Rician Noise Removal Approach for Brain MR Images Using Kernel Principal Component Analysis

Ashish Phophalia(✉) and Suman K. Mitra

Dhirubhai Ambani Institute of Information and Communication Technology, 382007 Gandhinagar, Gujarat, India
{ashish_phophalia,suman_mitra}@daiict.ac.in

Abstract. It has been observed that the noise accumulated in medical images due to various reasons during acquisition process is Rician in nature. A Rician noise removal method of Brain Magnetic Resonance (MR) Images using Kernel Principal Component Analysis (KPCA) is proposed in this paper. The proposed approach is non-parametric in nature. It explores the image space for *non-local* similar patch search and clusters them accordingly. The basis vectors are then learned using KPCA for each cluster which makes the proposed method data adaptive in nature. The approach has been applied to 2D phantom Brain MR images and experimental results are comparable to the other state-of-the-art methods in terms of various quantitative measures.

Keywords: Kernel Principal Component Analysis (KPCA) · Magnetic Resonance Image (MRI) · Rician noise removal

1 Introduction

Image Restoration is considered as one of the crucial ingredient of Medical Image Analysis systems. The possible sources for addition of noise are various parameters of the acquisition process such as flip angle, scan time, coil resistance, dielectric and inductive losses in sample, patient movement etc. [12]. MRI, being a non-invasive technique, offers many advantages in clinical analysis but the disturbances or *noise* induced in acquisition process degrade the quality of the signal. In Medical Image Denoising problem, the noise model is found to be Rician in nature which is different from commonly used distributions such as Gaussian, Poisson, etc. [8].

It has been shown that the intensities of MR images represent magnitude of underlying complex data which follows Rice distribution [7]. The real and imaginary parts are modeled as independently distributed Gaussian with means a_r and a_i respectively, with same variance σ^2. The probability density function (pdf) of Rician random variable y is defined as follows:

$$f_Y(y|a,\sigma) = \frac{y}{\sigma^2} e^{\left(-\frac{y^2+a^2}{2\sigma^2}\right)} I_0\left(\frac{ya}{\sigma^2}\right), y > 0 \tag{1}$$

M. Kryszkiewicz et al. (Eds.): PReMI 2015, LNCS 9124, pp. 545–553, 2015.
DOI: 10.1007/978-3-319-19941-2_52

where $a = \sqrt{a_r^2 + a_i^2}$ is underlying noise free signal amplitude and $I_n(z)$ is n^{th} order modified bessel function of first kind. When Signal to Noise Ratio (SNR, here it is a/σ) is high, the Rician distribution approaches a Gaussian; when SNR approaches to zero (that is only noise is present, $a \to 0$) the Rician distribution becomes Rayleigh distribution and the pdf turns out to be

$$f_Y(y|a \to 0, \sigma) = \frac{y}{\sigma^2} e^{\left(-\frac{y^2}{2\sigma^2}\right)} \tag{2}$$

Hence, the conventional methods for Rician noise removal first try to find the background portion in the medical images where no signal is assumed. Hence, one can use Rayleigh distribution in background portion and Gaussian distribution in the rest (where SNR is assumed to be high enough) [9,16]. However under the noisy condition, it is difficult to find proper background in the image.

Recent methods use the principle of non-local self similarity for image restoration task, where the first step involves finding out the similar patches (in terms of some predefined criteria such as Euclidean distance) that are similar to a given reference patch from the image [1]. Thereafter, an orthonormal basis is inferred for each patch and shrinkage is performed on the coefficients when the patch is projected on that basis, coefficients are sparse in nature as described in [4,6,14].

Out of recently proposed techniques, BM3D [4] is most popular. BM3D technique creates a 3D stack of similar patches, projects it onto a 3D basis (tensor product of 2D-DCT and 1D-Haar), and performs hard thresholding of these coefficients followed by basis inversion, thereby allowing a coupled update of the coefficients [4]. Another class of methods such as [5,13], first to cluster similar patches and then learn basis for each cluster instead of searching the similar patches for each underlying reference patch. However, due to nature of noise, straight forward implication of natural image denoising methods has not been advocated for medical images. The NLM method has been extended for Medical Image denoising problem in [11] where bias correction needs to be considered. BM3D has been extended using a suitable invertible transformation of the medical data into another domain where data behaves like Gaussian distributed in resultant domain. The most commonly known such kind of transformation for this purpose is Anacombe's Transformation, also known as Variance Stabilization Technique (VST). Recently, VST has been proposed in [7] for Rician distributed data and BM3D method is referred as BM3D+VST method. The BM3D+VST method can be summarized mathematically as follows:

$$\hat{y} = VST^{-1}(BM3D(VST(z, \sigma), \sigma_{VST}), \sigma) \tag{3}$$

where VST^{-1} denotes the inverse VST, σ_{VST} is the stabilized standard deviation induced by VST and z denotes the additive white Gaussian noise whose true intensity is represented by y. However, BM3D+VST is extended to 3D medical data as BM4D method in [10]. This manuscript focuses on 2D data denoising methods only.

The aim of this article is to explore a direct technique that can handle Rician noise suitably giving rise to noise removal as good as BM3D+VST, if not better.

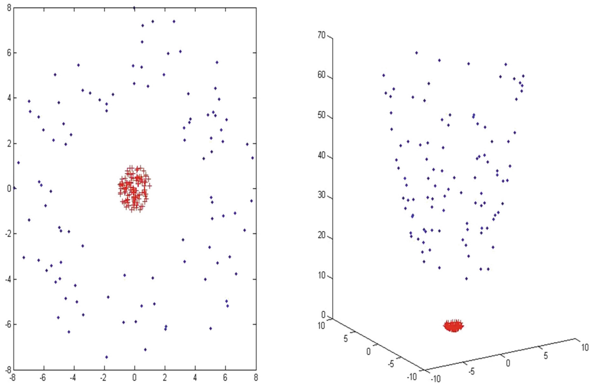


Fig. 1. Transformation of two circular data sets into higher dimension space using kernel method where separation between them is more prominent and can be classified using linear hyper-surface.

We have extended PCA based method using Rough Set based clustering proposed in [13] to Rician noise model and bias term correction is also made, referred as ER-PCA in the paper. We have proposed a new Kernel based PCA (KPCA) method for Rician noise. However, we have adopted the clustering strategy used in [13], which is non-local approach in *true-sense*. As per our knowledge, KPCA has not been applied for Rician noise removal in medical image yet. The kernel based methods can find non-linearity of data in Feature Space. Recently, kernel based methods have been used in Medical imaging in [2,15,19]. However, choice of appropriate kernel for given data is undecidable. In the current proposal, Gaussian kernel is used and the performance of noise removal technique is at par with the state-of-the-art methods.

The paper has been arranged in following manner: Sect. 2 presents proposed method using KPCA. Section 3 compares proposed method with other state-of-the-art methods. The manuscript is concluded in Sect. 4.

2 Proposed Method Using KPCA

A non-parametric variant of PCA, known as Kernal Principal Component Analysis (KPCA) has been explored for Rician noise removal. The KPCA tries to explore structure in the data in Feature Space instead of Image Space itself and tries to capture higher-order dependencies in the data. In Fig. 1, two class data is shown in circular form and transformed to higher dimension for classification purpose, where transformation is $\phi(x) : (x_1, x_2) \rightarrow (x_1, x_2, x_1^2 + x_2^2)$. Hence, one can find a discriminating plane (linear surface) in higher dimensions which is not possible in two dimensions for given data points.

In KPCA, this nonlinearity is introduced by first mapping the data into another space F using a nonlinear map $\phi : R^N \rightarrow F$, before standard linear PCA is carried out in F using the mapped samples $\phi(x_k)$. The map ϕ and the

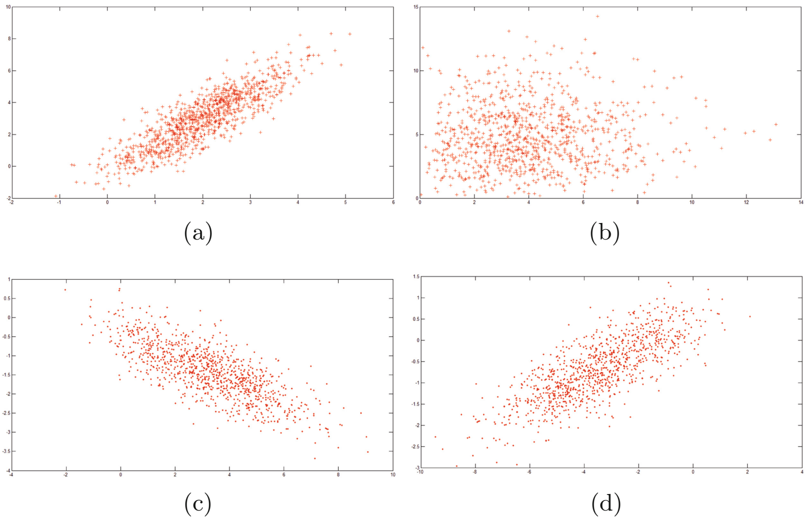


Fig. 2. Reconstruction using PCA and KPCA over synthetic data with Rician noise. (a) Synthetic Data, (b) Rician Noisy Data, (c) Reconstruction using PCA and (d) Reconstruction using KPCA.

space F are determined implicitly by the choice of a kernel function k, which acts as a similarity measure. This mapping computes the dot product between two input samples x and y mapped into F:

$$k(x; y) = \phi(x).\phi(y) \tag{4}$$

One can show that if k is a positive definite kernel, then there exists a map ϕ into a dot product space F such that Eq. 4 holds. The space F then has the structure of a so-called Reproducing Kernel Hilbert Space (RKHS) [2].

The identity Eq. 4 is important for KPCA since PCA in F can be formulated entirely in terms of inner products of the mapped samples. Thus, we can replace all inner products by evaluations of the kernel function. This has two important consequences: first, inner products in F can be evaluated without computing $\phi(x)$ explicitly. This allows to work with a very high-dimensional, possibly infinite-dimensional RKHS F. Second, if a positive definite kernel function is specified, we need to know neither ϕ nor F explicitly to perform KPCA since only inner products are used in the computations. Commonly used positive definite kernel functions are *polynomial kernel* of degree $d \in N, k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}.\mathbf{y})^d$ or $k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}.\mathbf{y} + 1)^d$ or Gaussian kernel of width $\sigma > 0$, $k(\mathbf{x}, \mathbf{y}) = exp\left(- \|\mathbf{x} - \mathbf{y}\|^2 / 2\sigma^2\right)$. In all the experiments, Gaussian kernel has been used which is isotropic stationary in nature and also satisfies Mercer's Theorem [19].

A synthetic experiment has been performed as shown in Fig. 2 where Rician noise added in the synthetic data. However, KPCA (with Gaussian kernel) is able to preserve orientation of the data in a better way as compared to PCA based reconstruction.

The outline of present work can be described as follows:

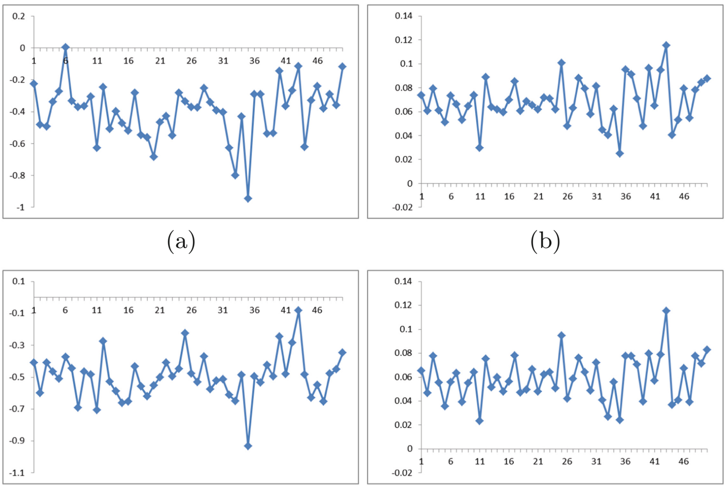
1. Get the clusters of patches from the given noisy image using Rough set based method (as described in [13]).
2. For each cluster, get the basis vectors using KPCA method along pixel positions. For patches of size $p \times p$, kernel matrix would be of size $p^2 \times p^2$. Hence, the method is data adaptive in nature.
3. Project the noisy image patches on the obtained basis vectors in the KPCA domain.
4. Apply coefficient shrinkage method on these projected patches to get the denoised patches. Transform them back to image space.
5. Remove the bias term from each pixel of the denoised image.

$$I_{unbiased} = \sqrt{max(\hat{I}(i,j)^2 - 2h^2, 0)} \tag{5}$$

where h is the standard deviation of noise and $\hat{I}$ is the image obtained by step (4).

3 Experimental Results

This Section encompasses the qualitative and quantitative evaluations of the proposed method along with some of the state-of-the-art methods. The experiments



(a) (b) (c) (d)

Fig. 3. Difference comparison of KPCA with reference to BM3D+VST method (at zero level vertically) for 50 slices for noise standard deviation equal to 15 (a) T1 images with PSNR difference values, (b) T1 images with MSSIM difference values, (c) T2 images with PSNR difference values and (d) T2 images with MSSIM difference values.

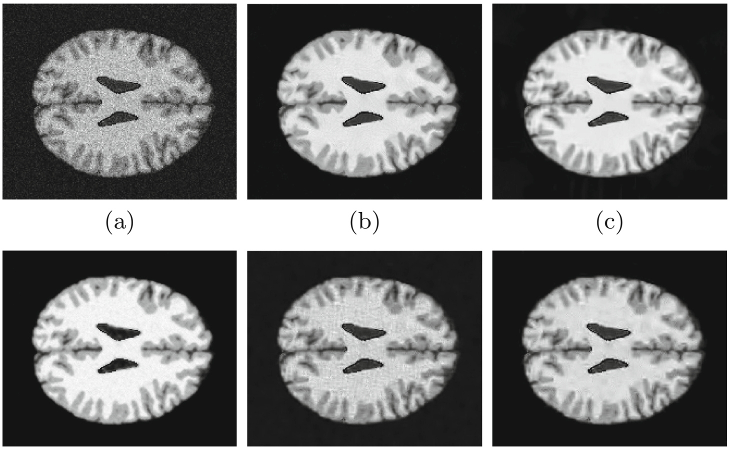


Fig. 4. (a) Synthetic Noisy T1 Image with Rician noise standard deviation=15 and PSNR =22.7220 dB, Denoised image using (b) UNLM method, PSNR = 34.4622 dB, (c) BM3D+VST method, PSNR = 34.2393 dB, (d) RS-NLM method, PSNR = 32.5856 dB, (e) ER-PCA method, PSNR = 33.8155 dB, (f) KPCA method, PSNR = 34.0241 dB.

Table 1. Performance comparison of proposed denoising strategy with different approaches on various quantitative measures under *Rician Noise assumption* in Brain Web database (**slice** = **70** & **100**, Modality = T1, image size = 181 × 217 and patch size = 5 × 5). Best figures are shown in Bold.

Noise SD	Methods	Slice 70				Slice 100			
		PSNR	RMSE	MSSIM	FSIM	PSNR	RMSE	MSSIM	FSIM
5	Noisy	32.4293	37.1660	0.6134	0.9296	32.2588	38.6549	0.5564	0.8922
	UNLM [11]	39.0519	8.0889	0.9832	0.9845	40.1551	6.2744	0.9882	0.9887
	BM3D+VST [7]	**40.9727**	**5.1937**	0.9602	0.9843	41.4921	4.6118	0.9602	0.9857
	RS-NLM [13]	39.8595	6.7163	**0.9851**	**0.9853**	**41.5829**	**4.5164**	**0.9914**	**0.9913**
	ER-PCA	40.4514	5.8606	0.9791	0.9764	39.9719	6.5447	0.9689	0.9563
	KPCA	40.2107	6.1946	0.9197	0.9797	41.2223	4.9073	0.9866	0.9850
10	Noisy	26.4115	148.5702	0.4717	0.8149	26.2398	154.5629	0.4183	0.7567
	UNLM [11]	35.9894	16.3733	0.9608	0.9643	36.9916	12.9993	0.9707	0.9724
	BM3D+VST [7]	**36.3738**	**14.9866**	0.9040	0.9607	36.8590	13.4025	0.9132	0.9653
	RS-NLM [13]	35.8260	17.0011	**0.9631**	**0.9645**	**37.2231**	**12.3246**	**0.9762**	**0.9770**
	ER-PCA	35.7387	17.3464	0.9389	0.9439	36.3168	15.1846	0.9597	0.9484
	KPCA	36.1061	15.9395	0.9586	0.9522	36.6642	14.0172	0.9682	0.9628
15	Noisy	22.8950	333.8752	0.3744	0.7177	22.7220	347.4434	0.3331	0.6495
	UNLM [11]	33.5147	28.9475	0.9299	**0.9391**	**34.4622**	**23.2732**	0.9453	**0.9498**
	BM3D+VST [7]	**33.7666**	**27.3162**	0.8583	0.9368	34.2393	24.4992	0.8684	0.9447
	RS-NLM [13]	32.1179	39.9292	0.9273	0.9244	32.5856	35.8523	**0.9472**	0.9448
	ER-PCA	33.2440	30.8093	0.9133	0.9178	33.8155	27.0103	0.9377	0.9287
	KPCA	33.4097	29.6557	**0.9323**	0.9262	34.0241	25.7438	0.9469	0.9404

Table 2. Performance comparison of proposed denoising strategy with different approaches on various quantitative measures under *Rician Noise assumption* in Brain Web database (**slice = 70** & **100**, Modality = T2, image size = 181 × 217 and patch size = 5 × 5). Best figures are shown in Bold.

Noise SD	Methods	Slice 70				Slice 100			
		PSNR	RMSE	MSSIM	FSIM	PSNR	RMSE	MSSIM	FSIM
5	Noisy	32.4349	37.1185	0.6257	0.9365	32.2639	38.6095	0.5691	0.9052
	UNLM [11]	34.4831	23.1617	0.9822	0.9813	35.2666	19.3385	0.9869	0.9858
	BM3D+VST [7]	**40.4738**	**5.8305**	0.9648	**0.9861**	**41.0752**	**5.0764**	0.9663	0.9885
	RS-NLM [13]	36.9814	13.0300	**0.9856**	0.9835	37.6322	11.2166	**0.9915**	**0.9900**
	ER-PCA	39.8618	6.7127	0.9783	0.9727	39.1934	7.8297	0.9610	0.9473
	KPCA	37.8578	10.6487	0.8002	0.9782	38.1996	9.8429	0.7610	0.9797
10	Noisy	26.4322	147.8642	0.4956	0.8356	26.2550	154.0201	0.4408	0.7757
	UNLM [11]	32.9818	32.7262	0.9618	0.9623	33.8132	27.0246	0.9710	0.9687
	BM3D+VST [7]	**35.7377**	**17.3504**	0.9181	**0.9681**	35.8044	17.0860	0.9683	0.9637
	RS-NLM [13]	34.5041	23.0502	**0.9691**	0.9676	35.2411	19.4522	**0.9799**	**0.9766**
	ER-PCA	34.8288	21.3894	0.9432	0.9323	34.5457	22.8303	0.9262	0.9008
	KPCA	35.0519	20.3182	0.8527	0.9567	**36.1329**	**15.8413**	0.9184	0.9727
15	Noisy	22.9275	331.3825	0.4131	0.7519	22.7460	345.5293	0.3676	0.6776
	UNLM [11]	31.4832	46.2121	0.9346	0.9408	32.1181	39.9271	0.9472	0.9456
	BM3D+VST [7]	**32.8504**	**33.7321**	0.8769	**0.9496**	**33.1694**	**31.3427**	0.8855	**0.9567**
	RS-NLM [13]	31.9206	41.7849	**0.9446**	0.9452	32.6973	34.9423	**0.9601**	0.9543
	ER-PCA	31.7529	43.4297	0.9034	0.8973	31.7770	43.1894	0.8989	0.8718
	KPCA	32.3606	37.7592	0.9363	0.9346	32.8539	33.7049	0.9516	0.9439
20	Noisy	20.4499	518.2594	0.3540	0.6871	20.2642	611.8738	0.3162	0.6059
	UNLM [11]	30.0502	64.2771	0.9063	0.9205	30.5757	56.9519	0.9199	0.9216
	BM3D+VST [7]	**30.7168**	**55.1319**	0.8426	**0.9303**	**30.9691**	**52.0201**	0.8508	**0.9398**
	RS-NLM [13]	29.4113	74.4654	0.9109	0.9104	30.1086	63.4192	0.9293	0.9137
	ER-PCA	29.6008	71.2860	0.8723	0.8785	29.6002	71.2947	0.8647	0.8527
	KPCA	30.1448	62.8934	**0.9129**	0.9144	30.6031	56.5941	**0.9316**	0.9245
25	Noisy	18.5384	910.4194	0.3095	0.6362	18.3487	951.0736	0.2774	0.5520
	UNLM [11]	28.6394	88.9483	0.8777	0.9012	29.1108	79.7989	0.8914	0.8987
	BM3D+VST [7]	**29.0589**	**80.7598**	0.8109	**0.9114**	**29.2912**	**76.5527**	0.8269	**0.9227**
	RS-NLM [13]	26.4734	146.4670	0.8599	0.8492	26.6486	140.6762	0.8696	0.8372
	ER-PCA	28.0567	101.7219	0.8576	0.8824	28.1251	100.1314	0.8529	0.8685
	KPCA	28.1995	98.4298	**0.8792**	0.8814	28.3402	95.2919	**0.8932**	0.8785

have been carried out on 2D monochrome phantom human brain MRI images obtained from Brain Web Database [3]. The parameters are as follows: RF = 20, protocol = ICBM, slice thickness = 1 mm, volume size = 181 × 217 × 181. The experimental set up considers Rician noise model at different noise levels along with two modalities, namely T1 and T2. The simulated database provides the ground truth image for evaluating denoising performance which most of the time is unavailable with real database. The Rician noise addition and bias correction are done as suggested in [10] and [11] respectively. The evaluation measures used are Peak-Signal-to-Noise Ratio (PSNR), Root Mean Square

Error (RMSE), Mean Structural Similarity Index (MSSIM) [17] and Feature Similarity Index (FSIM) [18].

For comparison purpose, several state-of-the-art methods are considered: Unbiased Non Local Means (UNLM method) presented in [11], BM3D+VST method proposed in [4], Rough Set based Non Local Means (RS-NLM) method proposed in [13] and PCA based method proposed in the [13] has been extended in this work for Rician noise, referred as Extended Rough set based PCA method (ER-PCA). The parameters of all methods are kept default as suggested by respective authors. In all the experiments, patch size is kept as 5×5. The proposed KPCA method does not use VST method. Tables 1 and 2 represent quantitative results for two slices 70 and 100 of T1 MR and T2 MR images respectively. The ER-PCA performance is comparable to UNLM and BM3D+VST methods. The proposed KPCA method outperforms ER-PCA and preserves structure better than other state-of-the-art method. Figure 3 shows difference of PSNR and MSSIM measure for KPCA method with reference to BM3D+PCA (zero level on vertical axis) of 50 slices (from 61^{st} to 110^{th} slice of database mentioned above) with noise standard deviation equal to 15 for both T1 and T2 modalities. Negative value indicates BM3D+VST performs better and, in reverse, positive value is indicator of better performance of KPCA method. From Fig. 3, PSNR of KPCA fall below BM3D+VST method whereas it better preserves structure of the image in terms of MSSIM measure. This is also visually evident in Fig. 4 for the slice 100 of T1 modality at noise level 15.

4 Conclusion

In this paper, an approach for removal of Rician noise from brain MR images using Kernel PCA has been proposed. Being a manifold learning method, KPCA explores a suitable transformation for image representation through sparse bases. This method learns basis vectors from data itself unlike BM3D+VST method where basis vectors are kept fixed. The limitation of KPCA method is the selection of suitable kernel which is yet unanswered. If the nature of data is not known a-prior than one can try various kernels to find a suitable one. However, commonly used Gaussian kernel in KPCA, found to perform comparable with other state-of-the-art methods. The PCA based method proposed in [13] has also been implemented to remove Rician noise, but it fails to attain superior performance over KPCA. The proposed method is implemented on synthetic data for quantitative evaluation since ground truth data is available for the same.

References

1. Buades, A., Coll, B., Morel, J.M.: A non local algorithm for image denoising. In: IEEE Computer Vision and Pattern Recognition, pp. 60–65 (2005)
2. Charpiat, G., Hofmann, M., Schölkopf, B., et al.: Kernel methods in medical imaging. Handbook of Biomedical Imaging (2010)

3. Collins, D.L., Zijdenbos, A., Kollokian, V., Sled, J., Kabani, N., Holmes, C., Evans, A.: Design and construction of a realistic digital brain phantom. IEEE Trans. Med. Imaging **17**(3), 463–468 (1998)
4. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3d transform domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007)
5. Dong, W., Li, X., Zhang, D., Shi, G.: Sparsity-based image denoising via dictionary learning and structural clustering. In: 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 457–464. IEEE (2011)
6. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. **15**(12), 3736–3745 (2006)
7. Foi, A., Noise estimation and removal in mr imaging: the variance-stabilization approach. In: ISBI, pp. 1809–1814 (2011)
8. Gudbjartsson, H., Patz, S.: The rician distribution of noisy mri data. Magn. Reson. Med. **34**(6), 910–914 (1995)
9. He, L., Greenshields, I.R.: A nonlocal maximum likelihood estimation method for rician noise reduction in mr images. IEEE Trans. Med. Imaging **28**(2), 165–172 (2009)
10. Maggioni, M., Katkovnik, V., Egiazarian, K., Foi, A.: Nonlocal transform domain filter for volumetric data denoising and reconstruction. IEEE Trans. Image Process. **22**(1), 119–133 (2013)
11. Manjon, J.V., Caballero, J.C., Marti, G.G., Marti-Baonmati, L., Robles, M.: Mri denoising using non local means. Med. Image Anal. **12**, 514–523 (2008)
12. McVeigh, E., Henkelman, R., Bronskill, M.: Noise and filtration in magnetic resonance imaging. Med. Phys. **12**, 586 (1985)
13. Phophalia, A., Rajwade, A., Mitra, S.K.: Rough set based image denoising for brain mr images. Sig. Process. **103**, 24–35 (2014)
14. Rajwade, A., Rangarajan, A., Banerjee, A.: Image denoising using the higher singular value decomposition. IEEE Trans Pattern Ana. Mach. Intell. **35**(4), 849–862 (2013)
15. Rubio, E.L., Nunez, M.N.F.: Kernel regression based feature extraction for 3d mr image denoising. Med. Image Anal. **15**, 498–513 (2011)
16. Sijbers, J., den Dekker, A.J., Scheunders, P., Dyck, D.V.: Maximum likelihood estimation of rician distribution parameters. IEEE Trans. Med. Imaging **17**(3), 357–361 (1998)
17. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process. **13**(4), 600–612 (2004)
18. Zhang, L., Zhang, L., Mou, X., Zhang, D.: Fsim: A feature similarity index for image quality assessment. IEEE Trans. Image Process. **20**(9), 2378–2386 (2011)
19. Zimmer, V.A., Lekadir, K., Hoogendoorn, C., Frangi, A.F., Piella, G.: A framework for optimal kernel-based manifold embedding of medical image data. Comput. Med. Imaging Graph. **41**, 93–107 (2015)