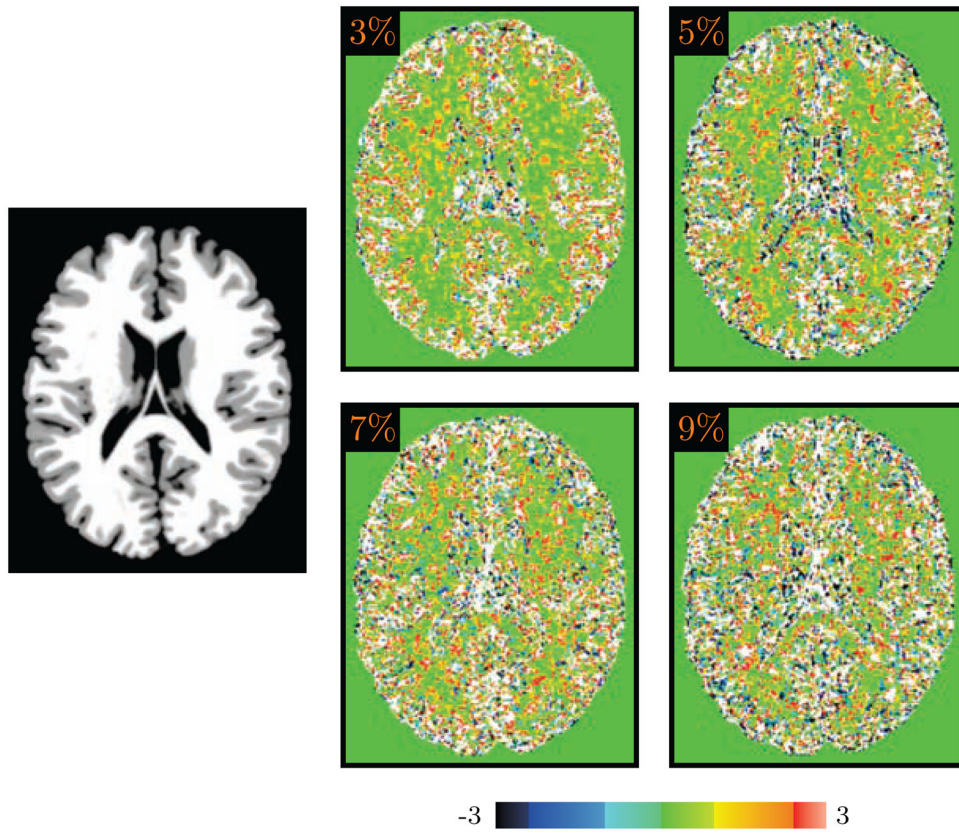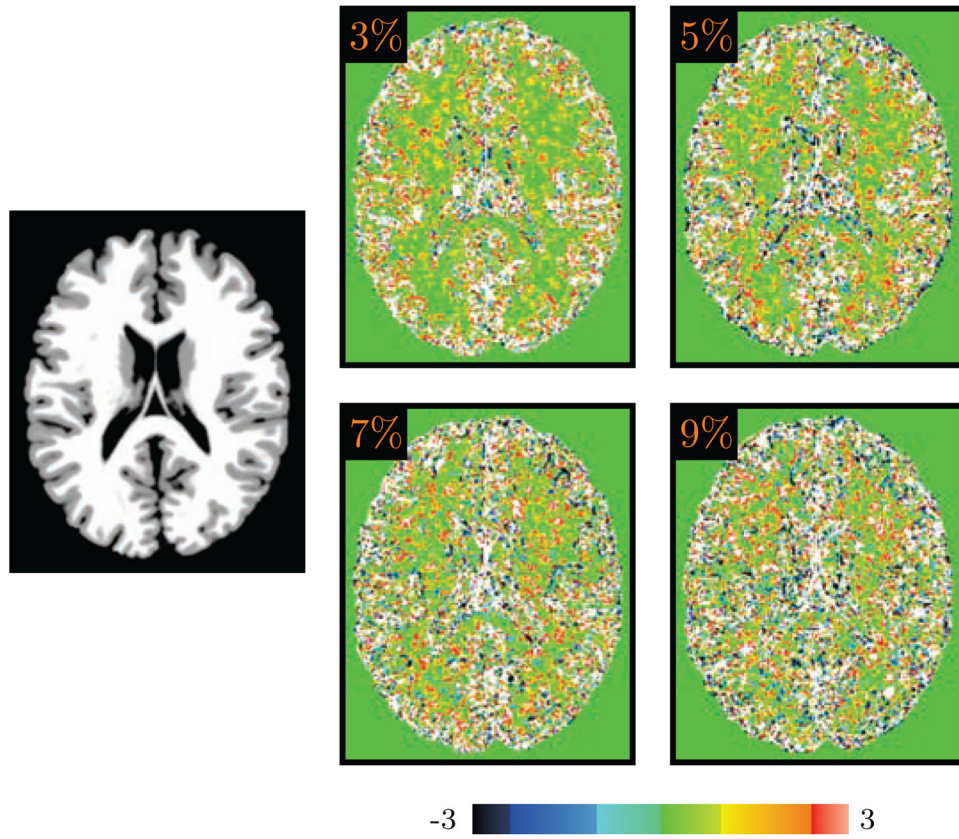
Figure 4. Comparison between NLM and CNLM
Warm colors indicate CNLM performs better than NLM; cool colors indicate otherwise. All results are based on the baseline synthetic dataset.
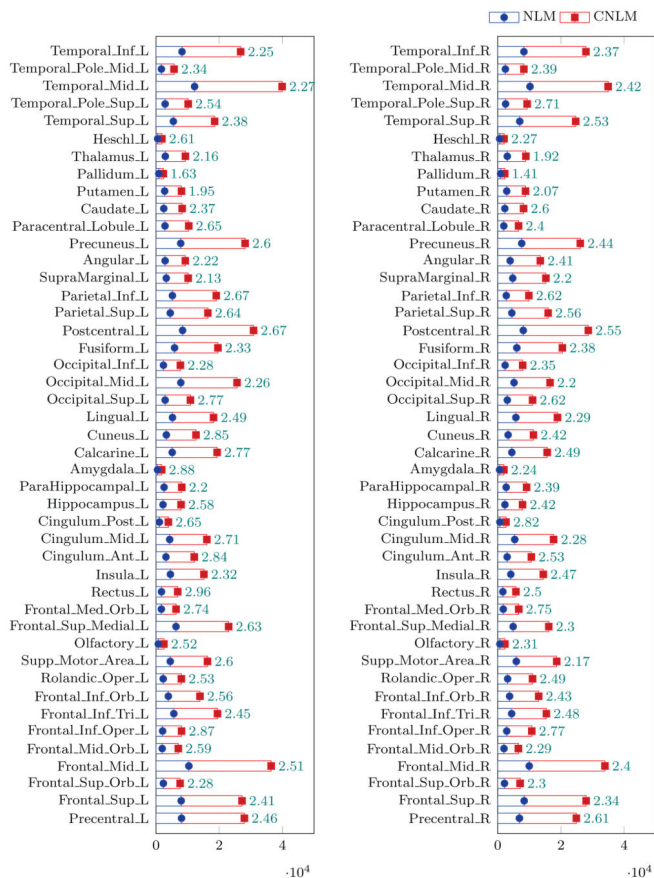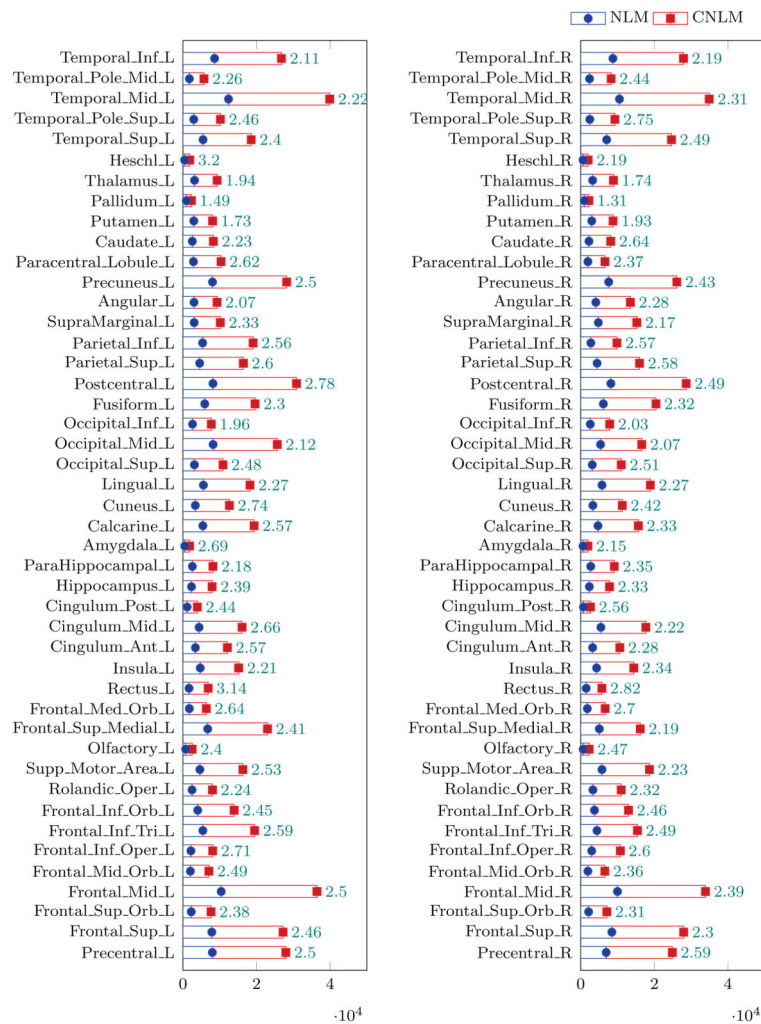
Figure 5. Regional comparison of CNLM and NLM

The results were computed based on the baseline synthetic dataset with 9% noise. The bar length represents the number of voxels where one method performs better than the other one. The number at each bar represents the ratio between the length of the CNLM (red) bar to that of the NLM (blue) bars.
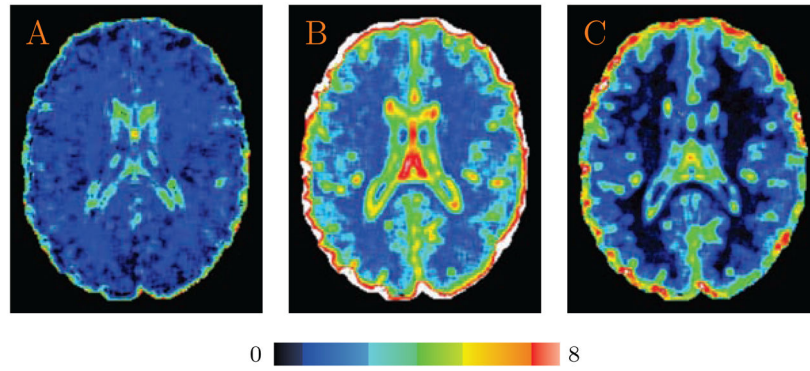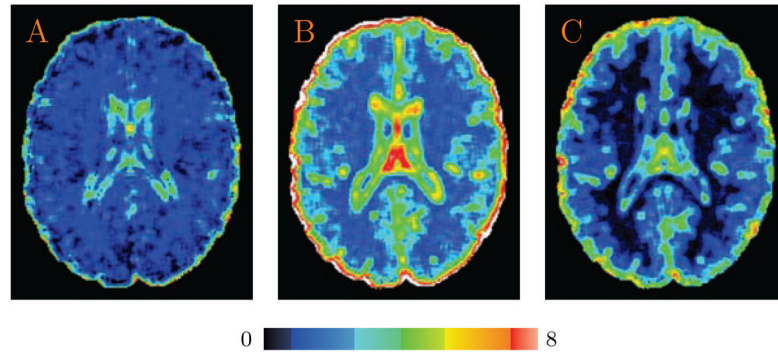
Figure 6. Structural Similarity

CNLM is able to identify a greater number of blocks with matching structures than NLM. The images show the sum of the 10 greatest weight values at each voxel for (A) NLM denoising and (B) CNLM denoising. (C) Subtraction of (A) from (B). The results are based on the baseline synthetic dataset with 3% noise.

Figure 7. Influence of the Number of Co-Denoising Images
The PSNR changing trend in relation to the number of co-denoising images for 3%, 5%, 7%, 9% noise evaluated based on the transformed synthetic data.

Figure 8. Comparison between NLM and CNLM
Warm colors indicate CNLM performed better than NLM; cool colors indicate otherwise.

All results are based on the transformed synthetic dataset.

Figure 9. Regional comparison of CNLM and NLM

The results were computed based on the transformed synthetic dataset with 9% noise. The bar length represents the number of voxels where one method performs better than the other one. The number at each bar represents the ratio between the length of the CNLM (red) bar to that of the NLM (blue) bars.

Figure 10. Structural Similarity
CNLM is able to identify a greater number of blocks with matching structures than NLM. The images show the sum of the 10 greatest weight values at each voxel for (A) NLM denoising and (B) CNLM denoising. (C) Subtraction of (A) from (B). The results are based on the transformed synthetic dataset with 3% noise.
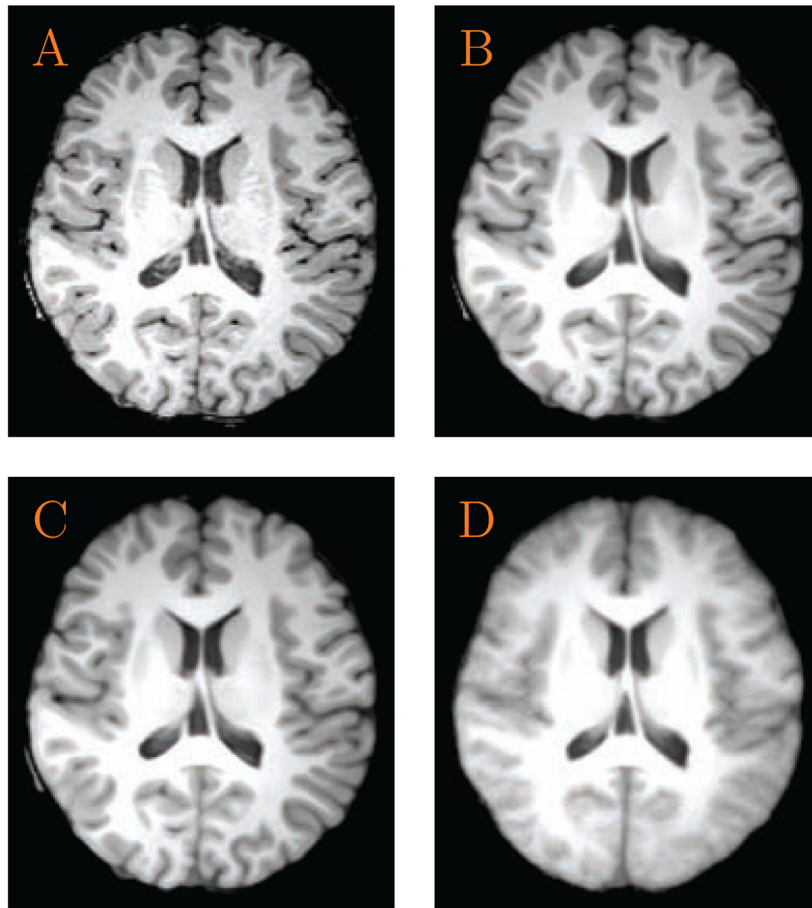
Figure 11. Denoising of Real Data
(A) The target noisy image; (B) CNLM-denoised image; (C) NLM-denoised image; (D)
Average image. Close-up views of cortical structures from various image slices are shown in
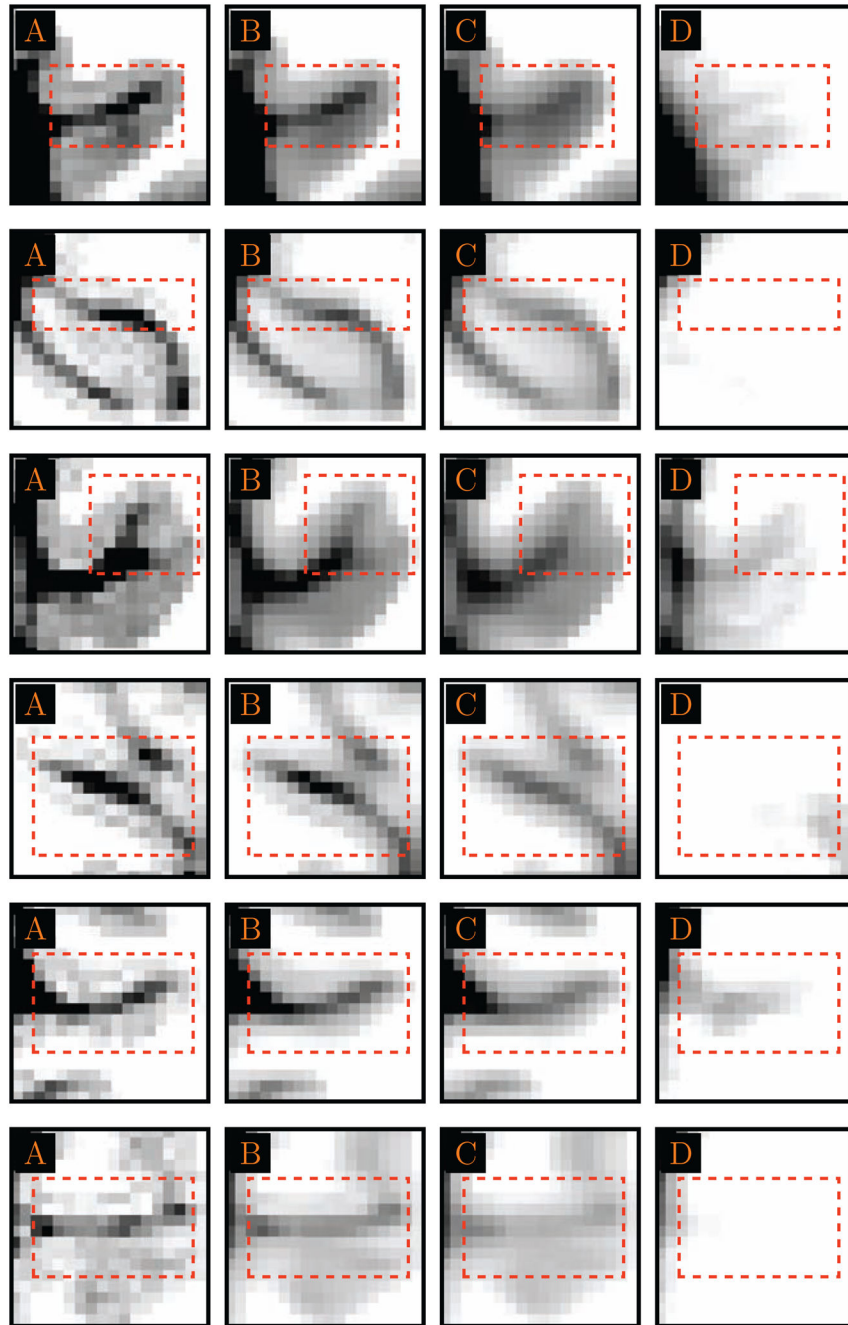Fig. 12.

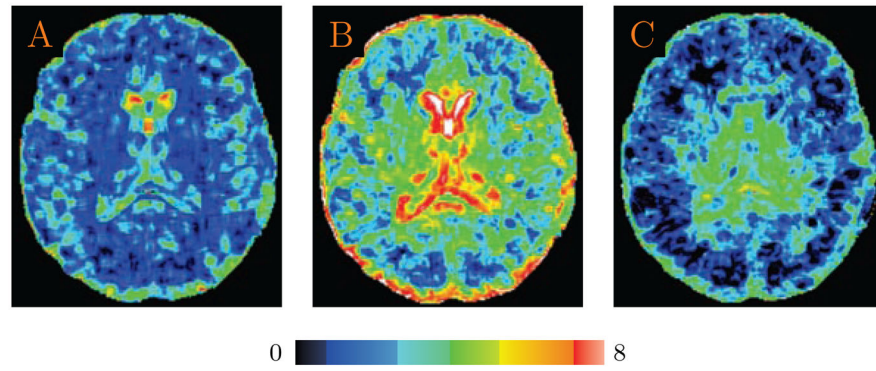**Figure 12. Close-Up Views**
Regional close-up views of Fig 11.

Figure 13. Structural Similarity
CNLM is able to identify a greater number of blocks with matching structures than NLM. The images show the sum of the 10 greatest weight values at each voxel for (A) NLM denoising and (B) CNLM denoising. (C) Subtraction of (A) from (B). The results are based on the real data.
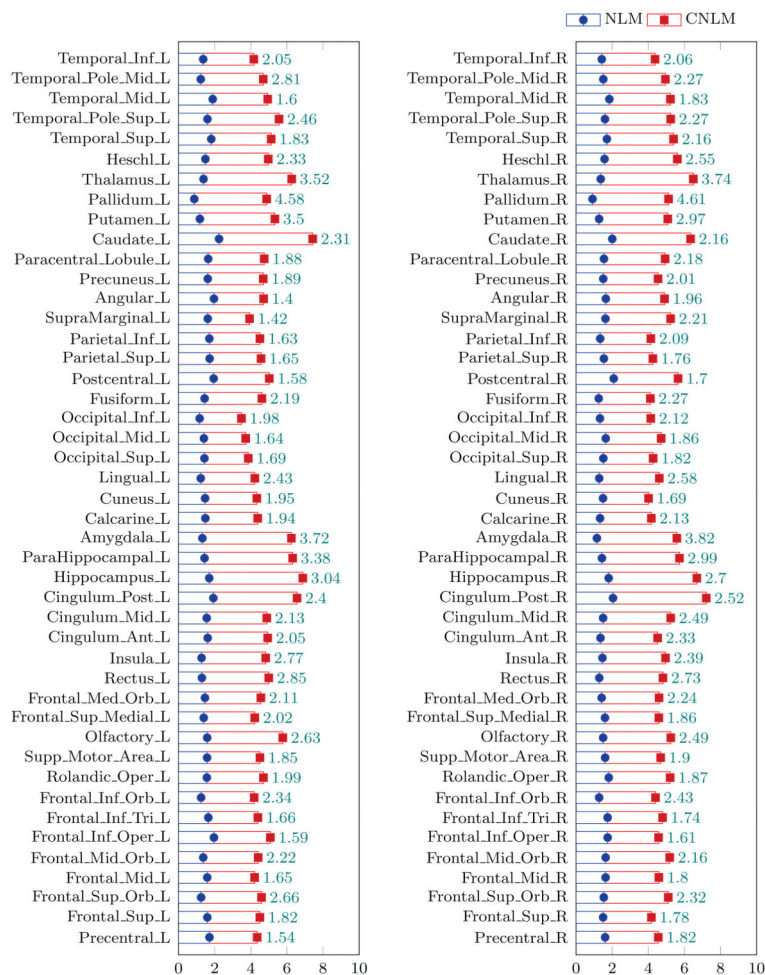
Figure 14. Regional comparison of CNLM and NLM
The results were computed based on the real data. The bar length represents the mean of the sum of the top 10 weights. The number at each bar represents the ratio between the length of the CNLM (red) bar to that of the NLM (blue) bars.