Rician nonlocal means denoising for MR images using nonparametric principal component analysis

Dong Wook Kim1, Chansoo Kim2, Dong Hee Kim1 and Dong Hoon Lim3*

Abstract
Denoising is always a challenging problem in magnetic resonance imaging (MRI) and is important for clinical diagnosis and computerized analysis, such as tissue classification and segmentation. The noise in MRI has a Rician distribution. Unlike additive Gaussian noise, Rician noise is signal dependent, and separating the signal from the noise is a difficult task. In this paper, we propose a useful alternative of the nonlocal mean (NLM) filter that uses nonparametric principal component analysis (NPCA) for Rician noise reduction in MR images. This alternative is called the NPCA-NLM filter, and it results in improved accuracy and computational performance. We present an applicable method for estimating smoothing kernel width parameters for a much larger set of images and demonstrate that the number of principal components for NPCA is robust to variations in the noise as well as in images. Finally, we investigate the performance of the proposed filter with the standard NLM filter and the PCA-NLM filter on MR images corrupted with various levels of Rician noise. The experimental results indicate that the NPCA-NLM filter is the most robust to variations in images, and shows good performance at all noise levels tested.

Keywords: image denoising, magnetic resonance (MR) image, nonlocal means (NLM), nonparametric principal component analysis (NPCA), Rician noise

1 Introduction
Magnetic resonance (MR) images are affected by several types of artifact and noise sources, such as random fluctuations in the MR signal mainly due to the thermal vibrations of ions and electrons. Such noise markedly degrades the acquisition of quantitative measurements from the data. The noise in MR images obeys a Rician distribution [1-3]. Unlike additive Gaussian noise, Rician noise is signal dependent, and consequently separating the signal from the noise is difficult.

There is an extensive literature on Rician noise reduction in magnetic resonance imaging (MRI), varying from the use of traditional smoothing filters to more elegant methods. Most conventional mask-based denoising filters, such as Gaussian and Wiener filters [4], are conceptually simple. However, they will most likely fail to reduce Rician noise in MRI, as they usually assume that the noise is Gaussian. Restored images may often look blurred and may be corrupted by artifacts that are usually visible around the edges. One way to overcome the problems of simple smoothing is to use a nonlocal means (NLM) filter [5-10]. These methods make use of the self-similarity of images, in that many structures show up more than once in the image. The NLM filter takes advantage of the high degree of redundancy of any natural image and produces an optimal denoising result if the noise can be modeled as Gaussian. Unfortunately, the method requires computation of the weighting terms for all possible pairs of pixels, making it computationally expensive. A number of recent reports on NLM denoising focused on shortcuts to make the method computationally practical [11-14]. One of the most compelling strategies is to exclude many weight computations between the image neighborhood feature vectors. Azzabou, et al. [11] and Tasdizen [13,14] proposed the so-called PCA-NLM filter, which uses the lower dimensional subspace of the space of image neighborhood vectors in conjunction with NLM using principal component analysis (PCA). More important, this approach was
also shown to result in increased accuracy over those that use the full-dimensional ambient space.

There are, however, some applications for which the PCA-NLM filter is not recommended because PCA is sensitive to image features and the presence of noise in the data, and the PCA-NLM filter is, therefore, highly dependent on the settings of its parameters.

In this paper, we propose a nonparametric PCA-NLM filter that is a useful alternative to the PCA-NLM filter for Rician noise reduction in MR images. The proposed filter uses PCA with ranked data instead of the original pixel data. We refer to this as the NPCA-NLM filter. We estimate the subspace dimensionality from parallel analysis [15-17] based on the artificial rank correlation matrix. In contrast to the method reported by Tasdizen [13,14], our estimation does not require the assumption of a Gaussian distribution and produces a more robust subspace dimensionality regardless of the images being denoised. We also propose a nonparametric method for optimal smoothing kernel width selection.

2 Background
2.1 Rician noise in MRI
In MRI, raw data are intrinsically complex valued and corrupted with zero mean Gaussian distributed noise with equal variance [1]. MR magnitude images are formed by simply taking the square root of the sum of the square of the two independent Gaussian random variables (real and imaginary images) pixel by pixel. After this nonlinear transformation, MR magnitude data can be shown to have a Rician distribution [1-3,9,18]. For an MR magnitude image, the Rician probability density function of the measured pixel intensity \( x \) is given by

\[
p(x|A) = \frac{x}{\sigma^2} \exp \left( -\frac{x^2 + A^2}{2\sigma^2} \right) I_0 \left( \frac{A \cdot x}{\sigma^2} \right),
\]

where \( I_0 \) is the modified zeroth-order Bessel function of the first kind, \( A \) is the underlying noise-free signal amplitude, and \( \sigma \) denotes the standard deviation of the Gaussian noise in the real and imaginary images. When \( A/\sigma \) is high, the Rician distribution approaches a Gaussian; when \( A/\sigma \) approaches 0, i.e., when only noise is present, the Rician distribution becomes Rayleigh distributed and Equation 1 reduces to [1]

\[
p(x|A) = \frac{x}{\sigma^2} \exp \left( -\frac{x^2}{2\sigma^2} \right).
\]

2.2 NLM filter
Starting from a true, discrete image \( u \), a noisy observation of \( u \) at pixel \( i \) is defined as \( v(i) = u(i) + n(i) \). Let \( N_i \) and \( v(N_i) \) denote a square neighborhood centered around pixel \( i \) and the image neighborhood vector, the elements of which are the gray-level values of \( v \) at \( N_i \), respectively. Also, let \( S_i \) be a square search-window of fixed size centered around pixel \( i \). Then, the NLM filter [5,6] defines an estimator for \( u(i) \) as

\[
\hat{u}(i) = \frac{1}{Z_i} \exp \left( -\frac{\|v(N_i) - v(N_j)\|^2}{h^2} \right) v(j),
\]

where \( Z_i \) is a normalization constant and \( h \) acts as a smoothing parameter controlling the decay of the exponential function.

This method is too slow to be practical. The high computational complexity is due to the cost of weight calculation for all pixels in the image during the process of denoising. For every pixel being processed, the whole image is searched and the differences between corresponding neighborhoods computed.

3 Proposed NPCA-NLM denoising
3.1 NPCA approach
We include a brief description of the projections of the image neighborhood vectors onto a lower dimensional space by NPCA, which uses their ranks instead of the original data. Let \( \Omega \) denote the entire set of pixels in the image and let \( \Psi \) be a randomly chosen subset of \( \Omega \). Also, let \( R_i \) denote the rank vector of \( \{x_p, i = 1, \ldots, Q\} \) in each window of \( Q \) size, where \( Q = r \times r \). The principal components of \( Q \) rank vector can be obtained from the eigenvectors of the rank covariance matrix

\[
C_R = \frac{1}{|\Psi|} \sum_{i \in \Psi} (R_i - \bar{R})(R_i - \bar{R})^T,
\]

where \( \bar{R} = (1/|\Psi|) \sum_{i \in \Psi} R_i \) is the rank mean and \( |\Psi| \) is the number of elements in the set \( \Psi \). Let \( \{e_p : p = 1, \ldots, q\} \) be the eigenvectors of \( C_R \), i.e., the principal neighborhoods, sorted in order of descending eigenvalues. Let the \( d \)-dimensional NPCA subspace be the space spanned by \( \{e_p : p = 1, \ldots, d\} \). Then, the projection of the image neighborhood vectors onto \( d \)-dimensional NPCA subspace is

\[
R_i^d = \sum_{p=1}^d (R_i \circ e_p)e_p,
\]

where \( (R_i \circ e_p) \) denotes the inner product of the two vectors.

Let \( f_i^d = [(R_i \circ e_{i_1}), \ldots, (R_i \circ e_{i_d})]^T \) be the \( d \)-dimensional vector of projection coefficients. Then, because of the orthonormality of the basis

\[
\|R_i^d - R_j^d\|^2 = \|f_i^d - f_j^d\|^2.
\]
Then, the $d$-dimension NLM algorithm in NPCA space is

$$\hat{u}_d(i) = \sum_{j \in S_i} \frac{1}{Z^d_i} \exp \left( -\frac{\|\mathbf{f}_i^j - \mathbf{f}_i\|^2}{h^2} \right) \nu(j),$$

where $Z^d_i = \sum_{j \in S_i} e^{-\|\mathbf{f}_i^j - \mathbf{f}_i\|^2/h^2}$ is the new normalizing term. Note that the proposed approach with $d = Q$ is equivalent to the useful NLM filter when applied to ranks rather than the original observations, i.e., Equation 4 with $d = Q$ becomes Equation 2 when calculated on the ranks.

### 3.2 Optimal smoothing parameter selection

Given a noisy image and a combination of $\mathcal{N}$ and $d$, there exists an optimal choice of the parameter $h$ in Equation 3.2 that yields the best output in terms of peak signal-to-noise ratio (PSNR). Figure 1 shows the PSNR of the estimator output $\hat{u}_d$ as a function of $h$ for an image corrupted with Rician noise ($\sigma=30$).

Buades et al. [5,6] reported that the smoothing parameter $h$ depends on the standard deviation of the noise $\sigma$, and typically a good choice for 2D images is $h = 10\sigma$. Tasdizen [14] used simple linear regression based on the least squares method as an automatic way of choosing $h$ given an image neighborhood size and $d$. Table 1 shows the linear fit parameters for several choices of $d$ of 7 × 7 image neighborhoods. Parameters such as those shown in Table 1 can be precomputed for all $\mathcal{N}$ and $d$ of interest.

Unlike Tasdizen [14], the linear fit parameters in Table 1 do not require the assumption of a Gaussian distribution of the noise with which the images are corrupted. Therefore, we expect that $h$ produced by these parameters will yield results for a much larger set of images than that from which they were learned.

### 3.3 Automatic dimensionality selection

Determining the number of components to retain is a crucial problem when using PCA. Of several methods proposed for determining the significance of principal components, the K1 method proposed by Kaiser [20] is the best known and most commonly utilized in practice [21]. However, this method often overextracts components [22]. Another commonly used approach is based on Cattell’s Scree test [23], which involves the visual exploration of a graphical representation of the eigenvalues. This method has been criticized for its subjectivity, as there is no objective definition of the cutoff point between the important and trivial factors. Parallel analysis [15] is more accurate than the above methods for determining the number of retained components, but it tends to overexert components [22]. Tasdizen [14] proposed a modification of the parallel analysis algorithm for determining the number of components in PCA of image neighborhoods for denoising. One of the main drawbacks of parallel analysis is that the number of principal components to retain is highly dependent on the images and the noise. Therefore, different numbers of principal components are required for images with different features. A quick solution to this problem is to use a modified procedure for parallel analysis of ranked data. Let $\lambda_p$ for $1 \leq p \leq Q$ denote the eigenvalues of $C_R$ in Equation 3 sorted in descending order. Similarly, let $\sigma_p$ denote the sorted eigenvalues of the artificial

![Figure 1 PSNR as a function of the parameter $h$ for the brain image](http://jivp.eurasipjournals.com/content/2011/1/15)

| $d$ | 6     | 10    | 20    | 49    |
|-----|-------|-------|-------|-------|
| Slope ($\beta$) | 2.2   | 2.37  | 3.275 | 4.55  |
| Intercept ($\alpha$) | 4     | 21.75 | 26.375| 40.125|
rank covariance matrix. Parallel analysis estimates data dimensionality as
\[ d = \max\{1 \leq p \leq Q \mid \lambda_p \geq \alpha_p \} \].

Figure 2 shows the automatic dimensionality selection results for brain, body, and knee images with noise \( \sigma \) ranging from 5 to 50 in increments of 5. The numbers of significant components for PCA as shown in Figure 2 were computed as about 10, 14, and 16 for brain, body, and knee images, respectively; however, the numbers of significant components for NPCA were all computed as about 15. It is important to note that the numbers of components vary more significantly with noise levels for PCA than for NPCA. Therefore, the number of principal components for NPCA is more robust to variations in noise as well as in images than for PCA.

Figure 3 shows the PSNR of the estimator output \( \hat{u} \) as a function of \( d \) for a brain image that was corrupted with Rician noise \((\sigma = 10)\). The curve for PCA showed a steep increase until the peak \( d \), for example, \( d = 15 \), after which the curve declined significantly. In contrast, the curve for NPCA increased similarly until the peak point, after which it became considerably flat. Therefore, as expected, the incorrect determination of the number of components for PCA can result in remarkable PSNR loss.

4 Experiments and results
The proposed NPCA-NLM filter was tested using 256 × 256, 8-bits/pixel MR images, i.e., the brain, body, and knee images shown in Figure 4. The performance of the proposed filter was tested for various levels of noise corruption and compared with the standard NLM filter and recently proposed PCA-NLM filter.

We generated MR magnitude data by adding Rician noise to noise-free images. The Rician noise was created as \( y(t_i) = \sqrt{[y(t_i) + e_1]^2 + e_2^2} \), where \( y \) is the true signal and \( e_1 \) and \( e_2 \) are random numbers from a Gaussian distribution with zero mean and standard deviation \( \sigma \) [9]. Four levels of noise were tested with \( \sigma = [10, 20, 30, 40] \). Figure 5 shows close-up images of the test images in Figure 4, which were corrupted with Rician noise \( \sigma = 10 \) and 40.

In addition to the visual quality, the performance of the proposed filter was measured by the following criteria: PSNR, mean absolute error (MAE), and structural dissimilarity (DSSIM). For 8-bit images, PSNR and MAE are defined as follows:

\[
\text{PSNR} = 10 \cdot \log_{10} \left[ \frac{255^2}{M \cdot N \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (I(i,j) - \hat{I}(i,j))^2} \right],
\]

\[
\text{MAE} = \frac{1}{M \cdot N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left| I(i,j) - \hat{I}(i,j) \right|,
\]

where \( \hat{I}(i,j) \) and \( I(i,j) \) denote the pixel values of the restored image and the original image, respectively, at location \((i,j)\) and \( M \times N \) is the size of the image. Higher PSNR values indicate better restoration, and smaller MAE values indicate that the filter can preserve more details and edges. However, they were not very well
matched to perceived visual quality. DSSIM is a distance metric derived from structural similarity (SSIM), which takes into account the human visual system, and was defined as follows [24]:

$$DSSIM(x, y) = \frac{1}{1 - SSIM(x, y)},$$

where SSIM is given by

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1) (2\text{cov}_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1) (\sigma_x^2 + \sigma_y^2 + c_2)},$$

where $\mu_x$ and $\mu_y$ are means of $x$ and $y$, respectively; $\sigma_x^2$ and $\sigma_y^2$ are variances of $x$ and $y$, respectively; and $\text{cov}_{xy}$ is covariance of $x$ and $y$. The constants were set as follows: $c_1 = 0.01L$ and $c_2 = 0.03L$, and $L$ was 255 for 8-bits/pixel gray scale images.

4.1 Visual quality comparison

Figures 6 and 7 show the results of applying the three filters to noisy images corrupted with Rician noise $\sigma = 10$ in Figure 5a-c, and Rician noise $\sigma = 40$ in Figure 5d-f, respectively.

As shown in Figure 6, the three filters based on the NLM filter performed well on images with low noise variance ($\sigma = 10$). The differences in performance of these filters are difficult to distinguish in the restored images for low noise, but inspection of images with high noise in Figure 7 showed that the denoising effects of the NPCA-NLM filter and PCA-NLM filter were almost identical, except for slight blurring of the output from the PCA-NLM filter. However, the impact of noise on the standard NLM filter was clearly visible, and the restored images contained considerable noise spots.

4.2 Quantitative comparison

The quantitative performances in terms of PSNR, MAE, and DSSIM for all of the algorithms are given in Tables 2, 3, and 4, respectively. The results shown in these tables indicate that the NPCA-NLM filter had the best performance among the filters examined for the brain image over all noise ranges $\sigma = 10, 20, 30, 40$. The NPCA-NLM filter showed better performance than the NLM filter and PCA-NLM filter for body and knee images, which were corrupted with $\sigma = 10$ and 40. It should be noted that the performance gap between the NPCA-NLM filter and the PCA-NLM filter increased on images with high noise. This was expected, as the PCA-NLM filter is more sensitive to noise. Our filter with the PCA-NLM filter still showed good performance on images that were corrupted with $\sigma = 20$ and 30.
Figure 5 Close-up images corrupted by Rician noise.

Figure 6 Comparison of the restoration on corrupted images in Figure 5a-c.
Of the three filters investigated, the NPCA-NLM filter appeared to be the most robust to variations in images, performing well at all noise distributions tested.

### Table 2 PSNR values for various Rician noise levels

|              | σ = 10 | σ = 20 | σ = 30 | σ = 40 |
|--------------|--------|--------|--------|--------|
| **Brain image** |        |        |        |        |
| NLM          | 32.62  | 27.39  | 23.26  | 20.10  |
| PCA-NLM      | 32.64  | 27.29  | 23.61  | 20.47  |
| NPCA-NLM     | 32.84  | 27.63  | 23.65  | 20.48  |
| **Body image** |        |        |        |        |
| NLM          | 31.94  | 25.07  | 20.78  | 17.80  |
| PCA-NLM      | 31.94  | 25.13  | 20.93  | 17.87  |
| NPCA-NLM     | 32.08  | 25.08  | 20.94  | 17.96  |
| **Knee image** |        |        |        |        |
| NLM          | 31.80  | 27.99  | 24.43  | 21.43  |
| PCA-NLM      | 32.13  | 28.17  | 24.66  | 21.82  |
| NPCA-NLM     | 32.31  | 28.25  | 24.62  | 21.86  |

### Table 3 MAE values for various Rician noise levels

|              | σ = 10 | σ = 20 | σ = 30 | σ = 40 |
|--------------|--------|--------|--------|--------|
| **Brain image** |        |        |        |        |
| NLM          | 4.42   | 9.03   | 14.62  | 21.04  |
| PCA-NLM      | 4.40   | 9.07   | 14.09  | 20.06  |
| NPCA-NLM     | 4.29   | 8.88   | 14.02  | 20.03  |
| **Body image** |        |        |        |        |
| NLM          | 5.06   | 12.65  | 20.97  | 29.62  |
| PCA-NLM      | 5.06   | 12.59  | 20.63  | 29.35  |
| NPCA-NLM     | 4.99   | 12.64  | 20.65  | 29.11  |
| **Knee image** |        |        |        |        |
| NLM          | 4.93   | 7.98   | 11.84  | 16.66  |
| PCA-NLM      | 4.77   | 7.82   | 11.46  | 15.66  |
| NPCA-NLM     | 4.70   | 7.78   | 11.58  | 15.59  |

5 Conclusion

We proposed an NPCA-NLM filter, which is a useful alternative to the PCA-NLM filter for Rician noise reduction in MR images. The filter uses PCA with
Table 4 DSSIM values for various Rician noise levels

| Brain image | σ = 10 | σ = 20 | σ = 30 | σ = 40 |
|-------------|--------|--------|--------|--------|
| NLM         | 287.2  | 88.4   | 35.12  | 17.17  |
| PCA-NLM     | 288.6  | 85.1   | 38.21  | 18.75  |
| NPCA-NLM    | 302.7  | 93.2   | 36.86  | 18.81  |

| Body image  | σ = 10 | σ = 20 | σ = 30 | σ = 40 |
|-------------|--------|--------|--------|--------|
| NLM         | 304.0  | 59.5   | 22.35  | 11.55  |
| PCA-NLM     | 303.9  | 60.2   | 22.96  | 11.58  |
| NPCA-NLM    | 313.9  | 59.5   | 23.01  | 11.84  |

| Knee image  | σ = 10 | σ = 20 | σ = 30 | σ = 40 |
|-------------|--------|--------|--------|--------|
| NLM         | 236.7  | 98.8   | 44.64  | 22.59  |
| PCA-NLM     | 255.9  | 103.3  | 47.18  | 24.97  |
| NPCA-NLM    | 266.4  | 105.4  | 46.83  | 25.20  |

The authors declare that they have no competing interests.

Acknowledgements
This work was supported by the RACS 2010-2014, Production of fine-scale scenario of future climate change using regional climate models and analysis of uncertainties.

Author details
1Department of Statistics, Busan National University, Busan, Korea
2Department of Applied Mathematics, Kongju National University, Gongju, Korea
3Department of Information Statistics and RINS, Gyeongsang National University, Jinju 660-701, Korea

Competing interests
The authors declare that they have no competing interests.

References
1. Gudbjartsson H, Patz S: The Rician distribution of noisy MRI data. Magn Reson Med 1995, 34(6):1010-9.
2. Macovski A: Noise in MRI. Magn Resonance Med 1996, 36(3):494-497.
3. Nowak RD: Wavelet-based Rician noise removal for magnetic resonance imaging. IEEE Trans Image Process 1999, 8(10):1408-1419.
4. Khieddine A, Benhammouda K, Puech W: Digital image restoration by Wiener filter in 2D case. Adv Eng Softw 2007, 38:515-516.
5. Buades A, Coll B, Morel JM: A review of image denoising algorithms, with a new one. Multiscale Model Simul 2006, 4(2):490-530.
6. Buades A, Coll B, Morel JM: A non-local algorithm for image denoising. Proc IEEE Trans IEEE Conf Computer Vision and Pattern Recognition 2005, 60-65.
7. Orchard J, Ebrahim M, Wang A: Efficient nonlocal-means denoising using the SVD. Proc Int Conf Image Processing 2008, 1732-1735.
8. Coupe P, Hellier P, Kervrann C, Barillot C: Nonlocal means-based speckle filtering for ultrasound images. IEEE Trans Image Process 2009, 18(10):2221-2229.
9. He L, Greenshields JR: A nonlocal maximum likelihood estimation method for Rician noise reduction in MR images. IEEE Trans Med Imaging 2009, 28(2):165-172.
10. Manjon JV, Coupe P, Mari-Bonmati L, Collins DL, Robles M: Adaptive non-local means denoising of MR images with spatially varying noise levels. J Magn Reson Imaging 2010, 31:192-203.
11. Azzabou N, Parajios N, Guichard F: Image denoising based on adapted dictionary computation. Proc Int Conf Image Processing 2007, III:109-112.
12. Coupe 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 Trans Med Imaging 2008, 27(4):425-441.
13. Tasdizen T: Principal components for non-local means image denoising. Proc Int Conf Image Processing 2008, 1728-1731.
14. Tasdizen T: Principal neighborhood dictionaries for nonlocal means image denoising. IEEE Trans Image Process 2009, 18(12):2649-2660.
15. Horn JL: A rationale and test for the number of factors in factor analysis. Psycho-metrica 1965, 30(2):179-185.
16. Goroff LD: An improvement on horn’s parallel analysis methodology for selecting the correct number of factor to retain. Educ Psychol Meas 1993, 53(3):377-395.
17. Ledesma RD, Valero-Mora P: Determining the number of factors to retain in EFA-an easy-to-use computer program for carrying out parallel analysis. Practical Ass Res Eval 2007, 12(2):1-11.
18. Sijbers J, den Dekker AJ, Scheunders P, Dyck DV: Maximum-likelihood estimation of Rician distribution parameters. IEEE Trans Med Imaging 1998, 17(3):357-361.
19. Sen PK: Estimates of regression coefficient based on Kendall’s tau. J Am Stat Assoc 1968, 63:1379-1389.
20. Kaiser HF: The application of electronic computers to factor analysis. Educ Psychol Meas 1960, 20:141-151.
21. Fabrigar LR, Wegener DT, MacCallum RC: Strahan EJ: Evaluating the use of exploratory factor analysis in psychological research. Psychol Methods 1999, 3:272-299.
22. Zwick WR, Velicer WF: Comparison of five rules for determining the number of components to retain. Psychol Bull 1986, 99:432-442.
23. Catell RB: The scree test for the number of factors. Multivar Behav Res 1966, 1:245-276.
24. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP: Image quality assessment: From error visibility to structural similarity. IEEE Trans Image Process 2004, 13(4):600-612.

Received: 7 February 2011 Accepted: 14 October 2011 Published: 14 October 2011