ANISOTROPIC LMMSE DENOISING OF MRI BASED ON STATISTICAL TISSUE MODELS

G. Vegas-Sánchez-Ferrero[1,2], *A. Tristán-Vega*[1], *S. Aja-Fernández*[1], *M. Martín-Fernández*[1], *C. Palencia*[1], *R. Deriche*[2]

[1]Universidad de Valladolid, Spain. [2]Athena Project Team, INRIA Sophia Antipolis, France *

ABSTRACT

Linear Minimum Mean Squared Error Estimation (LMMSE) is a simple, yet powerful denoising technique within MRI. It is based on the computation of the mean and variance of the data being filtered according to a noise model assumed, which is usually accomplished by calculating local moments over squared neighborhoods. When these neighborhoods are centered in pixels corresponding to image contours, the estimation is not accurate due to the presence of two or more tissues with different statistical properties. We overcome this limitation by introducing an anisotropic LMMSE scheme: the gray levels of each tissue in the MRI volume are modeled as a Gamma-mixture, such that we can discriminate between the different matters to construct anisotropic neighborhoods containing only one kind of tissue. The potential of the Gamma distribution relies on its ability to fit both the Rician distribution traditionally used to model the noise in MRI and the non-central Chi noise found in modern parallel MRI systems.

Index Terms— LMMSE, MRI, Rician, non-central Chi.

1. INTRODUCTION

Magnetic Resonance Imaging (MRI) is one of the medical imaging modalities that exhibits the largest Signal to Noise Ratio (SNR). Even so, noise-induced artifacts in MRI are not negligible, degrading image quality in both qualitative and quantitative applications. On one hand, several new acquisition technologies have appeared lastly, aiming to speed-up the acquisition of MRI, and to reduce phase distortions originated by the strong sensitizing gradients used in Diffusion Weighted Imaging (DWI). These methods are commonly referred to as Parallel Magnetic Resonance Imaging (pMRI), and are based on the simultaneous acquisition of different parts of the subsampled k-space by different receiving coils. A well-known side effect of this methodology is the worsening of the final SNR, which is typically modeled through the g-factor [1]. On the other hand, imaging modalities such as DWI demand much larger amounts of data compared to anatomical MRI, and are also noisier. As a consequence, direct procedures for boosting the SNR in MRI, mainly the averaging of multiple independent repetitions of the volume, are no longer practical.

Given these concerns, the need for *a posteriori* denoising schemes have become clear to many authors. General purpose and Gaussian-based filters have been used, though Rician-adapted techniques have been proven more adequate for quantitative MRI. They include the conventional approach [2], Rician Maximum Likelihood estimation [3], Expectation-Maximization (EM) [4], wavelets [5], anisotropic diffusion [6], and Unbiased Non-Local Means [7].

Among these approaches, the LMMSE estimator for Rician data first described in [8] has shown to outperform the most common techniques. This method is a classical estimation theory approach: provided we can characterize the expected value and covariance matrix of both the signal to estimate and the noise, it obtains the best linear estimation (i.e. that with minimum variance) of the data. These parameters may be easily inferred from local moments computed as sample estimates inside image neighborhoods. Besides, the linear nature of the filter allows a very fast implementation, and moreover it may be adapted to cope with noise models deviating from Rician, mainly the non-central Chi (nc-χ) [9], arising with certain modern machinery and pMRI protocols [1].

An important limitation of LMMSE models relies on the way the local moments are estimated. For those voxels corresponding to image edges, the neighborhoods used typically comprise several different kinds of tissues. Hence, the value estimated for the variance is artificially increased due to the superposition of the effect of noise and the multi-modal distribution of gray levels corresponding to each matter. As a consequence, edge voxels are not properly denoised, since the noisy pattern is mixed-up with the fine details (gradients).

In this work we propose to estimate the probabilistic response of each tissue by Gamma distributions for pMRI protocols. These distributions provide a good fitting as was demonstrated by Patniak in [10]. With this estimation, the probability of belonging to each tissue class can be obtained and, thus, the local moments used for the calculation of the

*The authors acknowledge: Gerencia Regional de Salud, Consejeria de Sanidad, Junta Castilla y Leon for grants GRS 474/A/1O and VA376A11-2; Ministerio de Ciencia e Innovacion for grants CEN-20091044, TEC2010-17982 and MTM2007-63257; Proyectos de investigacion en salud, Accion Estrategica en Salud, Instituto de Salud Carlos III for grant PI11-01492 and Centro para el Desarrollo Tecnico Industrial, Ministerio de Ciencia e Innovacion, CEN-20091044 (cvREMOD). This work was partially supported by the French ANR project NucleiPark and the France-Parkinson Association.

LMMSE can be derived as the local moments conditioned to each of the tissues. As a consequence, the estimation of local variance is not biased by the presence of different tissues because only those pixels belonging to each tissue class contribute to the calculation of moments resulting in an anisotropic denoising. Results demonstrate that this approach outperforms isotropic LMMSE methods [9].

2. METHODS

2.1. Background

In the case of multiple-coil acquisition, the composite magnitude image can be obtained using Sums of Squares (SoS) as $M_L = \sqrt{\sum_{l=1}^{L} |S_l|^2}$, where each coil signal is modeled as $S_l = A_l + N_l(\sigma_n^2)$ where N_l is a complex Gaussian process with zero mean and variance σ_n^2. Defining $A_T = \sqrt{\sum_{l=1}^{L} |A_l|^2}$ and assuming the noise components are independent and identically distributed (IID), the magnitude signal M_L follows a nc-χ distribution [9]. When the square of the magnitude signal is considered, M_L^2, the PDF trivially becomes a nc-χ square distribution (nc-χ^2).

The LMMSE filter in [8] and the extended version of [9] rely on the following expression:

$$\widehat{A_T^2} = E\{A_T^2\} + \mathbf{C}_{A_T^2 M_L^2} \mathbf{C}^{-1}_{M_L^2 M_L^2} (M_L^2 - E\{M_L^2\}\}) \quad (1)$$

where $\mathbf{C}_{M_L^2 M_L^2}$ is the covariance of M_L^2 and $\mathbf{C}_{A_T^2 M_L^2}$ is the cross-covariance. The Characteristic function of a nc-χ^2 is:

$$\varphi_{M_L^2}(t) = E\{e^{itM_L^2}\} = \frac{1}{(1-2it\sigma_n^2)^L} e^{\frac{iA_T^2 t}{1-2it\sigma_n^2}} \quad (2)$$

Note that this result gives rise to a Gamma distribution in the absence of the original signal ($A_T = 0$). The PDF, f_X, and Characteristic function, φ_X, of a Gamma distributed random variable, X, of parameters α and β are:

$$f_X(x|\alpha,\beta) = \frac{x^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} e^{-\frac{x}{\beta}} u(x), \;\; \varphi_X(t) = \frac{1}{(1-i\beta t)^\alpha} \quad (3)$$

where $\alpha, \beta > 0$; $\Gamma(x)$ is the Euler Gamma function.

This result was noticed by Patniak in [10], where the approximation of nc-χ^2 by a Gamma distribution, was firstly proposed. As an example, in Fig. 1.a the Gamma PDF approximation is depicted for a set of synthetic images of size 512×512 obtained by SoS with incremental $A_T \in [0, 100]$, number of coils $L = 8$, and $\sigma_n = 20$. Note that the theoretical PDF of the nc-χ^2 and the Gamma approach are almost indistinguishable and both follow the empirical PDF of M_L^2.

This result shows the potential of the Gamma distribution for fitting the response of both the Rician distributed noise traditionally used to model conventional MRI and nc-χ distributed noise found in modern parallel MRI systems. We propose to estimate the probabilistic response of each tissue

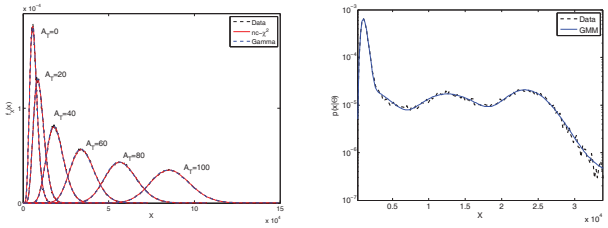


Fig. 1. (a) M_L^2 PDF for $A_T \in [0, 100]$, $L = 8$, and $\sigma_n = 20$. (b) Empirical PDF of a pMRI T1 image ($L = 8$ and $\sigma_n = 70$) and the proposed GMM for 5 elements in the mixture.

by a Gamma distributions. This way, the local moments required for the LMMSE can be calculated as the local moments conditioned to each of the tissues. In order to find the distributions of different tissues of the image, a Gamma mixture model is proposed and explained in the next section.

2.2. Gamma Mixture Model

The mixture of Gamma distributions (GMM) is calculated by means the EM algorithm [4] in the following way[1]: Let $\mathbf{X} = \{x_i\}$, $1 \leq i \leq N$ be a IID set of samples (pixel intensities of M_L^2). The GMM considers that these variables result from the contributions of J distributions: $p(x_i|\mathbf{\Theta}) = \sum_{j=1}^{J} \pi_j f_X(x_i|\Theta_j)$, where $\mathbf{\Theta}$ is a vector of the parameters of the GMM (π_j, Θ_j) and Θ_j are the parameters of the PDF, α_j and β_j and $\sum_{j=1}^{J} \pi_j = 1$.

The method is applied in the following way:

1. A first estimate of the hidden variables is obtained by k-means. For each cluster,j parameters $\Theta_j^{(0)}$ and $\hat{\pi}_j^{(0)}$, are calculated from samples.
2. Bayesian Inference Step. Calculate $\gamma_{i,j}$ as:

$$p(Z_i = j|x_i, \mathbf{\Theta}^{(n)}) = \gamma_{i,j} = \frac{\pi_j^{(n)} p(x_i|\Theta_j^{(n)})}{p(x_i|\mathbf{\Theta}^{(n)})}, \quad (4)$$

3. Expectation-Maximization Step. The estimate for $\hat{\alpha}_j$, $\hat{\beta}_j$ and $\hat{\pi}_j$ are calculated by the following equations:

$$\log(\hat{\alpha}_j) - \psi(\hat{\alpha}_j) = \log\left(\frac{\sum_i^N \gamma_{i,j} x_i}{\sum_i^N \gamma_{i,j}}\right) - \frac{\sum_i^N \gamma_{i,j} \log x_i}{\sum_i^N \gamma_{i,j}} \quad (5)$$

$$\hat{\pi}_j = \frac{1}{N} \sum_{i=1}^{N} \gamma_{i,j}, \;\; \hat{\beta}_j = \frac{1}{\hat{\alpha}_j} \frac{\sum_{i=1}^N \gamma_{i,j} x_i}{\sum_{i=1}^N \gamma_{i,j}} \quad (6)$$

4. Go to step 2 until $||\mathbf{\Theta}^{(n)} - \mathbf{\Theta}^{(n-1)}|| <$ Tolerance

As an example of the GMM, in Fig. 1.b the GMM is fitted to a pMRI T1 of Fig. 3 acquired with $L = 8$ and $\sigma_n = 70$ and the proposed GMM for 5 elements in the mixture.

[1]This method maximizes the log-likelihood function when some hidden discrete random variables (RVs), $\mathbf{Z} = \{Z_i\}$, are introduced to the model. These RVs take values in $\{1, \cdots, J\}$, their meaning is that the sample x_i belongs to the distributions class j when $Z_i = j$.

2.3. Anisotropic-LMMSE

In [9] it is shown that for uncorrelated and independent Gaussian noises the variance and covariance terms, $C_{M_L^2,M_L^2}$ and $C_{A_T^2,M_L^2}$ can be calculated as:

$$C_{M_L^2,M_L^2} = E\{M_L^4\} - E\{M_L^2\}^2 \quad (7)$$

$$C_{A_T^2,M_L^2} = E\{A_T^4\} + 2L\sigma_n^2 E\{A_T^2\} - E\{A_T^2\}E\{M_L^2\} \quad (8)$$

where the 2nd and 4th moments of A_T can be derived by the calculation of the same order moments of M_L^2 which is a nc-χ^2 distribution:

$$E\{A_T^4\} = E\{M_L^4\} - 4(E\{M_L^2\} - 2\sigma_n^2)(L+1)\sigma_n^2 - 4L(L+1)\sigma_n^4 \quad (9)$$

$$E\{A_T^2\} = E\{M_L^2\} - 2L\sigma_n^2 \quad (10)$$

These moments can be calculated by assuming local ergodicity as was done in [8, 9]. Hence, the expectation, $E\{\cdot\}$ can be replaced by local spatial mean, $\langle\cdot\rangle$, calculated in a neighborhood. Now, the LMMSE can be calculated just as a function of the local moments of M_L^2, $\langle M_L^2\rangle$ and $\langle M_L^4\rangle$.

The calculation of the components for the mixture model provides important information related to the probability of belonging to each tissue class. This information can be taken into account by calculating the conditioned mean for each tissue class in the local neighborhood, $\eta(x,y)$. When the moments are calculated in this way, only those pixels belonging to each tissue class contribute to the estimate of each local moment, resulting in a more accurate estimate of local moments. This result can be considered as a non isotropic way to calculate the local moments since each pixel is not equally treated when the local moments are calculated.

By introducing the probability of belonging to each tissue class in the calculation of the local moments, they are calculated as:

$$\langle M_L^2(x,y)\rangle = \sum_{j=1}^{J} P(Z=j)\,\langle M_L^2(x,y)|Z=j\rangle \quad (11)$$

where the local moment for each class is:

$$\langle M_L^2|Z=j\rangle = \frac{\sum_{(x',y')\in\eta(x,y)} M_L^2(x',y')p(Z(x',y')=j|M_L^2,\boldsymbol{\Theta}^{(n)})}{\sum_{(x',y')\in\eta(x,y)} p(Z(x',y')=j|M_L^2,\boldsymbol{\Theta}^{(n)})} \quad (12)$$

with

$$p(Z(x',y')=j|M_L^2,\boldsymbol{\Theta}^{(n)}) = \frac{\pi_j \cdot p(M_L^2|\Theta_j^{(n)})}{p(M_L^2|\boldsymbol{\Theta}^{(n)})} \quad (13)$$

The weight $P(Z=j)$, is the probability of belonging to the j-th tissue class for the neighborhood of (x,y) and is calculated as:

$$P(Z(x,y)=j) = \frac{\prod_{(x',y')\in\eta(x,y)} p(Z(x',y')=j|M_L^2,\boldsymbol{\Theta}^{(n)})}{\sum_{j=1}^{J}\prod_{(x',y')\in\eta(x,y)} p(Z(x',y')=j|M_L^2,\boldsymbol{\Theta}^{(n)})} \quad (14)$$

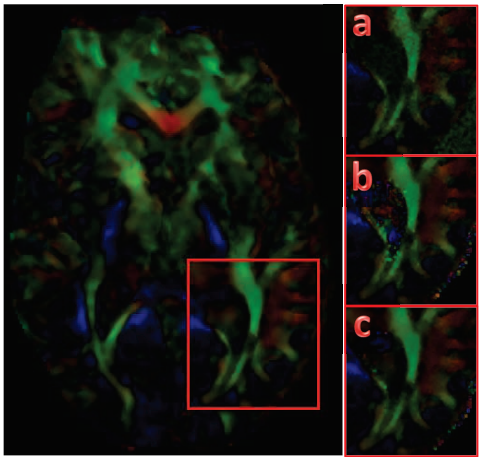


Fig. 2. (a) Noisy (SNR=15db) (b) nc-χ LMMSE (c) Proposed

3. RESULTS

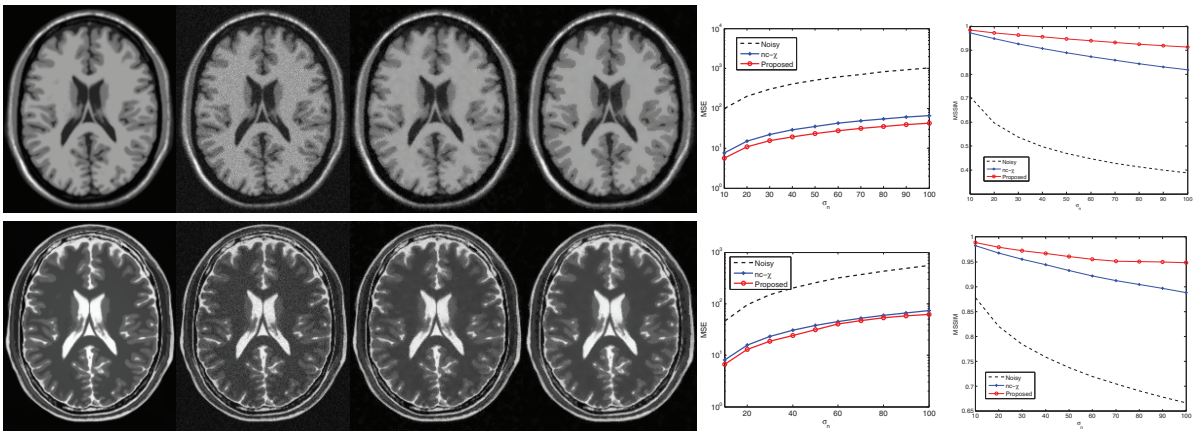
For the sake of validation, two experiments are considered to test the performance of the anisotropic LMMSE in comparison to the isotropic nc-χ LMMSE [9]. In the first one, 3D synthetic T1 and T2 images from BrainWeb MR [11] with intensity values in $[0-255]$ and 1 mm Slice thickness were considered. A simulation of 8-coil system was performed for an increasing $\sigma_n \in [10, 100]$ in each coil [1]. The image was reconstructed by using SoS. A neighborhood of $5\times5\times5$ for both the proposed and the nc-χ LMMSE filter in [9] was used. The number of elements in the GMM was 5.

Results for the experiment are shown in Fig. 3, in the top row, the T1 image is shown as well as the noisy image (Fig. 3.b) for $\sigma_n = 70$. The denoised image obtained by the isotropic nc-χ LMMSE is shown in Fig. 3.c. Note that this filter retains noise in regions with the presence of several tissues. the proposed method is shown in Fig. 3.d where the better performance in areas with more than one tissue can be appreciated at first sight. An analogous conclusion can be obtained for T2 images represented in the lower row of Fig. 3.

The quantitative evaluation was performed using the mean squared error (MSE) and the mean structural similarity[2] (MSSIM) [12]. Twenty independent trials were considered for each value of σ_n. The mean of both measures is represented in columns Fig. 3.e-f for T1 and T2.

Though the proposed method outperforms the isotropic one for both measures, for the MSE both methods do not exhibit a big difference. This is so because homogeneous areas are equally treated by both methods and pixels in the edge of different tissues are not explicitly reflected by the MSE. Nevertheless, for the MSSIM measure, the better performance of the anisotropic filter becomes evident as we increase σ_n. This result confirms the ability for detecting the structures of the image leading to a better estimate of local moments and, thus,

[2]This gives a similarity value in $[0, 1]$.



(a) Original (b) Noisy (c) nc-χ LMMSE (d) Proposed (e) MSE (f) MSSIM

Fig. 3. Quantitative comparison with synthetic T1 (first row) and T2 (second row) for increasing σ_n.

overcoming the main problem of LMMSE methods.

For the second experiment, a 15 gradients DWI phantom presented in [13] was considered. A simulation of 8-coil system was performed for a SNR=15 dB. In Fig. 2 the fractional anisotropy colored by orientation is represented. The zoomed area shows a better definition of areas with different orientations. Quantitatively, the MSE was calculated for the mean FA; noisy: 0.0271, nc-χ LMMSE: 0.0124, and the proposed: 0.0055. Additionally, the MSE of the trace of the diffusion tensor gives the following results; noisy: $7.38 \cdot 10^{-6}$, nc-χ LMMSE: $8.23 \cdot 10^{-7}$, and the proposed: $5.7 \cdot 10^{-7}$. Both measures show the better performance of the proposed method for the calculation of the diffusion tensor.

4. CONCLUSIONS

An anisotropic LMMSE method is proposed. The anisotropy is achieved by applying statistical models to the tissues based on a GMM model. This mixture have proved to properly fit the probabilistic behavior of noise in conventional MRI (Rician) and pMRI (nc-χ). Experiments have shown that the inclusion of statistical models lead to a better estimate of local moments by considering the local moments conditioned to the probability of belonging to each tissue class.

5. REFERENCES

[1] S. Aja-Fernández, A. Tristán-Vega, and S. Hoge, "Statistical noise analysis in GRAPPA using a parametrized non-central chi approximation model," *Magnet Reson. Med.*, vol. 65, no. 4, pp. 1195–1206, Apr. 2011.

[2] G. McGibney and M.R. Smith, "Unbiased signal-to-noise ratio measure for magnetic resonance images," *Med. Phys.*, vol. 20, no. 4, pp. 1077–1078, 1993.

[3] J. Sijbers and A. J. den Dekker, "Maximum likelihood estimation of signal amplitude and noise variance form MR data," *Mang. Reson. Imaging.*, vol. 51, pp. 586–594, 2004.

[4] T.L. Marzetta, "EM algorithm for estimating the parameters of multivariate complex Rician density for polarimetric SAR," in *Proc. of ICASSP*, 1995, vol. 5, pp. 3651–3654.

[5] R.D. Nowak, "Wavelet-based Rician noise removal for Magnetic Resonance Imaging," *IEEE Trans. Imag.e Proc.*, vol. 8, no. 10, pp. 1408–1419, Oct. 1999.

[6] K. Krissian and S. Aja-Fernández, "Noise-driven anisotropic diffusion filtering of MRI," *IEEE Trans. Imag. Proc.*, vol. 18, no. 10, pp. 2265 –2274, Oct. 2009.

[7] N. Wiest-Daesslé, S. Prima, P. Coupé, S.P. Morrissey, and C. Barillot, "Non-local means variants for denoising of diffusion-weighted and diffusion tensor MRI," in *MICCAI*. 2007, vol. 4792 of *LNCS*, pp. 344–351, Springer–Verlag.

[8] S. Aja-Fernández, C. Alberola-López, and C.-F. Westin, "Noise and signal estimation in magnitude MRI and Rician distributed images: A LMMSE approach," *IEEE Trans. Imag. Proc.*, vol. 17, no. 8, pp. 1383–1398, Aug. 2008.

[9] V. Brion, C. Poupon, O. Riff, S. Aja-Fernández, A. Tristán-Vega, J-F. Mangin, D. Le Bihan, and F. Poupon, "Parallel mri noise correction: An extension of the LMMSE to non central distributions," in *MICCAI (2)*, 2011, pp. 226–233.

[10] P. B. Patnaik, "The non-central χ^2- and f-distribution and their applications," *Biometrika*, vol. 36, pp. 202–232, 1949.

[11] D. Louis Collins, A.P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imaging*, vol. 17, no. 3, pp. 463–468, 1998.

[12] Z. Wang, Alan C. Bovik, Hamid R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[13] A. Tristán-Vega and S. Aja-Fernández, "Design and construction of a realistic DWI phantom for filtering performance assessment," in *MICCAI*. Sept. 2009, LNCS, pp. 951–958, Springer-Verlag.