



A new method for Rician noise rejection in sparse representation

Un nuevo método para el rechazo de ruido Rician en representación escasa

Morera-Delfin Leandro

Centro Nacional de Investigación y Desarrollo Tecnológico,
CENIDET, Morelos

E-mail: leomorerad@gmail.com

<https://orcid.org/0000-0002-9646-7179>

Abstract

One of the factors that affects the quality of medical images is noise in the acquisition process. Rician noise, for example, is present in MRI images and causes errors in measurements and interpretations of visual information. The objective of this work is to obtain high indexes of structural similarity (SSIM) and peak signal-to-noise ratio (PSNR) through digital filtering of Rician noise. In the design, clusters of low coefficients are used to eliminate information redundancies, the probability density function (pdf) of the Rician noise to estimate signal levels and minimization by conjugate gradient to achieve a greater approximation to the real signal. The model is applied by filtering longitudinal sequences of MRI studies at T2 acquisition time affected with Rician noise in a controlled manner. Different models of noise filtering were implemented and tested on the same test sequence. The proposed method achieves an iterative approach to the real image. As a result, the SSIM and PSNR parameters improve in a magnitude of 0.02 and 0.3dB over the estimate with Gaussian pdf. The system has as limiting the effectiveness of the estimation for high signal levels due to the increase of the standard deviation in the pdf of the Rician noise in the aforementioned levels, however it manages to surpass the performance of current models within the state of the art. The proposed model has the novelty of linking the grouping of coefficients and the estimation by means of the pdf of the Rician noise. The system helps to avoid errors in measurements and interpretations of data affected by Rician noise, in particular in MRI studies. It is concluded that the pdf of Rician noise behaves like a good estimator in a digital filtering model with grouping of coefficients. Despite having better performance for medium and low signal levels, the proposed system manages to overcome the results obtained by other filtering models described within the state of the art.

Keywords: Filtering, Rician noise, sparse representation, clustering, probability.

Resumen

Uno de los factores que afecta la calidad de las imágenes médicas es el ruido en el proceso de adquisición. El ruido Rician, por ejemplo, está presente en las imágenes de MRI y provoca errores en mediciones e interpretaciones de la información visual. El presente trabajo tiene como objetivo obtener elevados índices de similitud estructural (SSIM) y relación pico de señal sobre ruido (PSNR) mediante filtrado digital del ruido Rician. En el diseño se utilizan agrupamientos de coeficientes escasos para eliminar redundancias de información, la función de densidad de probabilidad (fdp) del ruido Rician para estimar niveles de señal y minimización mediante gradient conjugado para lograr mayor aproximación a la señal real. El modelo se aplica filtrando secuencias longitudinales de estudios MRI en tiempo de adquisición T2 afectadas con ruido Rician en forma controlada. Se implementaron diferentes modelos de filtrado de ruido y se probaron sobre la misma secuencia de prueba. El método propuesto logra una aproximación iterativa a la imagen real. Como resultado, los parámetros SSIM y PSNR mejoran en una magnitud de 0.02 y 0.3dB sobre la estimación con fdp gaussiana. El Sistema tiene como limitante la efectividad de la estimación para altos niveles de señal debido al aumento de la desviación típica en la fdp del ruido Rician en los niveles mencionados, sin embargo, logra superar el desempeño de modelos actuales dentro del estado del arte. El modelo que se propone tiene la novedad de vincular el agrupamiento de coeficientes y la estimación mediante la fdp del ruido Rician. El sistema contribuye a evitar errores en mediciones e interpretaciones de datos afectados por ruido Rician, en particular, en estudios de MRI. Se concluye que la fdp del ruido Rician se comporta como un buen estimador en un modelo de filtrado digital con agrupación de coeficientes. A pesar de tener mejor desempeño para niveles medios y bajos de señal, el sistema propuesto logra superar los resultados obtenidos por otros modelos de filtrado descritos dentro del estado del arte.

Descriptores: Filtrado, ruido Rician, representación escasa, agrupamiento, probabilidad.

INTRODUCTION

In Magnetic Resonance Imaging (MRI), the denoising process is an important step because it impacts the diagnostic accuracy. In many cases the signal-to-noise-ratio (SNR) of the acquired images must be improved. This is carried out by using denoising methods and by preserving most of the image features. In MRI, the images are contaminated with Rician noise with a Probability Density Function (PDF) that depends on the level of the original signal. The problem of noise filtering in MRI has been addressed in different works (Pizurica, 2003; Casaca, 2010; Pu, 2016; Weaver, 1991). In the signal processing literature, several popular denoising algorithms are based on thresholds for coefficients of some transformation (Sendur, 2002). These approaches attempt to preserve significant features of the image while removing most of the noise.

If the wavelet transform is applied on the MR magnitude data directly, both the wavelet and the scaling noisy coefficients are biased estimates of the noise-free coefficients. In Nowak (1999), it was suggested that the application of the wavelet transform on the squared MR magnitude image data (which is non-central chi-square distributed) would result in wavelet coefficients being no longer biased estimates of their noise-free counterparts. Although the bias still remains in the scaling coefficients, in many models it is not signal-dependent and the noise can, therefore, be easily removed (Pizurica, 2003; Nowak, 1999). Similarly, denoising methods based on anisotropic diffusion have been proposed (Perona and Malik, 1990; Samsonov, 2004) The difficulty with wavelet or anisotropic diffusion algorithms is the risk of over-smoothing the details, particularly in low SNR images (Tisdall, 2005).

Also other strategies have been proposed. In He (2009), the estimation of the noise-free coefficients of the MR's magnitude was carried out in a non-local mode using a log-likelihood function. The images are modeled as random fields where pixels with similar neighborhoods come from the same distribution. In, Sijbers (1998) the authors showed that the Maximum-Likelihood (ML) estimation, which is known to yield optimal results asymptotically, enhances the conventional estimation methods. The proposed method is unbiased for high signal-to-noise ratios (SNRs) and yields physical relevant results for low SNRs.

The concepts of non-local means and dictionary reconstruction have been used in Nair (2014); Aarya (2013); Elad (2006). In Nair (2014), a Non-Local Means Maximum Likelihood (NLML) estimation method was used to estimate the noise-free coefficients from the

MR's magnitude images by focusing on preserving edges and tissue boundaries. The method is an improvisation over non-local means maximum likelihood approach for Rician noise reduction in MR images. In Aarya (2013), an adaptive filter for Rician noise based on the probability distribution function of the noise and the SNR information of the image was proposed. The filter uses the local statistics of the neighborhood within a mask to perform denoising. Thus, the filter adapts according to the local SNR of the neighborhood. In Elad (2006), two training options are considered: using the corrupted image itself, or using a high-quality image database. The K Singular Values Decomposition (KSVD) was used to design an adaptive dictionary trained for natural real images, as well as on patches of the noisy image were also used.

Our method combines two important features, the Non-Local Maximum Likelihood (NLML) estimation and denoising via sparse and redundant representation in an unique process.

The KSVD method is introduced to construct an adaptive dictionary on the noisy image. The patches are grouped using the K-means algorithm and the Principal Components Analysis (PCA) to associate the characteristics (Annexed A). Afterward, the Rician distribution guides each group by using local statistics to perform the denoising process in a Maximum A Posteriori (MAP) formulation. As a restriction over the space of search we estimate the noise coefficients over a residual image resulting from the noise image and an initial filtered image.

The problem of recovering a high-quality image from one or several degraded (e.g, noisy, blurred, and/or down-sampled) versions can be generally formulated by:

$$y = Hx + n \tag{1}$$

Where:

- y = available distorted image
- H = degradation matrix
- x = original image vector and
- n = noise vector

H can be the identity matrix, a blurring operator or a blurring and down-sampling operator. The studies have been conducted to restore x from y , (Yang, 2010; Sendur, 2002). For an effective approach, the appropriate prior knowledge of the natural images must used (Rudin, 1992), similarly to the Tikhonov (1963) and Total Variation (TV) regularization methods (Rudin, 1992). However, these methods tend to over-smooth the

restored images due to the rigid rules of a piecewise function. An alternative is the use of sparsity-based regularization (Sendur, 2002). The model assumes that a signal $x \in \mathbb{R}^N$ can be represented as $x \approx \phi\alpha$, where $\phi \in \mathbb{R}^{N \times M}$ ($N < M$) is an over-complete dictionary and most entries of the coding vector α are zero or close to zero. The l_1 -norm based sparse coding problem is generally formulated in the following Lagrangian form:

$$\alpha_x = \arg \min_{\alpha} (\|x - \phi\alpha\|_2^2 + \lambda \|\alpha\|_1) \quad (2)$$

where the constant λ denotes the regularization parameter. With an appropriate selection of λ , we can obtain a good approximation error and an appropriate sparsity of α . In the scenario of Image Restoration (IR), what we observed is the degraded image signal y . To recover x , first y is sparsely coded with respect to ϕ for solving the minimization

$$\alpha_y = \arg \min_{\alpha} (\|y - H\phi\alpha\|_2^2 + \lambda \|\alpha\|_1) \quad (3)$$

Therefore, y is reconstructed by $\hat{y} = \phi\alpha_y$. The local sparsity constraint $\|\alpha\|_1$ in Eq. 3 may not lead to an accurate image reconstruction because image sparse coding coefficients α are not randomly distributed. It is due to the internal correlations or redundancies in natural images, (Casaca, 2010; Pu, 2014). In Pu (2014), a group sparse coding scheme to code similar patches simultaneously was proposed achieving impressive denoising results. The KSVD method reported in Aharon (2006) bring better results than the cosine transform (DCT) in the dictionary construction.

In our work, the Probability Density Function (PDF) of the observed image is linked with an initial noise estimation. The coefficients of the residual noise decomposition are used as a linker between the noisy and the reconstructed image. These two aspects are strong considerations to adapt the denoising method to the *Rician* taking advantage of an initial estimation, the processing of the residual noise in the domain of the sparse representation and the MAP formulation.

In this paper, the *Rician* noise distribution is analyzed. The noise reduction is based on the estimation of coefficients of the unknown real values of the signal. For fast processing and accuracy, the process is carried out with a sparse representation of the image. Similar patches of the image are grouped in clusters and treated independently. The main contribution of this work are:

- A new *Rician* denoising procedure in a K-SVD sparse representation domain using statistical relations.

- A direct prediction of the sparse coefficients of the signal after an iterative process with initial non-local estimation of the noise using residuals coefficients.
- A representation of the signal with clustering and dictionary reconstruction in the *Rician* denoise procedure.

The paper is organized as follows: Describes the patches decomposition of the signal and the sparse representation; exposes the treatment of the noise in the sparse representation and the PDF model of the noisy image; explains the proposed denoising method; discusses the results; and draws the conclusions.

SPARSE REPRESENTATION AND RICIAN MODEL

In the proposed model, the input image y is first filtered by using a low pass filter with output y_f . Then, the residual image, considering the initial noise approximation, is obtained by $y_r = y - y_f$. The patch at location i is extracted from an image of the form $x_i = R_i x$. The R_i matrix extracts the patch x_i of dimension $\sqrt{n} \times \sqrt{n}$. Given a dictionary $\phi \in \mathbb{R}^{N \times M}$ ($N < M$). Each patch can be sparsely represented as $x_i = \phi_i \alpha_{x_i}$ by solving an l_1 minimization problem $\alpha_{x_i} = \arg \min_{\alpha} (\|x_i - \phi_i \alpha\|_2^2 + \lambda \|\alpha\|_1)$, the image x can be represented by the set of sparse codes α_{x_i} with overlapped patches. A redundant patch-based representation of x is obtained by α_x as an over-determined system (Elad, 2006). The Eq. 4 yields an overall image reconstructed by averaging each patch.

$$x \approx \phi \alpha_x = \left(\sum_{i=1}^N R_i^T R_i \right)^{-1} \sum_{i=1}^N (R_i^T \phi \alpha_{x_i}) \quad (4)$$

NOISE IN THE SPARSE REPRESENTATION DOMAIN

The clean and the noisy image can be decomposed as a dictionary and the coefficients as the procedure developed in (Aharon, 2006). In a real situation, only the coefficients of the residual and noisy image can be acquired. The coefficients of the residual image exhibit similar characteristics to the *Rician* distribution. We propose a Gaussian filter with variance σ to obtain the residual image y_r . The dictionaries are learned from the example image patches by using the KSVD (Aharon, 2006). However, it has been shown that sparse coding with an over-complete dictionary is unstable (Elad, 2009). Especially, in the scenario of image restoration. A solution to this is to select k clusters with similar patches and then learn a KSVD sub-dictionary for each one. The similar patches are grouped in consecutive

rows of clusters in the matrix P_k . This is carried out on the noisy image, the filtered and the residual images.

In the next step, a β factor in the sparse plane is estimated from the residual image decomposition and then subtracted from the coefficients of the noisy image decomposition.

$$\alpha_y - \beta \approx \alpha_x \tag{5}$$

The residual image y_r is characterized by the previous knowledge of the *Rician* distribution. With a good estimation of β in y_r , the term $\hat{\alpha}_y \approx \alpha_x$ can be used instead of α_x in the Lagrangian model of reconstruction. Then, the coefficients of the clean image can be more approximated by using a minimization of Eq. 2, wherein a negative gradient method is proposed.

RICIAN NOISE DISTRIBUTION

If the real and imaginary parts of a signal are corrupted with zero-mean, equal variance and uncorrelated Gaussian noise, the envelope of the magnitude signal will follow a *Rician* distribution. This kind of noise is apparent in many practical situations. The signal magnitude M is expressed as:

$$M = \sqrt{(A + n_1)^2 + n_2^2} \tag{6}$$

where A is the clean signal level, n_1 and n_2 are uncorrelated Gaussian noise variables with zero mean and equal variance. The PDF of such image is a *Rician* distribution is expressed as:

$$P_M(M | A, \sigma_n) = \frac{M}{\sigma_n^2} \exp\left(-\frac{M^2 + A^2}{2\sigma_n^2}\right) I_0\left(\frac{AM}{\sigma_n^2}\right) u(M) \tag{7}$$

with I_0 the 0th-order modified Bessel function of the first kind and u the Heaviside step function. The moments of the distribution are difficult to calculate, the even-order (non-central) moments are simple polynomials (e.g. second-order moment). The signal with Rician noise is modeled with a probability density function (PDF) that changes the aperture with respect to the real level of signal A .

$$P(\alpha_y) \approx P(\alpha_x, \beta) \approx \prod_i \prod_j \frac{\alpha_{y(i,j)}}{\sigma_{n(i,j)}^2} \exp\left(-\frac{\alpha_{y(i,j)}^2 + (\alpha_{y(i,j)} - \beta_{i,j})^2}{2\sigma_{n(i,j)}^2}\right) I_0\left(\frac{(\alpha_{y(i,j)} - \beta_{i,j})\alpha_{y(i,j)}}{\sigma_{n(i,j)}^2}\right) \tag{13}$$

PROPOSED METHOD

The connection between the MAP estimator and sparse representation was established in (Sendur, 2002). In our model, the term $\theta = \alpha_y - \beta$ is defined. Using Eq. 5, for a given β , the MAP estimation of $\theta_{\alpha_y} \approx \alpha_x$ can be formulated as:

$$\theta_{\alpha_y} = \arg \max_{\theta} (\log(P(\theta | \alpha_y))) \tag{8}$$

$$\arg \max(\log(P(\theta | \alpha_y))) = \arg \max(\log(\alpha_y | \theta)) + \log(P(\theta)) \tag{9}$$

According to Eq. 3 and knowing that $P(\alpha_x, \beta | \alpha_y)P(\alpha_y) = P(\alpha_y | \alpha_x, \beta)P(\alpha_x, \beta)$, considering that for little values of β , $P(\alpha_x, \beta) = P(\alpha_y)$ and $P(\alpha_x, \beta | \alpha_y) = P(\alpha_y | \alpha_x, \beta)$ then the prior knowledge with a Gaussian distribution can be modeled as $P(\theta | \alpha_y)$.

$$P(\alpha_x, \beta | \alpha_y) = P(\alpha_y | \alpha_x, \beta) = \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\frac{1}{2\sigma_n^2} \|\alpha_y - (\alpha_x + \beta)\|_2^2\right) \tag{10}$$

The likelihood term is characterized by the Gaussian distribution, θ and β are assumed to be independent. In the prior probability $P(\theta | \alpha_y)$, θ reflects the variation of α_y from the estimation of β . If we take β as a good estimation, then $\theta_y = \alpha_x + \beta$ are basically the coefficients α_y of the noisy image. Thus, we can assume that θ follows an independently and identically distributed (*i.i.d.*) process, and the joint prior distribution of the image to be the *Rician* noise $P(\theta)$. The estimated image can be modeled using Eq. 7 for small values of β .

The MAP approximation considers the coefficients of the unknown clean image x in Eq. 10 and the *Rician* distribution considers α_y for modeling the observed image using the approximation of Eq. 5. Therefore, the proposed MAP formulation is:

$$\arg \max_{\theta} (\log(P(\alpha_x | \alpha_y))) = \arg \max_{\theta} (\log(P(\alpha_y | \alpha_x)) + \log(P(\alpha_x))) \tag{11}$$

The approximation to the MAP model of Eq. 10 is a classical Gaussian function. Then, we can use Eqs. 7 and 8 to approximate to $P(\alpha_y | \alpha_x, \beta)$ in the kind of *Rician* noise and the corrupted image can be modeled as:

$$P(\alpha_y | \alpha_x, \beta) = \frac{\alpha_y}{\sigma_n^2} \exp\left(-\frac{\alpha_y^2 + (\alpha_y - \beta)^2}{2\sigma_n^2}\right) I_0\left(\frac{(\alpha_y - \beta)\alpha_y}{\sigma_n^2}\right) \tag{12}$$

$$\hat{\alpha}_x = \sum_{i=1}^k \sum_{j=1}^p \log\left(\frac{\alpha_{y(i,j)}}{\sigma_n^2(i,j)}\right) I_0\left(\frac{(\alpha_{y(i,j)} - \beta_{i,j})\alpha_{y(i,j)}}{\sigma_n^2(i,j)}\right) - \sum_{i=1}^k \sum_{j=1}^p \frac{(\alpha_{y(i,j)}^2 + (\alpha_{y(i,j)} - \beta_{i,j})^2)}{\sigma_n^2(i,j)} \quad (14)$$

If the predicted values of β are considered independent of y , the term $P(\alpha_x, \beta)$ can be approximated using Eq. 5 in a discrete form as equation 13.

Where are the coefficients of the noisy image y , and β the estimate of the noise. Then, a good estimation of α_{y_i} and β_{i_j} can be computed as the weighted average of the sparse codes associated with the non-local similar patches (including patch i) to patch. For each patch x_{y_i} we calculate a set of similar patches, denoted by Ω_i . Then, Eq. 12 is applied for each cluster Ω_i . The problem is how to obtain a well approximation $\beta_{i,j}$ at position j of the group i . Applying logarithms to the MAP formulation in the corresponding localization at the cluster Ω_i , a good estimation of β maximizes the probability in the function 8 respect to α_x , a discrete formulation of the approximation using the Eq. 9 and 5 is eq. 14.

DETERMINATION OF β

The parameter β_i is determined on each cluster of $\alpha_{y_{ri}}$ β can be computed from the sparse codes of the residual output split in patches, the Eq. 14 is used to estimate α_x . The term ρ denotes the sparse codes related to the estimation of β as:

$$\varphi\alpha_{y_r} = \varphi\alpha_y - G(\varphi\alpha_y) \quad (15)$$

Where $G(\cdot)$ is a low pass filter operator and ρ is the sparse representation of y_r , ρ is characterized by the Rician distribution, then $y_r = \varphi\rho$.

$$\beta_i = \sum_{\rho \in \Omega_i} \omega_{i,k} \rho_{i,k} \quad (16)$$

where ω_{ik} is a weight. Similar to the non-local means approach, we set the weights to be inversely proportional to the distance between the coefficients ρ_i of the patch y_{ri} and the coefficients $\rho_{i,k}$ of the patches $y_{ri,k}$ in the k^{th} cluster of patches.

$$\hat{\alpha}_x^l = \frac{d}{d\beta} \left(\sum_{i=1}^k \sum_{j=1}^p \log\left(\frac{\alpha_{y(i,j)}}{\sigma_n^2(i,j)}\right) I_0\left(\frac{(\alpha_{y(i,j)} - \beta_{i,j})\alpha_{y(i,j)}}{\sigma_n^2(i,j)}\right) - \sum_{i=1}^k \sum_{j=1}^p \frac{(\alpha_{y(i,j)}^2 + (\alpha_{y(i,j)} - \beta_{i,j})^2)}{\sigma_n^2(i,j)} \right) \quad (18)$$

$$\omega_{i,k} = \frac{1}{w} \exp\left(-\frac{\|\hat{\rho}_i - \hat{\rho}_{i,k}\|_2^2}{h}\right) \quad (17)$$

$\alpha_{y_{ri}} = \varphi\hat{\rho}_i$ and $\alpha_{y_{ri,k}} = \varphi\hat{\rho}_{i,k}$ are the estimates of the patches ρ_i and $\rho_{i,k}$, h is a predetermined scalar and w is the normalization factor. With the non-local redundancies of natural images, we are able to achieve a good estimation of the unknown sparse vectors x_i with a good estimate of β , in the model of Eq. 14.

STOP CRITERION

The image estimation process is repeated until an approximation to the solution with error ξ is reached. The estimate of Eq. 14 is minimized with respect to β . In this case, we propose a numerical solution as the negative gradient. The negative differential of Eq. 14 is

$$-D_{\alpha_x} = \hat{\alpha}_x^l - \hat{\alpha}_x^{l-1} \quad x \quad x$$

Then a minimum differential is found numerically by using the conjugated gradient (Eq. 18).

$$|\hat{\alpha}_x^l - \hat{\alpha}_x^{l-1}| < \xi \quad (19)$$

l is the number of iterations $l = 1, \dots, l_{ite}$. When $l = 1$, β is calculated using the residual y_r and a new assignment is made. The estimated image is assigned: $\hat{\alpha}_y = \hat{\alpha}_x^l$.

THE ALGORITHM

1. Apply a filtering procedure to the noisy image y to obtain y_r using $y_r = y - G(y)$.
2. Provide a patch representation of the image using the extraction matrix R in the form of $y_i = R_i y$ and setting k groups using KMEANS and PCA classification.

3. Apply the KSVD algorithm and build the dictionary representation corresponding to y_r, y and obtain α_{y_r} and α_y in the sense $y_r = \varphi\alpha_{y_r}$ and $y = \varphi\alpha_y$ for each patch of y .
4. Calculate $\omega_{i,k}$ using Eq. 17 and obtain an initial estimation of β using Eq. 16.
5. Calculate the coefficients $\hat{\alpha}_x$ by using Eq. 14.
6. Restore the estimate image $\hat{x} = \varphi\hat{\alpha}_x$.
7. Determine the new residual $\hat{y}_r = \varphi\alpha_y - \varphi\hat{\alpha}_x$.
8. Assign \hat{x} to \hat{y} .
9. Repeat the process with the new residual as \hat{y}_r until the stop criterion is met.

RESULTS

The proposed method was applied on different MRI studies with relaxation time T2. The images are normalized to 256 gray levels. The Tables 1 and 2 give quantitative results of the comparison of our method respect to other referenced procedures. The quality factors considered to evaluate the results were the ratio of signal

peak over noise (PSNR) and the ratio of structural similarity (SSIM). In both cases, the proposed procedure is the most effective filtering the Rician noise with $\sigma = 30$ in Table 1 and $\sigma = 40$ in Table 2. Whith $\sigma = 30$ our method enhance the PSNR in 0.3 dB and SSIM in 0.01 as average respect to the model NCSR. Whith $\sigma = 40$ our method enhance the PSNR in 0.1 dB and SSIM in 0.1 as average respect to the model NCSR.

COMPARISONS

The first column of the Figure 1 contain the visual comparison of the result for the slice 3 in the Table 1. The Figure 2 shows the results of detail reproduction of our method with respect to other procedures that produces homogeneous areas. In the Figure 2c the procedure tends to homogenize the areas and eliminate fine detail. It is a model based in threshold over coefficients and dependences between decomposition bands. The Figure 2e shows a model that provides best results in the details reproduction. It employs also clustering techni-

Table 1. PSNR/SSIM comparison of the proposed model NLRD versus other methods with $\sigma = 30$

Method	Slice				
	1	2	3	4	5
Bilateral filter (Tomasi, 1998)	17.8492/0.2821	17.9380/0.2888	18.0132/0.2971	17.8524/0.3032	17.8980/0.3133
Wave atoms (Shashi, 2014)	18.8624/0.3208	18.9041/0.3459	19.3082/0.3649	19.1017/0.3638	19.0240/0.3486
Adaptive wavelet (Chang, 2000)	19.9755/0.5789	19.9251/0.5683	19.9551/0.5687	19.8701/0.5672	19.8772/0.5701
NLM (Wiest, 2008)	20.4395/0.3792	20.7070/0.3792	20.2079/0.3925	19.3309/0.3795	19.5331/0.3954
BM3D (Elahi, 2014)	21.6015/0.5942	21.5754/0.6102	21.4629/0.6260	21.7245/0.6326	21.959/0.6345
Wavelet LMS (Sathyanarayana, 2011)	22.6426/0.6596	22.3830/0.6284	22.4417/0.6358	22.5646/0.6471	22.7128/0.6713
ASDS' AR (W.-Dong, 2011)	21.0595/0.6285	21.1247/0.6303	21.0243/0.6397	20.2612/0.6596	19.6958/0.6509
ASDS (W.-Dong, 2011)	21.1890/0.6295	21.2986/0.6317	21.3359/0.6351	20.5224/0.6649	19.9718/0.6592
ASDS AR NL (W.-Dong, 2011)	20.7290/0.6135	20.5857/0.6200	20.1921/0.6305	19.4643/0.6304	18.9621/0.6228
NCSR (W.-Dong, 2013)	23.0987/0.6700	23.0133/0.6766	23.2129/0.7179	23.2232/0.7091	23.1175/0.7056
NLRD	23.4088/0.6919	23.2431/0.6821	23.4217/0.7153	23.5105/0.7159	23.4046/0.713

Note: Bold values indicate the best result NLRD

Table 2. PSNR/SSIM comparison of the proposed model NLRD versus other methods with $\sigma = 40$

Method	Slice				
	1	2	3	4	5
Bilateral filter (Tomasi, 1998)	16.9089/0.2466	16.9384/0.2525	17.0182/0.2499	16.9395/0.2621	16.9933/0.2703
Wave atoms (Shashi, 2014)	19.5205/0.3288	19.4310/0.3440	19.4958/0.3495	19.3397/0.3488	19.4218/0.3510
Adaptive wavelet (Chang, 2000)	18.2653/0.4777	18.2856/0.4982	18.3301/0.4836	18.2613/0.4903	18.3956/0.4864
NLM (Wiest, 2008)	19.5119/0.3258	19.5334/0.3494	19.1726/0.3384	19.0765/0.3435	19.1391/0.3327
BM3D (Elahi, 2014)	19.5697/0.5412	19.5839/0.5709	19.4429/0.5718	19.4253/0.5842	19.4549/0.5908
Wavelet LMS (Sathyanarayana, 2011)	19.2992/0.5837	19.1156/0.5677	19.0800/0.5625	19.0805/0.5709	18.8920/0.5673
ASDS' AR (W.-Dong, 2011)	19.7954/0.5380	19.8492/0.5491	19.8967/0.5740	19.9407/0.5951	20.0649/0.6025
ASDS (W.-Dong, 2011)	19.8506/0.5365	19.9209/0.5482	19.9733/0.5719	20.0225/0.5929	20.2000/0.6016
ASDS AR NL (W.-Dong, 2011)	19.5744/0.5404	19.5763/0.5476	19.6133/0.5767	19.5824/0.5937	19.5985/0.5958
NCSR (W.-Dong, 2013)	19.9763/0.5588	19.7573/0.5481	20.0824/0.5716	20.1234/0.5794	20.0236/0.5727
NLRD	20.0926/0.6019	19.8735/0.6257	20.1663/0.6309	20.2378/0.6262	20.2901/0.6463

Note: Bold values indicate the best result NLRD

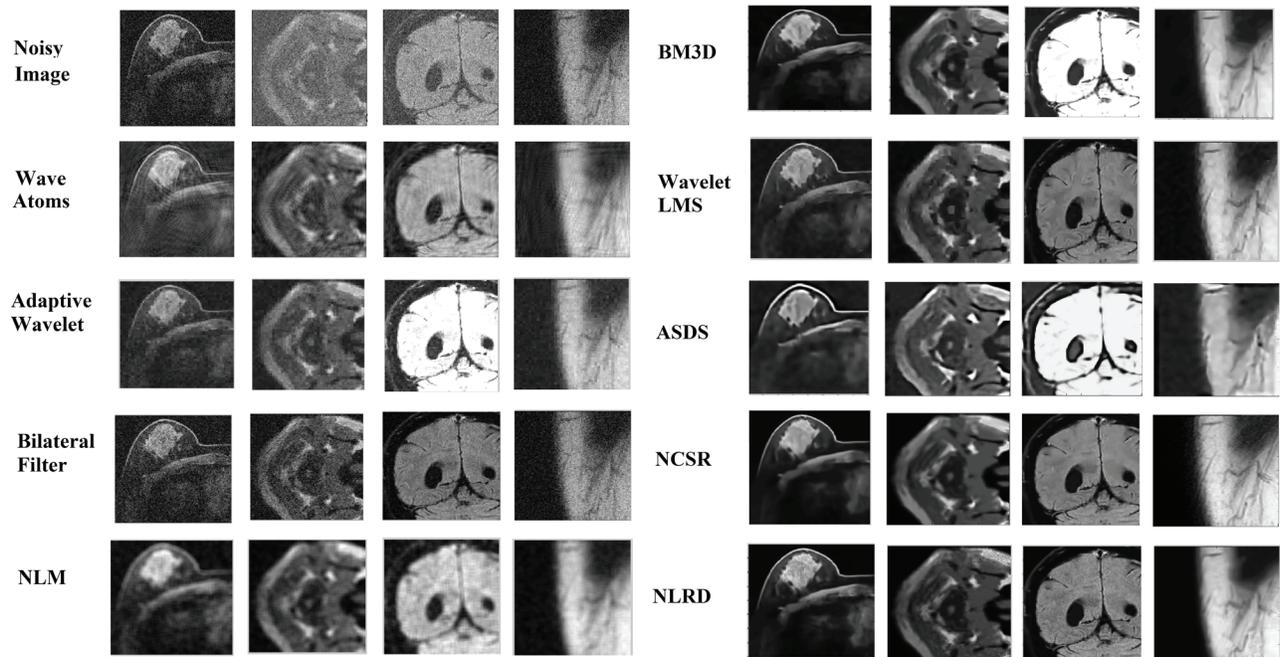


Figure 1. Visual comparison of results in Table 1, second column left to right

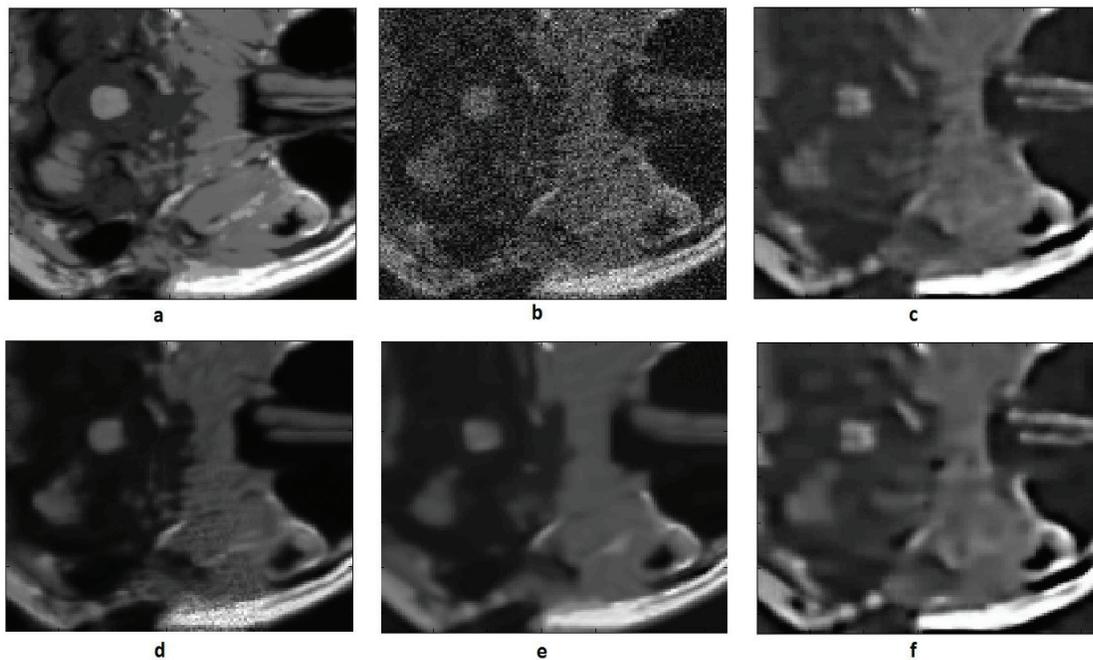


Figure 2. Visual comparison of results in Table 1, a) Original Image, b) Noisy image, c) Wavelet LMS method, d) Proposed NLRD method, e) NCSR method, f) ASDS method $\sigma = 40$

ques but the probabilistic inference is made with a fixed Gaussian function. The results of the Figure 2f lost fine details and the reconstruction tends to expand the structures. It is a model in which the group of coefficients are selected with an adaptive procedure for the estimation. In the case of the Figure 2d the result is a more detailed reconstruction of fine structures in areas from low to median-high levels of signal and in general a reconstruction more near to the original image. The estimation using a *Rician* noise PDF that change the width proportionally to the level of the signal makes more difficult the estimation with wide PDF. It is the reason for the presence of noise in high levels of signal. This effect could be avoided using equalization, normalizing or changing the ranges of the image's levels, this technique could be treated in future work.

In the proposed method, the approach of the exhaustive search and clustering is realized over α_{yr} . This give an increment of the SSIM and PSNR values over all the slices analyzed. The method is robust for high levels of noise, in our comparison it reach $\sigma = 40$. The results obtained in the Tables 1 and 2 shows an enhancement of the denoising techniques almost in chronological sense in relation with the publication of the concepts and procedures. The most recent methods use techniques for grouping and reduction of the information to be processed as in ASDS, NCSR and our proposed model Non Local Rician Denoise (NLRD).

The models under comparison use different techniques for noise treatment. The wavelet denoising method uses the background estimation for the level of noise and clusters similar coefficients. Then in the groups, the level of noise is subtracted. In this case, more fine bands are more affected by noise and the upper bands are assumed the desired signal. The PDF in the *Rician* case combines signal and noise in the model shown in the Eq. 7. Then, the MAP formulation 9 gives the possibility of $\hat{\alpha}_x$ estimation when a relationship is established between the unknown original coefficients α_x and the noisy coefficients α_y through of β . The MAP estimation is carried out over the sparse domain of clusters of patches of the image. The separation of the image in different groups of patches allows to carry out a precise approximation of β and the $\hat{\alpha}_x$ coefficients.

CONCLUSION

The proposed method combines the PDF of the Rician noise and the redundancies of the structures of the image for denoising. The spatial redundancy of the patches is exploited using the PCA classifier and the KSVD dictionary learning technique. The noise estimation was

carried out with a desired precision after some iterations. The Rician noise distribution is analyzed on the image decomposed in sub-regions and clusters. The mean of the group of patches is an important consideration because the Rician distribution changes the width according to the level of the real signal. Due to the width changes of the PDF the model must enhance the response for high levels of signal, despite of this the general result bring a good detail reproduction for the areas of the image. The utility of the sparse representation of the image to characterize the noise distribution was shown. For the estimation of local values, the redundancies around all the residual image are exploited. The method shows that the PDF of the Rician noise can be used in a MAP formulation using an estimation of the residual image as a link between the noisy and the estimated image. The levels of the noisy and estimated images are presents in the PDF. The methods of sparse representation, clustering and noise estimation selected for the procedure are robust for adaptation in the case of Rician noise. It also shows to outperform the results of methods in the state-of-the-art. In future works, the procedure will be tested with other models of noise distributions. Proposing a more precise initial reduction of the noise and the processing with multi-resolution decomposition of the image. Other interest of this work is the automatic determination of the hyper-parameters of the PDF by using the input image and the proposed model. Also more work is required to avoid the residual noise in high levels of signal using equalization, normalizing or changing the ranges of the image's levels.

ANEXO A

K-MEANS AND PCA CLUSTERING

The PCA algorithm is a projection method in patten recognition for decorrelation, (Fukunaga, 1991; Zhang, 2010). In our model it is carried out over a matrix of extracted patches as vectors in consecutive columns; $\hat{U} \in R^{M \times MN}$ where $(M + p - 1) \times (N + p - 1)$ are the dimensions of the input data. The K-MEANS algorithm gives k clusters Q_k from the matrix \hat{U}_k , $Q_k \in \hat{U}$. The covariance matrix of the centralized data \hat{U} is Ω . A group of values c_k where $k \in \{1 \dots n_c\}$ are selected with distance $d = \max(\hat{U})/n_c$, in our case $n_c = 80$. Each column U_i is associated with a center c_k in the form $\min \|\hat{U}_k - c_k\|_2$. The clusters Q_k are obtained and a new center c_k over Q_k is updated. The process is iterated until the number of patches in each cluster Q_k maintain a constant value.

$$\Omega = \Psi \Lambda \Psi^T$$

Where Ψ and Λ are the eigenvector matrix and the eigenvalues matrix. The eigenspace of the data \hat{U}_k is named Δ and calculated as:

$$\Delta_h = U_k \Psi$$

In order to find the more representative patches centers (i, j) of \hat{Q}_k a threshold t_h for the major elements in Λ is selected where, also the corresponding eigenvectors are used to find a minimum distance as:

$$ph(pos) = \min_{v_{(i,j)}} \left\| \Delta^T \hat{U}_k - \hat{\Delta}^T \hat{v}_l(j) \right\|_2, \text{ where } \hat{v}_l(j) \in \hat{U}_k$$

Where $\hat{v}_l(j)$ are each patch in \hat{U}_k , $\hat{\Delta}^T$ is calculated with the t_h condition, $ph(pos)$ denotes the more representative patches in the cluster \hat{U}_k .

REFERENCES

Aarya-I., J.D. (2013). Adaptive SNR filtering technique for Rician noise denoising in MRI. In: *Proc. of the 2013 Intl. Conf on Biomedical Engineering*, pp. 1-5. Amphur Muang. Recuperado de <https://ieeexplore.ieee.org/abstract/document/6687669>

Aharon-M., E.M. (2006). K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 4311-4322. <https://doi.org/10.1109/TSP.2006.881199>

Casaca-W, B.M. (2010). A decomposition and noise removal method combining diffusion equation and wave atoms for textured images. *Mathematical Problems in Engineering*, March, 1-20. <http://dx.doi.org/10.1155/2010/764639>

Chang-S., B.Y. (2000). Adaptive wavelet thresholding for image denoising and compression. *Trans. Image Processing*, 1532-1546. <https://doi.org/10.1109/83.862633>

Elad-M, A.A. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. on Image Processing*, 3736-3745. <https://doi.org/10.1109/TIP.2006.881969>

Elad-M., A.Y. (2009). A plurality of sparse representations is better than the sparsest one alone. *IEEE Trans. on Information Theory*, 4701-4714. <https://doi.org/10.1109/TIT.2009.2027565>

Elahi P.S. (2014). BM3D mridenoising equipped with noise invalidation technique. In: *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 6612- 6616. Florence. Recuperado de <https://doi.org/10.1109/ICASSP.2014.6854879>

Fukunaga K. (1991). Introduction to statistical pattern recognition. New York: Academic. Recuperado de <https://doi.org/10.1016/C2009-0-27872-X>

He-L., A.G. (2009). A nonlocal maximum likelihood estimation method for rician noise reduction in MR images. *IEEE Trans. on Medical Imaging*, 165-172. <https://doi.org/10.1016/j.sigpro.2013.12.018>

Nair J.J. (2014). A robust non local means maximum likelihood estimation method for Rician noise reduction in MR images. In: *Proc. of the 2014 International Conference on Communication and Signal Processing*, pp. 848-852. Melmaruvathur. Recuperado de <https://doi.org/10.1179/1743131X15Y.0000000008>

Nowak R.D. (1999). Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans. on Image Processing*, 1408-1419. <https://doi.org/10.1109/83.791966>

Perona P. and Malik J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 629-639. <https://doi.org/10.1109/34.56205>

Pizurica-A, P.W. (2003). A versatile wavelet domain noise filtration technique for medical imaging. *IEEE Transactions on Medical Imaging*, 323-331. <https://doi.org/10.1109/TMI>

Pu, Y.S. (2014). Fractional partial differential equation denoising models for texture image. *Sci. China Inf. Sci.*, 1-19. <https://doi.org/10.1007/s11432-014-5112-x>

Pu, Y.Z. (2016). A Texture image denoising approach based on fractional developmental mathematics. pattern analysis and applications, pp. 1-19. Recuperado de <https://link.springer.com/article/10.1007/s10044-015-0477-z>

Rudin-L., O.S. (1992). Nonlinear total variation based noise removal algorithms. *Phys. D. Nonlinear Phenomena*, 259-268. [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F)

Samsonov A.A., a.J. (2004). Noise-adaptive nonlinear diffusion ltering of MR images with spatially varying noise levels. *Magn. Reson. Med.*, 798-806. <https://doi.org/10.1002/mrm.20207>

Sathyanarayana-S, A.D. (2011). An adaptive LMS technique for wavelet polynomial threshold denoising. In: *Proc of SPIE 8063 Mobile Multimedia/Image Processing, Security, and Applications*, p. 8063. Recuperado de <https://doi.org/10.1117/12.881297>

Sendur-L, A.S. (2002). Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependency. *IEEE Trans. on Signal Processing*, 2744-2756. <https://doi.org/10.1109/TSP.2002.804091>

Shashi-J, S. (2014). Rician Noise reduction in MRI images using wave atom transform. *International Journal of Computer Science and Mobile Computing*, 215-222.

Sijbers-J., J.D. (1998). Maximum-likelihood estimation of Rician distribution parameters. *IEEE Trans. on Medical Imaging*, 357-361. <https://doi.org/10.1109/42.712125>

Tikhonov, A.N. (1963). Solution of incorrectly formulated problems and regularization method. *Soviet Math. Dokl.*, 1035-1038. <https://doi.org/10.4236/eng.2011.312145>

Tisdall, D.A. (2005). MRI denoising via phase error estimation. *Proc. of SPIE 5747, Med. Imaging 2005: Image Processing*, 646-654. Recuperado de <https://doi.org/10.1117/12.595677>

- Tomasi-C., A.M. (1998). Bilateral Filtering for Gray and Color Images. Proceedings of the 1998. In: *IEEE International Conference on Computer Vision*. Bombay, India. Recuperado de <https://doi.org/10.1109/ICCV.1998.710815>
- W.-Dong, L.Z. (2011). Image Deblurring and Super-Resolution by Adaptive Sparse Domain Selection and Adaptive Regularization. *IEEE Trans. on Image Processing*, 1838-1857. <https://doi.org/10.1109/TIP.2011.2108306>
- W.-Dong, L.Z. (2013). Nonlocally Centralized Sparse Representation for Image Restoration. *IEEE Trans. on Image Processing*, 1620-1630. <https://doi.org/10.1109/TIP.2012.2235847>
- Weaver J.B., X.Y. (1991). Filtering noise from images with wavelet transforms. *Magn. Reson. Med*, 288-295. <https://doi.org/10.1002/mrm.1910210213>
- Wiest N., P.S. (2008,). Rician noise removal by non-Local Means filtering for low signal-to-noise ratio MRI: applications to DT-MRI. *Med. Image Comput. Assist*, 171-179.
- Yang J., W.J. (2010). Image Super-Resolution Via Sparse Representation. *IEEE Trans. on Image Processing*, 2861-2873. <https://doi.org/10.1109/TIP.2010.2050625>
- Zhang L., D.W. (2010). Two-stage image denoising by principal component analysis with local pixel grouping. *Pattern Recognit*, 1531-1549. <http://dx.doi.org/10.1016/j.patcog.2009.09.023>