Signal LMMSE Estimation from Multiple Samples in MRI and DT-MRI

S. Aja-Fernández[1,2], C. Alberola-López[1], and C.-F. Westin[2]

[1] LPI, ETSI Telecomunicación, Universidad de Valladolid
[2] LMI, Brigham and Women's Hospital,Harvard Medical School

Abstract. A method to estimate the magnitude MR data from several noisy samples is presented. It is based on the Linear Minimum Mean Squared Error (LMMSE) estimator for the Rician noise model when several scanning repetitions are available. This method gives a closed-form analytical solution that takes into account the probability distribution of the data as well as the existing level of noise, showing a better performance than methods such as the average or the median.

1 Introduction

Magnetic Resonance Imaging (MRI) or Diffusion Weighted MRI (DW-MRI) provide the possibility of acquiring several –and fairly aligned– images of the same slice or even of the same volume. The number of scanning repetitions is usually known as NEX (number of excitations). These multiple samples may be used to estimate the magnitude image, as a way to reduce the level of noise as well as other type of artifacts. Although in literature some methods based on estimators using the Rician model have been reported, as that based on Maximum Likelihood (ML) [1], due to their complexity this task is often done using a simple average or median operator.

In this paper we propose an alternative Bayesian approach based on the Linear Minimum Mean Squared Error (LMMSE) estimator. If an accurate measure of the level of noise is feasible, the proposed estimator will be able to remove it more satisfactorily than the average and median operators and, although suboptimally with respect to the MLE, much more efficiently.

2 Rician Model and Signal Estimation

Due to the existence of uncorrelated Gaussian noise with zero mean and the same variance in both the real and imaginary parts of the complex k-space data, the magnitude signal of MR data may be modeled following a a Rician distribution, whose probability distribution function (PDF) for a 2D signal is as follows [2]

$$p_M(M_{ij}|A_{ij},\sigma_n) = \frac{M_{ij}}{\sigma_n^2}\, e^{-\frac{M_{ij}^2+A_{ij}^2}{2\sigma_n^2}}\; I_0\left(\frac{A_{ij}M_{ij}}{\sigma_n^2}\right)\; u(M_{ij}) \tag{1}$$

N. Ayache, S. Ourselin, A. Maeder (Eds.): MICCAI 2007, Part II, LNCS 4792, pp. 368–375, 2007.

being $I_0(.)$ the 0^{th} order modified Bessel function of the first kind, $u(.)$ the Heaviside step function and σ_n^2 the variance of noise. M_{ij} is the magnitude value of the pixel $\{i, j\}$ and A_{ij} the original value of the pixel without noise. Several samples of each slice will be considered, being $M_{ij}[k]$ the k-th scanning repetition of pixel $\{i, j\}$ in the actual slice. These repetitions are usually *fused* using an average or median operator. The effect of using the average operator may be easily observed in the areas of low SNR, like the background, where the Rician distribution tends to be Rayleigh. After the estimation, the signal value in the background pixels should be zero. However, when using the average operator, this value tends to be the mean of a Rayleigh PDF [2], i.e. $\sigma_n\sqrt{\pi/2}$. In a similar way, the median of a Rayleigh PDF is $\sigma_n\sqrt{4}$. In both cases, although there may be a smoothing of the noisy region, there is also a bias related with σ_n in the output values.

One feasible option is to use a ML estimator for the magnitude data [1], which is defined for multiple samples following a Rician distribution as $\widehat{A}_{ML} = \arg\max_A\{\log L\}$, being $\log L$ the log-likelihood function [1,3]. As this equation cannot be solved analytically, the maximum of the log-likelihood function must be found numerically. This task is computationally expensive, the more expensive the higher the number of images to be processed, especially when working with DTI data, with multiple slices and multiple gradient directions. Alternative methods to solve the ML estimator have been proposed, as the one based in Expectation-Maximization [3] or the work by Fillard *et al.* [4]. Other works use the Maximum a Posteriori, as Basu *et al.* [5].

We now propose a different approach to estimate the signal from the magnitude image, based on the LMMSE estimator. Instead of modeling A as unknown constant, we will consider it as a realization of a random variable which is functionally related to the observation. Although this approach may be suboptimal with respect to the MLE, the fact that a closed-form analytical solution is achievable, makes the whole process faster and more suitable when working with large amount of data -like DTI- where an optimization method to search $\widehat{A}_{ML}$ would be too slow.

3 LMMSE Estimation from Multiple Noisy Samples

The LMMSE estimator of a parameter θ using multiple samples is defined [6]

$$\hat{\theta} = E\{\theta\} + \mathbf{C}_{\theta x}\mathbf{C}_{xx}^{-1}(\mathbf{x} - E\{\mathbf{x}\}) \tag{2}$$

being $\mathbf{C}$ the covariance matrices and $\mathbf{x}$ the vector of available samples. The moments of the Rician distribution have a non-trivial integral expression. However, the even-order moments, are simple polynomial. In order to achieve a closed-form expression we will use A^2 instead of A. Consequently, all the moments to be used will be even. With this assumption in mind, the LMMSE estimator for the Rician distribution is

$$\widehat{A_{ij}^2} = E\{A_{ij}^2\} + \mathbf{C}_{A_{ij}^2 M_{ij}^2}\mathbf{C}_{M_{ij}^2 M_{ij}^2}^{-1}\left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\}\right) \tag{3}$$

For the sake of simplicity we will suppose that all the equations are pixelwise, removing the subindexes $\{i, j\}$. Assuming that N measures are taken of every pixel, $\mathbf{M} = [M[1]\ M[2]\ \cdots\ M[N]]^T$ is the measure vector. $\mathbf{M}^2$ must be understood *element-wise*, i.e. $\mathbf{M}^2 = [M^2[1] \cdots\ M^2[N]]^T$. $\mathbf{C}_{M^2M^2}$ is the $N \times N$ covariance matrix of $\mathbf{M}^2$, defined

$$\mathbf{C}_{M^2M^2} = E\{\left(\mathbf{M}^2 - E\{\mathbf{M}^2\}\right)\left(\mathbf{M}^2 - E\{\mathbf{M}^2\}\right)^T\}$$

After some algebra and replacing expectations by their sample estimator $\langle . \rangle$, we can finally write the covariance matrix as

$$\mathbf{C}_{M^2M^2} = \left(\langle \mathbf{M}^4 \rangle + 4\sigma_n^4 - 4\sigma_n^2\langle \mathbf{M}^2 \rangle - \langle \mathbf{M}^2 \rangle^2\right)\mathbf{1}_N\mathbf{1}_N^T - \left(4\sigma_n^4 - 4\sigma_n^2\langle \mathbf{M}^2 \rangle\right)\mathbf{I}_N \tag{4}$$

where $\langle \mathbf{M}^a \rangle = \frac{1}{N}\sum_{k=1}^{N} M^a[k]$, $\mathbf{1}_N$ is an all 1 vector of length N, and $\mathbf{I}_N$ is the $N \times N$ identity matrix. Matrix $\mathbf{C}_{A^2M^2}$ is

$$\begin{aligned}\mathbf{C}_{A^2M^2} &= E\{\left(A^2 - E\{A^2\}\right)\left(\mathbf{M}^2 - E\{\mathbf{M}^2\}\right)^T\} \\ &= \left(\langle \mathbf{M}^4 \rangle + 4\sigma_n^4 - 4\sigma_n^2\langle \mathbf{M}^2 \rangle - \langle \mathbf{M}^2 \rangle^2\right)\mathbf{1}_N^T\end{aligned}$$

Finally, for each point in the image, the estimator will be

$$\widehat{A^2} = \langle \mathbf{M}^2 \rangle - 2\sigma_n^2 + \mathbf{C}_{A^2M^2}\mathbf{C}_{M^2M^2}^{-1}\left(\mathbf{M}^2 - \langle \mathbf{M}^2 \rangle\right) \tag{5}$$

This equation must be understood pixelwise (say $M_{ij}[k]$ and A_{ij}) in the two dimensional case or voxelwise (say $M_{ijl}[k]$ and A_{ijl}) in the three dimensional case. Note that the value of the variance of noise σ_n^2 value must be properly estimated somehow. Several methods have been reported in literature [7]. New robust methods are also emerging for this task, making the proposed method useful.

4 Validation

Some synthetic experiments have been carried out in order to validate the LMMSE estimator previously introduced. Firstly we will compare it with other fusion methods over a single MR image. The image in Fig. 1-(a) will be considered as the *ground truth*. The image is corrupted with synthetic Rician noise with different values of σ_n. For each value, 10 independent noisy images are created, say $I_n[k]$, $k = 1, \cdots, 10$. This images are fused using the following methods: average of the images, (I_a), median of the images (I_m) and LMMSE estimator of eq. (5), say I_e.

To compare the performance of the different approaches two structural quality indexes have been used: the Structural Similarity (SSIM) index [8] and the Quality Index based on Local Variance (QILV) [9]. Both give a measure of the structural similarity between the golden standard and the other images. However, the former is more sensitive to the level of noise in the image and the latter to any possible blurring of the edges. Both indexes are bounded; the closer to 1,

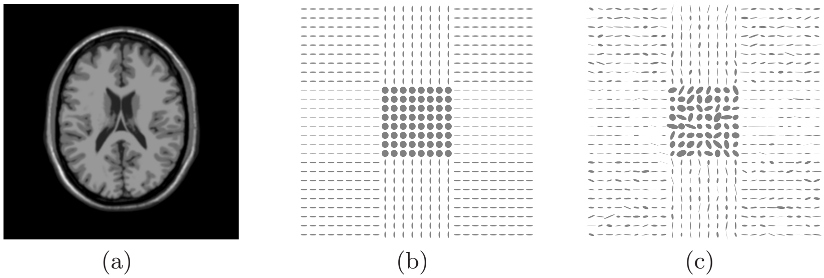


Fig. 1. Data used for the experiments. (a) MR image (256 gray levels) from BrainWeb database (`http://www.bic.mni.mcgill.ca/brainweb/`). (b) Synthetic 2D tensor field. (c) Noisy 2D tensor field, $\sigma_n = 100$ (SNR=2.62dB). Figures (b) and (c) created using Teem software. (`http://teem.sourceforge.net/`).

Table 1. Quality measures. Average of 100 realizations with $\sigma_n = 10$. The LMMSE shows the better results in the structural measures, and the lower MSE and level of noise.

	MSE	SSIM	QILV	LN
I_n	158.5421	0.3882	0.9890	9.9517
I_a	100.0603	0.4723	0.9918	9.9887
I_m	93.4674	0.4693	0.9932	9.4919
I_e	**33.7761**	**0.5227**	**0.9986**	**5.1882**

(a) Whole Image

	MSE	SSIM	QILV	LN
I_n	112.4711	0.7362	0.9882	9.9517
I_a	31.6995	0.9090	0.9912	9.9887
I_m	32.8593	0.8983	0.9924	9.4919
I_e	**15.6514**	**0.9209**	**0.9982**	**5.1882**

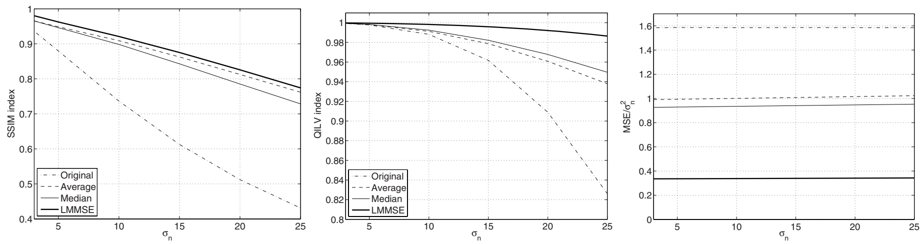
(b) Image without background

the better the quality. In addition we will also use the Mean Square Error (MSE) and the Level on Noise (LN) in the output image, measured as $LN = \frac{(\hat{\sigma}_n)_o}{(\sigma_n)_i}$, being $(\hat{\sigma_n})_o$ the estimated standard deviation of noise in the resulting image and $(\sigma_n)_i$ the standard deviation of noise in the original image.

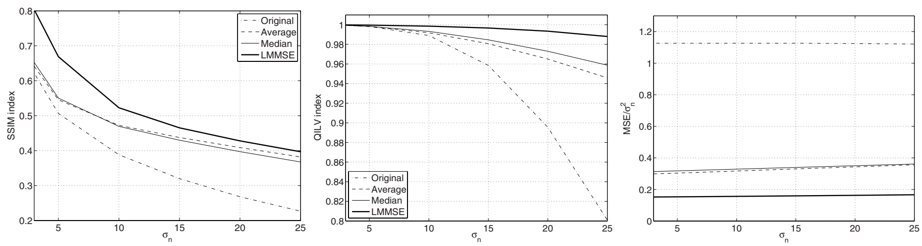
The average of 100 realizations for each method and each value of σ_n has been considered. Results are on Fig. 2 and Fig. 3-(a). The whole image and the image without background have been considered separately. As an illustration, numerical results for $\sigma_n = 10$ are shown in Table 1.

The proposed method shows a better performance in all the indexes and for all the levels of noise. It is the one with the lower MSE and the higher values of SSIM and QILV. When comparing the LN, it is the only one in which a reduction of noise is noticeable. However, for small levels of noise (say $\sigma_n < 5$ for a 256 gray-level image), the use of the average and the median may be a feasible alternative.

As we have stated previously, the method here presented is suboptimal with respect to a nonlinear counterpart or even the MLE. With respect to the latter, Fig. 3-(b) shows some comparative results for different values of σ_n. However, a numerical optimization is needed for the MLE, a fact that is avoided with our alternative simple (and linear) solution.



(a) SSIM (whole image) (b) QILV (whole image) (c) MSE/σ_n^2 (whole image)



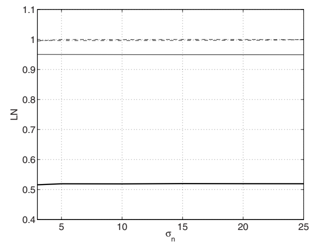
(d) SSIM (without background) (e) QILV (without background) (f) MSE/σ_n^2 (without background)

Fig. 2. Quality measures of the resulting fused images. In all the cases the proposed method shows a better performance.

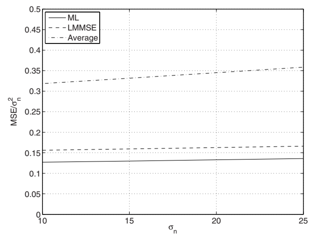
When working with DT-MRI, some scalar measures like the Fractional Anisotropy [10] are directly related to the eigenvalues of the diffusion tensors. To study the effect of the proposed method over these eigenvalues, a synthetic data set has been created: a 128×128 2D tensor field, as shown in Fig. 1-(b), where tensors are depicted using ellipses. Tensors with three different eigenvalue combinations were chosen

$$\boldsymbol{\lambda}_a = [1.9\ 10^{-3}, 0.4\ 10^{-3}]\ \boldsymbol{\lambda}_b = [2\ 10^{-3}, 0.1\ 10^{-3}], \boldsymbol{\lambda}_c = [2\ 10^{-3}, 2\ 10^{-3}]\}$$

and the diffusion weighted images (DWI) were simulated using the Stejskal-Tanner equation [10,11]. Different number of gradients have been considered, with a constant baseline with a level of 1000. The DWI are corrupted with Rician noise, Fig. 1-(c), and the tensors are re-estimated, using a Least Squares approach. Different values of σ_n have also been used. For the experiments 3, and 15 gradient directions, $\sigma_n \in [30, 210]$ and $N = 10$ will be considered. We define the signal to noise ratio as SNR (dB) $= 10 \log_{10} \left(S^2/\sigma_n^2\right)$, and we will define the power of the signal as $S^2 = \min\left(S^2_{i,j,k}\right)$, in our case $S^2 = 1.83\ 10^4$. The error is defined as the absolute distance of the estimated eigenvalues to the original values. For each number of gradients and each SNR value the average of 100 experiments is considered. In Fig. 4 the mean and the standard deviation of the error are shown. From the results it can be seen that in all the cases the bias and the variance of the estimation is smaller when using the LMMSE

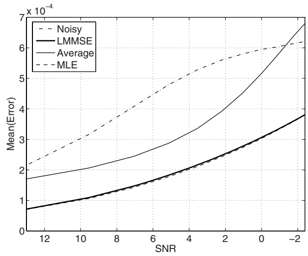


(a) Level of Noise

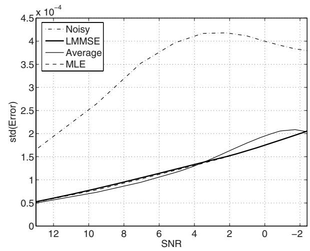


(b) MSE/σ_n^2

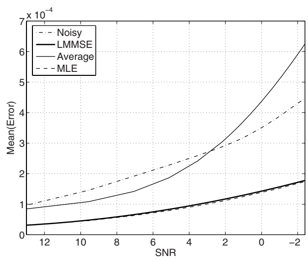
Fig. 3. (a) Level of noise: estimated standard deviation of noise of the output image normalized by the standard deviation of noise in the original image. Same legend as Fig. 2. the LMMSE scheme is the only one that shows a significant reduction in the level of noise. (b) Comparison of estimators using the MSE/σ_n^2 of the image without background.



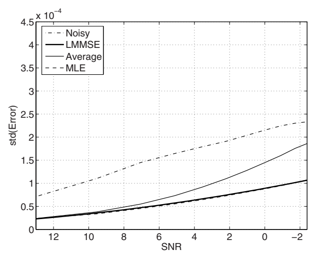
(a) Error mean, 3 direc.



(b) Error STD, 3 direc.



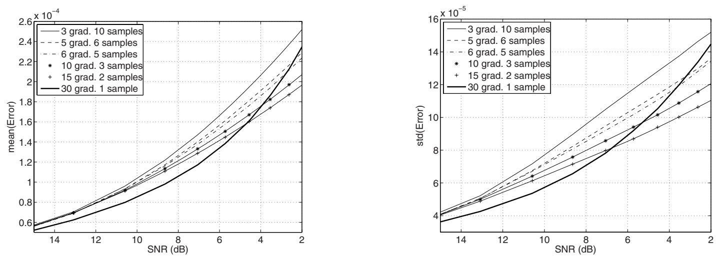
(c) Error mean, 15 direc.



(d) Error STD, 15 direc.

Fig. 4. Mean and standard deviation of the eigenvalue estimation error for 3 and 15 gradient directions

filter. Note that the effect of using more gradient directions positively affects to the variance or error of the original data and the data fused with LMMSE, but it hardly affects to the variance of the error when using the average. As in the previous experiment, results are just slightly better for the MLE but again a numerical optimization is needed to obtain the solution. In our experiments–using an optimized MATLAB code–, for 10 samples and 15 gradients LMMSE



(a) Error Mean (b) Error STD

Fig. 5. Mean (left) and standard deviation (right) of the error of estimation of the eigenvalues for different combination of number of gradient direction and number of samples. LMMSE estimator has been used for fusion.

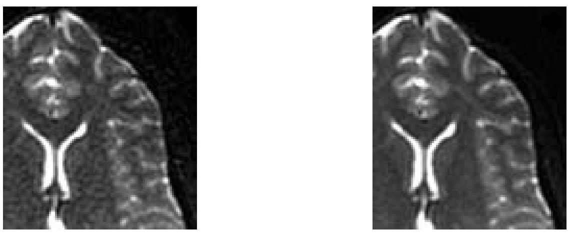


Fig. 6. Fusion of MR images from an EPI volume. Original (left) and filtered (right).

is about 28 times faster than MLE using the EM method [3]. In Fig. 5 the same measures are shown again for LMMSE, this time for a constant number of 30 scans, distributed in multiple samples and multiple sample directions. Although a larger number of gradients always improves the estimation, if only one sample is used the results show a bias when the SNR decreases.

Finally, to show the performance of the proposed scheme over real data, we have chosen a SENSE EPI data set, scanned in a 3.0 Tesla GE system, 51 gradient directions, 8 baselines, SENSE EPI. Voxel dimensions: 1.7 x 1.7 x 1.7 mm. We have selected an axial slice and fusing the 8 baselines using the LMMSE estimator. Results are on Fig. 6. Most of the noise in the original image has been removed after fusion. All the internal structures has been preserved, as well as the edges.

5 Conclusions

This paper introduced a method to estimate the magnitude signal from several acquisitions of both MRI and DWI based on the LMMSE estimator. To that end, a Rician assumption has been made. The proposed filtering method outperformed the average and the median methods, specially for moderate and high noise levels. Although a ML approach may have a slightly better performance in terms of error, the LMMSE provides a closed-form analytical solution, which is faster and more suitable when working with large amounts of data. Reducing the number of operations per pixel is a task of paramount importance when these large data sets must be dealt with. In addition, when working with DTI, the proposed method has also shown an important reduction in the bias of the first eigenvalue of the diffusion tensor, which makes some scalar measures like the Fractional Anisotropy more reliable.

Acknowledgements

The authors acknowledge the CICyT for research grants TIC2001-3808-C02-02 and TEC2004-06647-C03-01, the FIS for grants G03/135 and PIO-41483, the European Commission for the funds associated to the NoE SIMILAR (FP6-507609)m the MEC/Fulbright Comission for research Grant FU2005-0716 and NIH for grants R01MH074794 and P41RR13218.

References

1. Sijbers, J., den Dekker, A.J.: Maximum likelihood estimation of signal amplitude and noise variance form MR data. Magn. Reson. Imag. 51, 586–594 (2004)
2. Simon, M.K.: Probability distributions involving Gaussian random variables. Kluwer Academic Publishers, Boston, MA (2002)
3. DeVore, M.D., Lanterman, A.D., O'Sullivan, J.A.: ATR performance of a Rician model for SAR images. In: Proc. of SPIE 2000, ISSU 4050, Orlando, pp. 34–37 (2000)
4. Fillard, P., Arsigny, V., Pennec, X., Ayache, N.: Clinical DT-MRI estimation, smoothing and fiber tracking with Log-Euclidean metrics. In: IEEE International Symposium on Biomedical Imaging, pp. 786–789. IEEE Computer Society Press, Los Alamitos (2006)
5. Basu, S., Fletcher, T., Whitaker, R.: Rician noise removal in diffusion tensor MRI. In: Proceedings of MICCAI, vol. 1, pp. 117–125 (2006)
6. Kay, S.M.: Fundamentals of statistical signal processing: Estimation theory. Prentice Hall, New Jersey (1993)
7. Sijbers, J., den Dekker, A.J., Van Audekerke, J., Verhoye, M., Van Dyck, D.: Estimation of the noise in magnitude MR images. Magn. Reson. Imag. 16(1), 87–90 (1998)
8. 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)
9. Aja-Fernández, S., San-José-Estépar, R., Alberola-López, C., Westin, C.: Image quality assesment based on local variance. In: Proc. of the 28th IEEE EMBC, New York, USA, pp. 4815–4818. IEEE Computer Society Press, Los Alamitos (2006)
10. Westin, C.F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R.: Processing and visualization for diffusion tensor MRI. Medical Image Analysis 6, 93–108 (2002)
11. Stejskal, E.O., Tanner, J.E.: Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. Journal of Chemical Physics 42, 288–292 (1965)