Noise and Signal Estimation in Magnitude MRI and Rician Distributed Images: A LMMSE Approach

Santiago Aja-Fernández, Carlos Alberola-López, and Carl-Fredrik Westin, *Member, IEEE*

Abstract—A new method for noise filtering in images that follow a Rician model—with particular attention to magnetic resonance imaging—is proposed. To that end, we have derived a (novel) closed-form solution of the linear minimum mean square error (LMMSE) estimator for this distribution. Additionally, a set of methods that automatically estimate the noise power are developed. These methods use information of the sample distribution of local statistics of the image, such as the local variance, the local mean, and the local mean square value. Accordingly, the dynamic estimation of noise leads to a recursive version of the LMMSE, which shows a good performance in both noise cleaning and feature preservation. This paper also includes the derivation of the probability density function of several local sample statistics for the Rayleigh and Rician model, upon which the estimators are built.

Index Terms—Linear minimum mean square error (LMMSE) estimator, MRI filtering, noise estimation, Rician noise.

I. INTRODUCTION

NOISE in magnitude magnetic resonance (MR) images is usually modeled by means of a Rician distribution, due to the existence of zero-mean uncorrelated Gaussian noise with equal variance in both the real and imaginary parts of the complex k-space data [1], [2]. This noise may affect the performance of different postprocessing techniques applied to MR data, such as segmentation, registration or tensor estimation in diffusion tensor MRI (DT-MRI) [3]. Accordingly, a great amount of noise-removal methods has been reported in the literature.

One of the first attempts proposed to estimate the magnitude MR image from a noisy image is due to Henkelman [4] who investigated the effect of noise on MR magnitude images. The author showed that the noise influence leads to an overestimation of the signal amplitude and proposed a correction scheme based on image intensities. The so-called *conventional approach* (CA) was proposed by McGibney *et al.* [5] utilizing the noise properties of the second-order moment. Sijbers *et al.* [6]–[8] estimate the Rician noise level and perform signal reconstruction using a maximum likelihood (ML) approach. A similar method is used by Jiang and Yang [9]. Expectation maximization (EM) formulations with Rician noise assumptions have been used in SAR imaging [10], [11]. Fillard *et al.* [12] recently proposed another ML-based scheme that uses an anisotropic regularization term to exploit the spatial correlation, and apply it to correct the bias in tensor estimation in DT-MRI.

Other approaches use wavelet-based methods for noise removal, as Nowak's [13]—in which the authors assume an underlying Rician model—or the one due to Pižurica *et al.* [14]. Lysaker *et al.* [15] proposed a method for noise removal in MRI using fourth-order partial differential equations. McGraw *et al.* [16] use a weighted total-variation-norm denoising scheme and Ahn *et al.* [17] propose a template-based filtering procedure; none of this approaches use a Rician noise model. Basu *et al.* [3] use a Perona–Malik-like smoothing filter combined with a local Rician data attachment term (effectively trying to remove the intensity bias locally), assuming a known noise level for the Rician noise model. Using local statistics as priors, Awate and Whitaker [18] introduced a method to denoise a MR image using nonparametric neighborhood statistics. Recently, Koay and Basser in [19] developed a correction scheme that allows to analytically estimate the signal, also assuming the Rician model, and in [20] Martin-Fernandez *et al.* proposed a Wiener-filter approach with a correction for MRI data.

In this paper we propose a new method for noise filtering of MRI data, and by extension, of images that follows a Rician distribution, by using the linear minimum mean square error (LMMSE) estimator for Rician noise.

As it is well-known, LMMSE estimators are a tradeoff between optimality and simplicity; we obtain a closed-form analytical solution for our estimator which makes the filtering process far simpler than other estimation techniques that find the solution via an iterative optimization scheme. Results from our method are satisfactory as the experiments in Section V show. Additionally, the goodness of the filter will be intrinsically related to the accuracy of the estimation of the noise variance in the images. In the paper we will present four novel techniques to estimate this parameter; although three of them rely on the assumption of a uniform background, none of the cases require any sort of background segmentation.

The paper is organized as follows. Section II is a background section on Rician and Rayleigh distributions and their estimators. In Section III, the LMMSE estimator for Rician noise is presented. Then, Section IV describes the different procedures to estimate the noise. Some experiments have been added in Section V. A number of appendices have been included for readability purposes.

S. Aja-Fernández is with the LPI, ETSI Telecomunicación, Universidad de Valladolid, and also with LMI, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215 USA.

C. Alberola-López is with the LPI, ETSI Telecomunicación, Universidad de Valladolid, Valladolid, Spain.

C.-F. Westin is with the LMI, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02215 USA.

II. BACKGROUND

A. Rician Distributed Data

If both real and imaginary parts of a signal are corrupted with zero-mean uncorrelated Gaussian noise with equal variance, the envelope of the magnitude signal will follow a Rician distribution. This kind of noise may be encountered in many practical situations, such as MR data, speckle [10], communication fading problems [21], and many others. The magnitude signal may be expressed

$$M = \sqrt{(A + n_1)^2 + n_2^2} \qquad (1)$$

being M the magnitude image, A the original signal level if no noise is present, and n_1 and n_2 uncorrelated Gaussian noise variables with zero mean and equal variance σ_n^2. The probability density function (PDF) of such an image is, as indicated, a Rician distribution [1], [22]

$$p_M(M \mid A, \sigma_n) = \frac{M}{\sigma_n^2} e^{-\frac{M^2 + A^2}{2\sigma_n^2}} I_0\left(\frac{AM}{\sigma_n^2}\right) u(M) \qquad (2)$$

being $I_0(\cdot)$ the 0th-order modified Bessel function of the first kind and $u(\cdot)$ the Heaviside step function. Although, generally speaking, the moments of this distribution are difficult to calculate, the even-order (non central) moments are simple polynomials as, for instance, the second-order moment

$$\mu_2 = E\{M^2\} = A^2 + 2\sigma_n^2 \qquad (3)$$

In the image background, where the signal-to-noise ratio is zero due to the lack of water-proton density in the air, the Rician PDF simplifies to a Rayleigh distribution [22] with PDF

$$p_M(M \mid \sigma_n) = p_M(M \mid A = 0, \sigma_n) = \frac{M}{\sigma_n^2} e^{-\frac{M^2}{2\sigma_n^2}} u(M). \quad (4)$$

B. Parameter Estimation of the Rician Distribution

In this section, we review a number of methods to estimate parameters in a Rician distribution. Although many approaches have been reported in the literature (see, for instance, the papers cited in the introduction) we will focus now only on those based on a stochastic model. Most of the papers cited hereafter focus on MRI processing, although some of them deal with other type of images, such as SAR [10].

The *conventional approach* (CA) [5] accounts for the relation between noise and signal of the second-order moment in a Rician distribution, see (3), so the signal can be estimated as

$$\hat{A} = \sqrt{\langle M^2 \rangle - 2\sigma_n^2} \qquad (5)$$

being $\langle M^2 \rangle$ the sample second-order moment. Other approaches use the *maximum likelihood estimator* (ML) [6]–[9] where the signal is estimated by maximizing the likelihood function. Koay and Basser in [19] propose an analytically exact solution assuming a Rician noise model based on image statistics and the noise level. In these three methods the variance of noise σ_n^2 is an input value, so it must be known before hand, or at least, it must be estimated. The *expectation-maximization* (EM) method [10], [11] provides a recursive scheme that aims

at estimating the noise variance and the signal simultaneously through maximization of the expected log likelihood.

Note that the last three schemes (ML, EM, and Koay's) are originally designed to estimate the signal from multiple samples. When only one image is available, the statistics must be calculated locally. In addition, in the three methods the solution is found via an iterative optimization scheme.

The different methods reported to estimate σ_n^2 from the magnitude data in Rician distributions may be roughly divided into two vast groups, namely, those that use a single magnitude image and those that use multiple images. In this paper we will mainly deal with methods in the former group.

The estimation of noise in MRI using a single image is usually done out of the background pixels, where the signal is assumed to be zero. According to (3), in the areas where the signal is zero $A^2 = 0$, and then $E\{M^2\} = 2\sigma_n^2$. So a straightforward estimator—based on the method of moments—will be [7], [13], [23]

$$\widehat{\sigma_n^2} = \frac{1}{2N} \sum_{i=1}^{N} M_i^2 \qquad (6)$$

being N the number of points considered for the estimation. Equation (6) is in fact, an unbiased estimator of σ_n^2 with variance

$$\text{Var}\left(\widehat{\sigma_n^2}\right) = \frac{\sigma_n^4}{N}. \qquad (7)$$

This estimator is also the maximum likelihood estimator of this parameter for the Rayleigh distribution [7].

An alternative way to estimate σ_n in nonsignal areas is obtained by using the first-order moment of the Rayleigh PDF. An unbiased estimator then will be [7]

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}} \frac{1}{N} \sum_{i=1}^{N} M_i. \qquad (8)$$

In [9], the authors propose an estimator based on the method of moments for a Rician distribution

$$\widehat{\sigma_n^2} = \frac{1}{2}(\langle M^2 \rangle - (2\langle M^2 \rangle^2 - \langle M^4 \rangle)^{1/4}).$$

Recently, Sijbers *et al.* [24], [25] proposed a new method to estimate σ_n from the image histogram. As in the background the image distribution is Rayleigh, and since the mode of this distribution equals the parameter σ_n, then $\widehat{\sigma_n} = M_{\max}$, being $M_{\max}$ the mode of the histogram of the image. Noise can also be estimated using multiple images taken under identical conditions [23]. Finally, in [26] noise is estimated by using the information of several images and how they match a model.

III. SIGNAL ESTIMATION: THE LMMSE ESTIMATOR

The method proposed in this paper to estimate the signal from the magnitude image is based on the linear minimum mean square error (LMMSE) estimator [27]. Our aim is to find a closed-form estimator for a signal that follows a Rician distribution. This is in contrast to many other estimation techniques—such as ML and EM—which find the solution via an iterative optimization scheme. Closed-form solutions make estimation methods computationally more efficient than

optimization-based solutions. In addition, the estimator will be based on local statistics, so one single (noisy) image—as opposed to several perfectly aligned images—will suffice to estimate the signal.

As previously stated, the moments of the Rician distribution have a nontrivial integral expression but for even-order moments, which are simple polynomials. In order to achieve a closed-form expression we will use A^2 instead of A. Consequently, all the moments to be used hereafter will be even.

A. LMMSE Filter Derivation for the General Rician Model

The LMMSE estimator[1] of a parameter θ is defined [27]

$$\hat{\theta} = E\{\theta\} + \mathbf{C}_{\theta\mathbf{x}}\mathbf{C}_{\mathbf{xx}}^{-1}(\mathbf{x} - E\{\mathbf{x}\}) \qquad (9)$$

being $\mathbf{x}$ the vector of available samples, $\mathbf{C}_{\mathbf{xx}}$ the covariance matrix of $\mathbf{x}$ and $\mathbf{C}_{\theta\mathbf{x}}$ the cross-covariance vector. Rewriting (9) for a 2-D signal with a Rician distribution

$$\widehat{A_{ij}^2} = E\{A_{ij}^2\} + \mathbf{C}_{A_{ij}^2 M_{ij}^2}\mathbf{C}_{M_{ij}^2 M_{ij}^2}^{-1}\left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\}\right) \quad (10)$$

where A_{ij} is the unknown intensity value in pixel ij and $\mathbf{M}_{ij}$ the observation vector. If the estimator is simplified to be pointwise, vectors and matrices become scalar values. Then

$$\begin{aligned}
&\mathbf{C}_{M_{ij}^2 M_{ij}^2} \\
&= E\left\{\left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\}\right)\left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\}\right)^T\right\} \\
&= E\left\{\left(M_{ij}^2 - E\{M_{ij}^2\}\right)^2\right\} = E\{M_{ij}^4\} - E\{M_{ij}^2\}^2 \\
&= C_{M_{ij}^2 M_{ij}^2} \\
&\mathbf{C}_{A_{ij}^2 M_{ij}^2} \\
&= E\left\{\left(A_{ij}^2 - E\{A_{ij}^2\}\right)\left(\mathbf{M}_{ij}^2 - E\{\mathbf{M}_{ij}^2\}\right)^T\right\} \\
&= E\left\{\left(A_{ij}^2 - E\{A_{ij}^2\}\right)\left(M_{ij}^2 - E\{M_{ij}^2\}\right)\right\} \\
&= C_{A_{ij}^2 M_{ij}^2}
\end{aligned}$$

and making use of the model in (1)

$$C_{A_{ij}^2 M_{ij}^2} = E\{A_{ij}^4\} + 2E\{A_{ij}^2\}\sigma_n^2 - E\{A_{ij}^2\}E\{M_{ij}^2\}.$$

Finally, the LMMSE estimator is

$$\begin{aligned}
\widehat{A_{ij}^2} = \; & E\{A_{ij}^2\} \\
& + \frac{E\{A_{ij}^4\} + 2E\{A_{ij}^2\}\sigma_n^2 - E\{A_{ij}^2\}E\{M_{ij}^2\}}{E\{M_{ij}^4\} - E\{M_{ij}^2\}^2} \\
& \times \left(M_{ij}^2 - E\{M_{ij}^2\}\right).
\end{aligned} \qquad (11)$$

Assuming local ergodicity, the expectation may be replaced by its sample estimator $\langle \cdot \rangle$, that can be defined

$$\langle I_{ij}\rangle = \frac{1}{|\eta_{ij}|}\sum_{p\in\eta_{ij}} I_p \qquad (12)$$

[1]As it is well known, this type of estimators model the parameter to be estimated as a sample of a random variable, the parameters of which have a known relation with the parameters of the observation [27].

with η_{ij} a square neighborhood around the pixel ij. This estimation may be also done using nonsquare weighted windows, such as Gaussian functions. Using the relations from (1)

$$\begin{aligned}
E\{M_{ij}^2\} &= E\{A_{ij}^2\} + 2\sigma_n^2 \\
E\{M_{ij}^4\} &= E\{A_{ij}^4\} + 8\sigma_n^2 E\{A_{ij}^2\} + 8\sigma_n^4
\end{aligned}$$

and $\langle \cdot \rangle$, the LMMSE estimator may finally be written as

$$\widehat{A_{ij}^2} = \langle M_{ij}^2\rangle - 2\sigma_n^2 + K_{ij}\left(M_{ij}^2 - \langle M_{ij}^2\rangle\right) \qquad (13)$$

with K_{ij}

$$K_{ij} = 1 - \frac{4\sigma_n^2\left(\langle M_{ij}^2\rangle - \sigma_n^2\right)}{\langle M_{ij}^4\rangle - \langle M_{ij}^2\rangle^2}. \qquad (14)$$

Note that the σ_n^2 value must be properly estimated. This task is usually done from a selected region from the background pixels, where the signal is assumed zero. In Section IV, some new automatic methods will be presented.

Since only pointwise dependence has been considered in the filter, the extension to an arbitrary number of dimensions is straightforward by changing the estimation neighborhood. For example, for 3-D images

$$\langle I\rangle = \frac{1}{|\eta_{ijk}|}\sum_{p\in\eta_{ijk}} I_p. $$

Some experiments have been done in Section V-A to illustrate the LMMSE filtering performance.

IV. ESTIMATION OF NOISE USING LOCAL STATISTICS

The performance of the estimator previously described, as well as other methods in the literature, is directly related to the quality of the estimate of the noise variance σ_n^2. Noise estimation schemes described in Section II-B (but the one in [24]) have some disadvantages: they need the background pixels of the image to be manually selected and the estimation is done considering that the signal is always zero in the background pixels. These two considerations make the estimate sensitive to errors and artifacts. In this section, we will propose a new approach based on local statistics. We will estimate the noise using the distributions of some sample local statistics of the image, such as the sample second-order moment, the sample mean and the sample variance.

Using the second-order moment and the mean we will develop two estimators that can be used only when the image has a background where the distribution can be assumed Rayleigh (i.e., a nonsignal background). This is the case of many MR magnitude images. Then, we will define a variance-based estimator that can be used with any kind of image that follows a Rician distribution.

A. Noise Estimator Based on the Local Second-Order Moment

Taking expectations in (1) for zero-mean and equal variance noise components, we obtain

$$E\{M_{ij}^2\} = E\{A_{ij}^2\} + 2\sigma_n^2 \qquad (15)$$

Fig. 1.   Original image (left) and noisy counterpart (right) with $\sigma_n = 10$.

Fig. 3.   Original magnitude MRI from the BrainWeb database (left) and noisy one (right) with $\sigma_n = 10$.
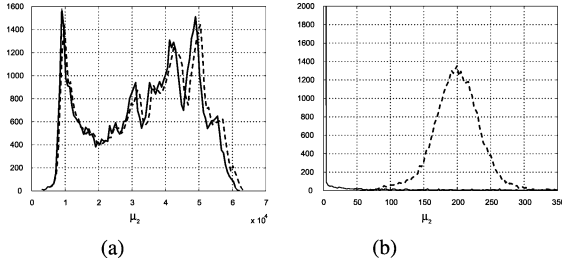

Fig. 2.   Distribution of the local second-order moment. (a) Lena image without noise (solid) and with Rician noise with $\sigma_n = 10$ (dashed). (b) MR image without noise (solid) and with noise (dashed), with $\sigma_n = 10$. A 7×7 window has been used.

being $E\{I_{ij}^2\}$ the local second-order moment of an image I. It may be estimated using a neighborhood η_{ij} centered around the pixel under analysis giving rise to $\widehat{\mu_{2}}_{ij} = \widehat{E\{I_{ij}^2\}} = \langle I_{ij}^2 \rangle$. For the sake of simplicity we will use a square neighborhood, as in (12).

If we assume that σ_n is constant throughout the image, the effect of the noise over the local second-order moment distribution will be a shift to higher values. In those regions where $A_{ij} = 0$, it holds

$$E\left\{M_{ij}^2\right\} = 2\sigma_n^2.$$

This fact has been used in the past to estimate the noise from MRI by selecting areas in the background (see Section II-B). In this new approach, we will not be working with selected regions, but with the distribution of the local statistics of the whole image.

As an effect of the noise, the shape of the local second-order moment distribution remains fairly unchanged in the noisy image with respect to that in the original image, but for a right-shift of the whole distribution. See, for instance, the distribution of this moment for the image in Fig. 1 when Rician noise is present; this is depicted in Fig. 2(a).

However, in MR images—see, for instance, Fig. 3 taken from the BrainWeb database [28]—due to the presence of an extensive background in which the signal is virtually zero, the distribution of the local second-order moment of the original image will have a maximum in the origin, as shown in Fig. 2(b) solid line. If the image is corrupted with Rician noise with $\sigma_n = 10$, the effect, according to (15), will be a shift of the maximum from zero to $2\sigma_n^2$. This effect can be seen in Fig. 2(b), dashed line, where the distribution of the local second-order moment shows a bell shape with its maximum in 200, with in fact is $2\sigma_n^2$ when $\sigma_n = 10$.
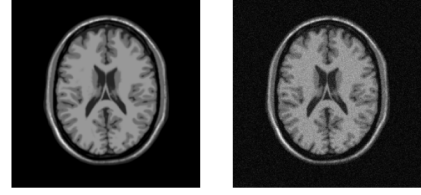
So, the position of this maximum may be used in order to estimate σ_n^2

$$\widehat{\sigma_n^2} = \frac{1}{2}\arg\max\{p(\widehat{\mu_2})\}$$

being $p(\widehat{\mu_2})$ the distribution of the sample (local) second-order moment of the noisy image. This maximum may be calculated as the mode of the distribution.

To analyze this assumption let us calculate the theoretical value of the mode. In MR images, the pixels in the background follow a Rayleigh distribution. So, let $R_i(\sigma^2)$, $i = \{1, \ldots, N\}$ be a set of independent and identically distributed (IID) Rayleigh random variables. Then [29]

$$\sum_{i=1}^{N} R_i^2(\sigma^2) \sim \gamma(N, 2\sigma^2)$$

$$S = \frac{1}{N}\sum_{i=1}^{N} R_i^2(\sigma^2) \sim \gamma\left(N, \frac{2\sigma^2}{N}\right)$$

i.e., the sample local second-order moment of a Rayleigh distribution follows a Gamma distribution with parameters $\alpha = N$ and $\beta = ((2\sigma^2)/N)$. The mode of this distribution is defined as $\mathrm{mode}(S) = (\alpha - 1)\beta$, so, the mode of the sampling second-order moment will be

$$\mathrm{mode}(\widehat{\mu_2}_{ij}) = \frac{N-1}{N}2\sigma_n^2.$$

If we redefine the estimator of the moment as

$$\widehat{\mu_2} = \frac{1}{N-1}\sum_{i=1}^{N} R_i^2(\sigma^2) \qquad (16)$$

then

$$\mathrm{mode}(\widehat{\mu_2}_{ij}) = 2\sigma_n^2$$

and

$$\widehat{\sigma_n^2} = \frac{1}{2}\mathrm{mode}(\widehat{\mu_2}_{ij}). \qquad (17)$$

To verify this hypothesis, an experiment has been carried out for 40 different variances of noise. The MR image in Fig. 3 has been normalized between 0 and 1. Local moments are calculated using 7×7 windows. For each σ_n^2, we would take the average value of 100 simulations. In Fig. 4(a), the estimated variance is depicted versus the original one. In Fig. 4(b), as a quality measure, we show the ratio $Q = \widehat{\sigma_n}/\sigma_n$.
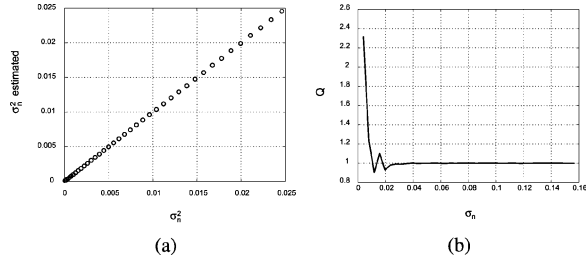
Fig. 4. Estimation of variance of noise using the mode of the local second-order moment. (a) $\widehat{\sigma_n^2}$ versus σ_n^2. (b) $\widehat{\sigma_n}/\sigma_n$ versus σ_n.
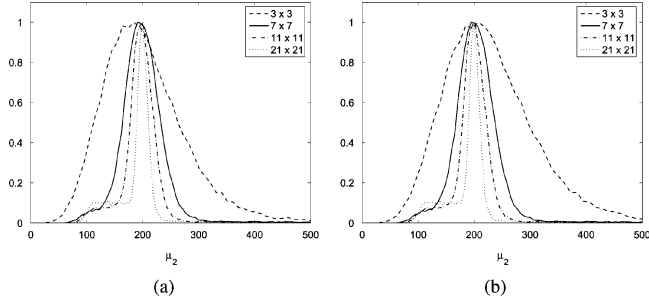

Fig. 5. (Normalized) Local second-order moment distribution of the noisy image with $\sigma_n = 10$ using 3×3 window (dashed), 7×7 window (solid), 11×11 window (dash-dotted), and 21×21 window (dotted). (a) Normalizing by N. (b) Normalizing by $N - 1$, as in (16).
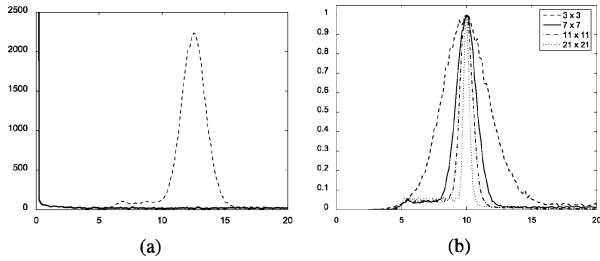

Fig. 6. (a) Sample local mean distribution of the image without noise (solid-bold) and with noise (dashed), with $\sigma_n = 10$. (b) Normalized distribution of $\sqrt{(2/\pi)}\widehat{\mu_1}_{ij}$ of the noisy image with $\sigma_n = 10$. Window size: 3×3 (dashed), 7×7 (solid), 11×11 (dash-dotted), and 21×21 (dotted).

Let us now study how the window size affects the estimation. According to (7), the variance of the estimator is inversely proportional to the number of points used. So, as far as ergodicity is maintained within the window, the estimation should be better for larger window sizes. However, the influence of the window size is not paramount in this new estimator, as only the maximum of the distribution is considered.

As an illustration, we have calculated the local second-order moment distribution of the noisy image using square windows of different sizes. The result is in Fig. 5. The effect of increasing the size of the window is that the function goes narrower, as an effect of the decrease in the variance. However, its maximum is fixed to the same value.

B. Estimator Based on the Local Mean

Following a similar reasoning to the one done in the previous section, we can define a local measure based on the estimator of

(8). This estimator is also based on the assumption that the distribution in the background of an MR image follows a Rayleigh distribution. Due to the influence of the pixels in this background, when the local mean distribution of a non-noisy MRI is depicted, it presents a maximum around zero [see Fig. 6(a), solid line]. When the image is corrupted with Rician noise, this maximum is now shifted to a value related to σ_n [see the dotted line in Fig. 6(a)]. As the mean of a Rayleigh distribution is defined as $\mu_1 = \sigma_n\sqrt{\pi/2}$, following a similar reasoning as the one done in the previous section, we can define a new estimator

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}} \arg\max\{p(\widehat{\mu_1})\}$$

being $p(\widehat{\mu_1})$ the distribution of the sample local mean. The maximum may be calculated as the mode of the distribution

$$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}} \text{mode}(\widehat{\mu_1}_{ij}). \tag{18}$$

To study this assumption, let us analyze the PDF of the sample local mean. Let $R_i(\sigma^2), i = \{1, \ldots, N\}$ be a set of IID random variables with Rayleigh distribution and

$$S = \sum_{i=1}^{N} R_i(\sigma^2).$$

The sum of Rayleigh variables is a classical-hard-to-find problem in communications. Some approximations are usually employed, as the one in [30]

$$p_S(x) = \frac{x^{2N-1}e^{-x^2/2bN}}{2^{N-1}N^N b^N (N-1)!} \tag{19}$$

$$b = \frac{\sigma^2}{N}[(2N-1)!!]^{1/N} \tag{20}$$

which can be approximated by $b \approx \sigma^2(2/e)$ (see Appendix A). The mode of the distribution is defined as $\text{mode}(S) = \arg\max_x\{p_S(x)\}$, so

$$\text{mode}(S) = \sigma_n\sqrt{\frac{2(2N-1)N}{e}}$$

which can be approximated when $N \gg 1$ by $\text{mode}(S) \approx \sigma_n(2N/\sqrt{e})$. Therefore, as the sample mean is defined $\widehat{\mu_1} = (1/N)S$ the noise estimator becomes

$$\widehat{\sigma_n} \approx \frac{\sqrt{e}}{2} \text{mode}(\widehat{\mu_1}_{ij}). \tag{21}$$

Note that in this approximation, the factor is $(\sqrt{e}/2) = 0.8244$, and in (18) the factor is $\sqrt{(2/\pi)} = 0.7979$. This small difference is due to both the approximation in the PDF and to the simplifications made to obtain the mode. In what follows we will use the estimator in (18), as it experimentally shows a better behavior.

To test this estimator, a new experiment has been carried out for 40 different standard deviations of noise. The MR image has been normalized between 0 and 1. Local moments are calculated using 7×7 windows. For each σ_n, we would take the average value of 100 simulations. In Fig. 7(a), the estimated standard
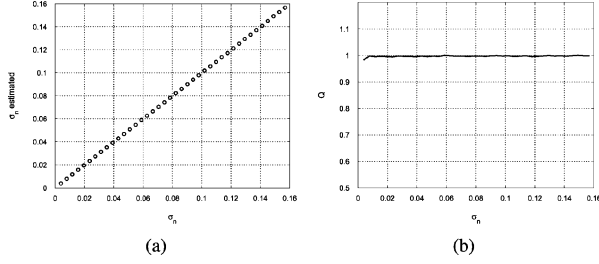
Fig. 7. Estimation of standard deviation of noise using the mode of $\widehat{Q_{E_{ij}}}$. (a) $\widehat{\sigma_n}$ versus σ_n. (b) $Q = \widehat{\sigma_n}/\sigma_n$ versus σ_n.
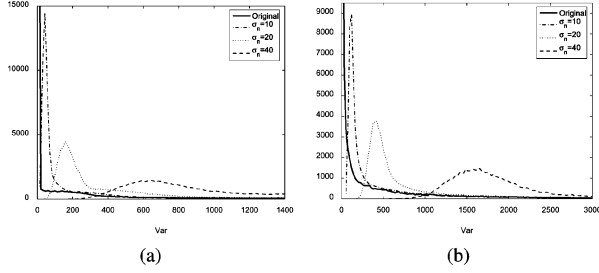

Fig. 8. Sample local variance distribution. (a) MRI with Rician noise. (b) Lena Image with Rician noise. Original image (solid), Rician noise with $\sigma_n = 10$ (dash-dotted), with $\sigma_n = 20$ (dotted) and with $\sigma_n = 40$ (dashed).

deviation is depicted versus the original one. In Fig. 7(b), as a quality measure, we show the ratio $\widehat{\sigma_n}/\sigma_n$.

As in the previous case, the effect of changing the size for the window used for estimation—see Fig. 6(b)—is a change in the width of the distribution, but in any case, the maximum stays on the same value.

C. Estimation Based on the Local Variance

The estimators previously introduced require a Rayleigh distributed area, i.e., an area in which $A_{ij} = 0$. In this section, we introduce a new estimator that can be used in signals with general Rician noise. However, we will first study the case of images with zero background.

The distribution of the local variance for most non-noisy images, is a decreasing function with its maximum in zero (see Appendix B for details). When corrupted with noise, the distributions present a *bell-shape distribution* whose maximum is located in a point that should be related to σ_n^2. An illustration is depicted in Fig. 8. We will take advantage of this feature for both estimators.

1) Local Variance in MR Data: In the background of MR data, the variance reduces to the variance of a Rayleigh distribution, which is defined

$$\sigma_M^2 = \sigma_n^2 \left(2 - \frac{\pi}{2}\right). \tag{22}$$

In the pixels belonging to the background then

$$\sigma_n^2 = \sigma_M^2 \left(2 - \frac{\pi}{2}\right)^{-1}. \tag{23}$$

If the background of the image is properly segmented, it can be used to estimate the variance of noise using (23). However, the segmentation is not necessary if we follow a philosophy as the one stated in the previous sections. If the variance is estimated locally, say $\sigma_{M_{ij}}^2$, we can define the variance of noise as

$$\widehat{\sigma_n^2} = \left(2 - \frac{\pi}{2}\right)^{-1} \text{mode}\left(\sigma_{M_{ij}}^2\right). \tag{24}$$

Actually, the maximum values of the distributions in Fig. 8(b) are located at in $(2 - (\pi/2))\sigma_n^2$.

To give a theoretical justification to this solution, it would be interesting to study the distribution of the sample variance when the samples are Rayleigh random variables (if they are Gaussian, the result is well known [29]). The distribution of the sample local variance for Rayleigh random variables is (see Appendix C for details)

$$p_V(x) = C_V e^{-\frac{xN}{2\sigma^2}} \sum_{k=0}^{N-1} \frac{(2N-2-k)!}{k!(N-1-k)!} \left[\left(1 + \frac{e}{2}\right) \frac{xN}{2\sigma^2}\right]^k \tag{25}$$

being

$$C_V = \frac{N}{2\sigma^2} \frac{1}{(N-1)!} \frac{(2/e)^{N-1}}{(1+2/e)^{2N-1}}.$$

In order to find the maximum independently of the value of σ, we make the change $t = (xN/2\sigma^2)$

$$f_V(t) = C_V e^{-t} \sum_{k=0}^{N-1} \frac{(2N-2-k)!}{k!(N-1-k)!} \left[\left(1 + \frac{e}{2}\right) t\right]^k. \tag{26}$$

We have numerically calculated the maximum of this function which is

$$t_{\max} \approx 0.26 \times (N-1) \approx \left(1 - \frac{2}{e}\right)(N-1).$$

In Fig. 9, the function in (26) has been depicted for $N = 25$ (dash-dotted), $N = 49$ (dashed) and $N = 81$ (solid). To avoid the dependence with N, the x axis has been normalized by $N - 1$. The maxima of the functions for the different values of N are all located on the same point, around 0.26. So, the mode of the distribution, after the variable change, will be

$$\text{mode}(V) = x_{\max} \approx 0.52 \frac{N-1}{N} \sigma^2 \approx \left(2 - \frac{4}{e}\right) \frac{N-1}{N} \sigma^2.$$

Note that the N in the denominator is due to the fact that we have used the biased variance. Similarly as in the case of the sample mean, due to the approximation, the solution is not exactly what it should be, as $(2 - (\pi/2)) \approx 0.42$. Previously, in (21), we have seen that $\sqrt{e}/2$ was in fact an approximation of $\sqrt{2/\pi}$, due to some approximation in the analysis. Note that the parameter we are using for the variance analysis is now $e/2 = 2(\sqrt{e}/2)^2$. If we use $2(\sqrt{2/\pi})^2 = 4/\pi$ instead, the numerical analysis sets the maximum in

$$t_{\max} \approx 0.21(N-1) \approx \left(1 - \frac{\pi}{4}\right)(N-1)$$

being the mode

$$\text{mode}(V) \approx 0.42 \frac{N-1}{N} \sigma^2 \approx \left(2 - \frac{\pi}{2}\right) \frac{N-1}{N} \sigma^2.$$
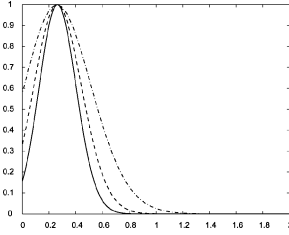
Fig. 9. (Normalized) function of (26) for $N = 25$ (dash-dotted), $N = 49$ (dashed) and $N = 81$ (solid). The x axis has been normalized by $N - 1$.
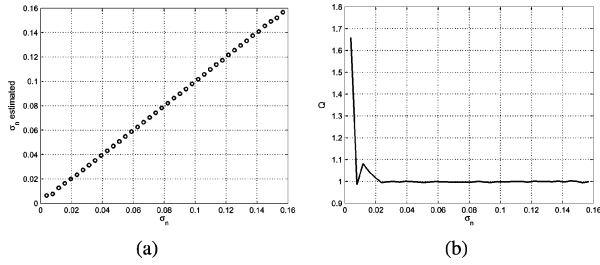

Fig. 11. Normalized local variance distribution of the Lena image with Rician Noise ($\sigma_n = 15$) using 5×5 (dashed), a 9×9 (solid), and 21×21 (dash-dotted) windows. $(N - 3)/(N - 1)$ correction has been applied.

Fig. 10. Estimation of standard deviation of noise using the mode of the sample variance. (a) $\widehat{\sigma_n}$ versus σ_n. (b) $Q = \widehat{\sigma_n}/\sigma_n$ versus σ_n.

Due to experimental results, we define the estimator as in (24).

To test this estimator, a new experiment has been carried out for 40 different standard deviation of noise. The MR image has been normalized between 0 and 1. Local moments are calculated using 7×7 windows. For each σ_n, we would take the average value of 100 simulations. In Fig. 10(a), the estimated standard deviation is depicted versus the original one. In Fig. 10(b), as a quality measure, we show the ratio $\widehat{\sigma_n}/\sigma_n$.

2) Local Variance in Rician Images: Let us now focus on those images with Rician noise in which the assumption of a low SNR background does not hold. As a result, the variance of noise cannot be estimated using the second-order moment, nor the local mean, but it can be done using the variance.

Although most of the images we usually deal with are images that do not have a uniform background, they do have a great amount of uniform areas. By *uniform area*, we mean areas of the image without borders and with soft transitions. This means that, if no texture is present, the distribution of the local variances will have its maximum value in the vicinity of zero. To illustrate this statement, see for instance the sample local distributions of Lena and the MR (non-noisy) images in Fig. 8 (solid line), where it can be seen that most values are nearly zero, as expected. This exponentially decreasing histogram may be observed in a wide range of images, from natural scenes to medical images (see Appendix B). An exception to this is obviously the case of pictures of complex textures, for which intensity variability is considered signal content. We must remark that this assumption does not mean that the estimation is to be done in one specific area assumed constant; the estimation will be done using the whole variance histogram and it will not require any previous selection of any area.
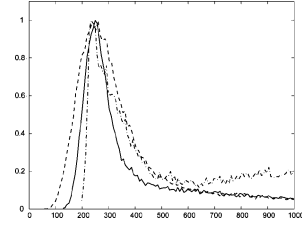
With this assumption in mind, we will first approximate the variance of a Rician model, to afterwards study the sample variance. The mean of a Rician distribution is given by

$$\mu_1 = E\{M\} = \sigma_n \sqrt{\frac{\pi}{2}} e^{-\frac{A^2}{4\sigma_n^2}}$$
$$\times \left[\left(1 + \frac{A^2}{2\sigma_n^2} \right) I_0 \left(\frac{A^2}{4\sigma_n^2} \right) + \frac{A^2}{2\sigma_n^2} I_1 \left(\frac{A^2}{4\sigma_n^2} \right) \right]. \quad (27)$$

In order to obtain an approximation of the behavior of the variance of the image, we will consider the asymptotic expansion of the Bessel functions [31]

$$I_n(z) \approx \frac{e^z}{\sqrt{2\pi z}} \left(1 - \frac{4n^2 - 1}{8z} \right), \quad z \to \infty. \quad (28)$$

It implies that $A_{ij}^2 \gg \sigma_n^2$. After some algebra

$$\sigma_M^2 \approx \sigma_n^2 \left(1 - \frac{5}{4} \frac{\sigma_n^4}{A^2} - \cdots \right). \quad (29)$$

A natural image, like Lena, has a distribution of $E\{A_{ij}^2\}$ as the one in Fig. 8(a). This means that if we take local statistics $E\{A_{ij}^2\} > \sigma_n^2$ for the great majority of the pixels, and even $E\{A_{ij}^2\} \gg \sigma_n^4$. So, in those points, we can say that

$$\sigma_{M_{ij}}^2 \approx \sigma_n^2 \quad (30)$$

and, accordingly, if the variance of the signal is estimated locally, the variance of noise can be estimated as

$$\widehat{\sigma_n^2} = \text{mode}\left(\sigma_{M_{ij}}^2 \right). \quad (31)$$

The maximum values of the distributions in Fig. 8(a) are located at σ_n^2.

What we have done here is similar to consider the rough approximation of the noise being Gaussian when the SNR is high. This way, the sample variance has a gamma distribution [29] with mode

$$\text{mode}(\sigma_M^2) = \sigma_n^2 \frac{N - 3}{N - 1}.$$

In Fig. 11, the local variance distribution of a noisy image has been calculated using square windows of different sizes: 5×5 (dashed), 9×9 (solid) and 21×21 (dash-dotted) applying the $(N - 3)/(N - 1)$ correction.
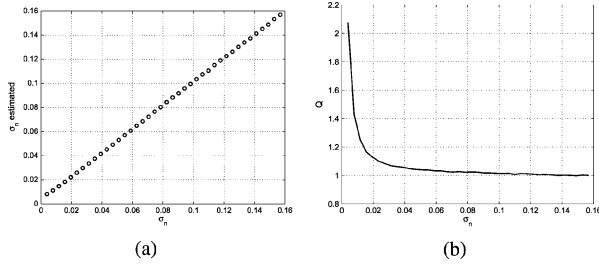
Fig. 12. Estimation of variance of noise using the mode of the local sample variance in the Lena image. (a) $\widehat{\sigma_n}$ versus σ_n. (b) $\widehat{\sigma_n}/\sigma_n$ versus σ_n.

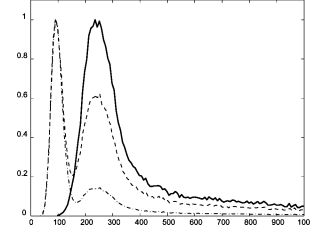Fig. 13. Noise image with $\sigma_n = 15$ and different background size added.

Fig. 14. Normalized local variance distribution of the Lena image with Rician noise. Without extra background (solid) with background (dashed) and with a greater background (dash-dotted).

To completely justify this result, we need to calculate the distribution of the sample variance for the Rician model. This is not an easy task, though some approximations have been done in Appendices D and E, where we show that the sample variance PDF for this model may be written as

$$p_V(t) \approx C_T e^{-t} \sum_{m=0}^{\frac{2N-1}{4}} \binom{\frac{2N-1}{4}}{m} \times t^m (1+N)^m \Gamma(2N - 1/2 - m)$$

with $t = (xN)/(2\sigma_n^2)$. Numerically, it is easy to proof that $t_{\max} = ((N-3)/2)$. So the mode of the variance distribution will be

$$\text{mode}(V) = \frac{N-3}{2} \frac{2\sigma_n^2}{N} = \frac{N-3}{N} \sigma_n^2.$$

Note that if the unbiased sample estimator is used, $\text{mode}(V) = (N-3)/(N-1)\sigma_n^2$.

To verify this hypothesis an experiment has been carried out for 40 different variances of noise. The Lena image has been normalized between 0 and 1. Local variance is calculated using 7×7 windows. For each σ_n we would take the average value of 100 simulations. In Fig. 12(a), the estimated variance is depicted versus the original one. In Fig. 12(b), as a quality measure, we show the ratio $\widehat{\sigma_n}/\sigma_n$.

D. Connection Between the Two Variance Models

Two models have been presented for signal estimation based on the mode of the local sample variance. The first one is based on the assumption that the image has a background where the Rician model tends to be Rayleigh; the second one is for images without such a background.

To study the behavior of the local variance in relation with the background, an experiment has been done. A black background has been added to the Lena image, having a MR-like image. Rician noise with $\sigma_n = 15$ has been added (see Fig. 13). The distribution of the local variance is shown in Fig. 14. We can see that there is a maximum set around σ_n^2, as the second model stated, but in addition, a new maximum appears when the background is added. This new maximum will be set on $\sigma_n^2(2 - (\pi/2))$, the value predicted by the first model. So, depending on the background, the local variance distribution will be a multimodal distribution, being the local maxima placed on $\sigma_n^2(2 - (\pi/2))$ and σ_n^2, respectively. The global maxima will be in the mode with the highest coverage.

V. EXPERIMENTAL RESULTS

A. Synthetic Experiments

Some experiments are carried out to illustrate the LMMSE filtering behavior. To be able to compare the results to a *ground truth*, we work with synthetic images artificially corrupted with noise. A magnitude MR image originally noise-free (from the BrainWeb database [28]) with 256 gray levels, is corrupted with Rician noise, see Fig. 15(a) and (b), following the model in (1). The noisy image is processed using different techniques. Many methods to restore a noisy MR image have been reported. For our experiments we will select those related somehow with the method proposed in this paper. First, those stochastic schemes based on the Rician model:

1) the CA by McGibney *et al.* [5];
2) the ML estimator [6]–[9];
3) the EM method [10], [11];
4) the *analytically exact solution*, proposed by Koay and Basser in [19].

In all the cases where the variance of noise is needed, it is manually set to its optimal value. Note that the ML and EM estimators, as well as the method by Koay and Basser, are designed to work with several samples of the same image. As in the present experiment we suppose only one sample is available, the statistics are computed using local neighborhoods. In all cases, 5×5 neighborhoods have been used. To compare the filters with other techniques, as follows.

5) *Adaptive Wiener Filtering* [32], using a 5×5 neighborhood. In order to achieve the best performance of the filter, σ_n is manually set to the actual value.

6) The *wavelet domain noise filter* for medical imaging proposed by Pižurica *et al.* in [14]. The best results for this experiment are achieved using $K = 5$ and 5×5 window size.

And finally, the LMMSE based schemes, as follows.

7) The *LMMSE Estimator*, as proposed in (13), with σ_n manually set to the actual value. 5×5 neighborhood.

8) The *LMMSE Estimator*, as proposed in (13), with σ_n automatically estimated using (18). 5×5 neighborhood both for filtering and noise estimation.

9) The *LMMSE Estimator*, as proposed in (13), with overestimation of noise, manually setting the standard deviation of noise to $\sigma_n + 20$.

10) Although once the image is filtered with the LMMSE estimator the output model is no longer Rician, or at least nothing assures it is, we can think in making the filter recursive. As the noise is dynamically estimated in each iteration, the filter should reach a steady state as the estimated noise gets smaller and smaller. As a result, if a proper noise estimation is done, the filter should stop modifying the input image once the noise is eliminated. We define then a recursive LMMSE (RLMMSE). For the experiments, eight and 50 iterations are considered, and a 5×5 neighborhood has been used both for filtering and noise estimation.

To compare the restoration performance of the different methods, two quality indexes are used: the Structural Similarity (SSIM) index [33] and the Quality Index based on Local Variance (QILV) [34]. Both give a measure of the structural similarity between the ground truth and the estimated 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. This way we are able to assess the noise cleaning and border preserving capability of the different schemes. Both indexes are bounded; the closer to one, the better the image. In addition, the mean square error (MSE) is also calculated. To avoid any bias due to background, these three quality measures are only applied to those areas of the image inside the skull.

Table I shows the experimental results of the average of 100 experiments for two different values of σ_n. The best value of each column has been highlighted. Some graphical results for $\sigma_n = 10$ are shown in Fig. 15.

When compared with other schemes considering a Rician noise model with proper noise estimation, the LMMSE and the RLMMSE show a better performance in terms of noise cleaning (a larger SSIM) while the edges are preserved (the QILV value gets better). However, when noise estimation is incorrectly done [as in Fig. 15(h) for the case of noise overestimation], the noise in the background is even amplified while the inner edges are blurred.

The noise cleaning performance of the ML, EM and Koay schemes are good, but, as the QILV index points out, they cause image blurring. Consequently, image information is lost at the border and the image edges. This performance is not due to the schemes themselves, but to the fact that they are originally designed to estimate the signal from multiple samples. When only one image is available, the statistics must be calculated locally. Consequently this local estimation produces some edge smoothing, in some cases similar to the one produced by a Gaussian filter. LMMSE, although it is also based on local statistics does not show this edge-blurring behavior. This is one of the arguments for using LMMSE when only one sample is available.

It is interesting to study the performance of the Wiener filtering; although it slightly blurs the image, it shows an overall
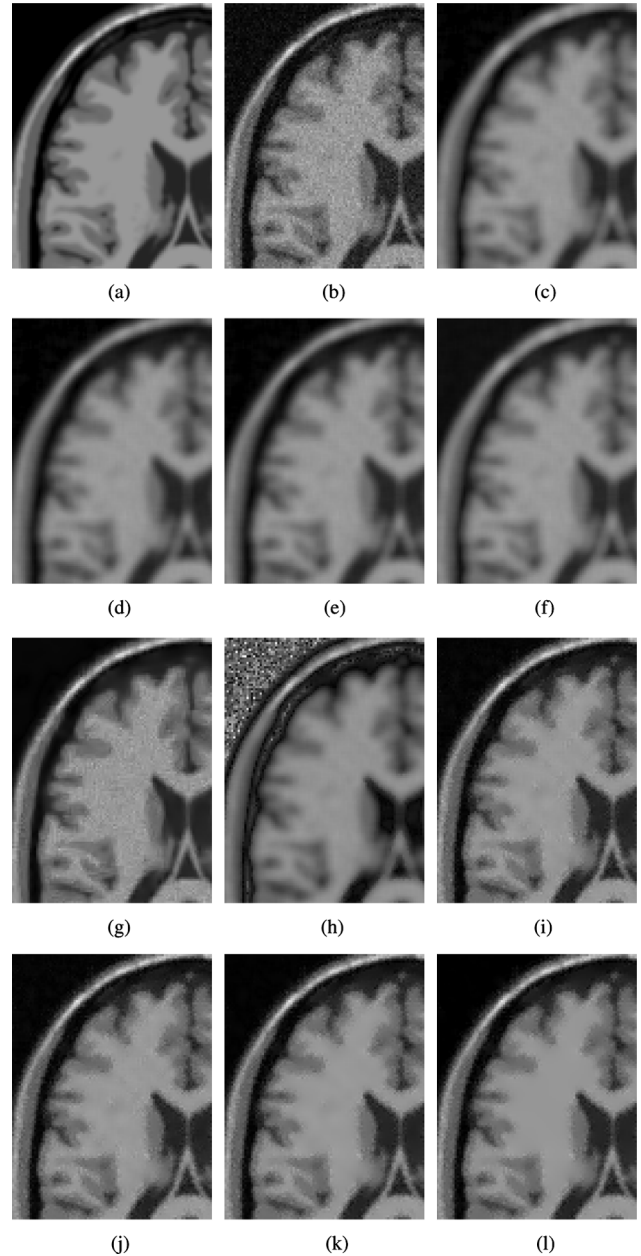

Fig. 15. Experiment with synthetic noise. MRI from Brainweb (detail). (a) Original Image. (b) Image with Rician noise with $\sigma_n = 10$. (c) Conventional approach. (d) ML estimator. (e) EM method. (f) Koay's method. (g) Wavelet Domain Noise filter. (h) LMMSE estimator with noise overestimation ($\sigma_n = 30$). (i) LMMSE estimator with $\sigma_n = 10$ (manually set). (j) LMMSE estimator automatic noise estimation. (k) Recursive LMMSE (eight iterations). (l) Recursive LMMSE (50 iterations). The LMMSE filters with correct noise estimation show the best performance, as confirmed by the numerical results in Table I.

good performance. Its performance is worse than the LMMSE, because the Wiener filter is based on a Gaussian noise model. This mismatch between the Rician model and the Gaussian model is not too large in structural MRI, but becomes more important in other kind of images, like diffusion weighted images (DWI).

On the other hand, the Wavelet based filter shows also a good behavior, both in edge preservation and noise cleaning, though the quantitative quality indices are a bit lower than those of the

TABLE I
QUALITY MEASURES FOR THE SYNTHETIC EXPERIMENT: SSIM, QILV, AND MSE FOR THE MR IMAGE WITH RICIAN NOISE. TWO DIFFERENT σ_n VALUES HAVE BEEN CONSIDERED. IN BOLD FACE, THE BEST VALUE OF EACH COLUMN. LMMSE-BASED SCHEMES SHOW BETTER RESULTS IN TERMS OF NOISE REMOVAL AND EDGE PRESERVATION

|	$\sigma_n = 10$			$\sigma_n = 20$		
	SSIM	QILV	MSE	SSIM	QILV	MSE
Noisy	0.7904	0.9890	100.2940	0.5722	0.9251	395.0881
CA	0.8491	0.6430	192.1922	0.8236	0.6543	205.6214
EM	0.8685	0.6513	144.1177	0.8373	0.6342	168.0615
Koay	0.8424	0.6026	212.4360	0.8024	0.5259	290.9409
ML	0.8681	0.6516	144.2500	0.8370	0.6354	168.1712
Wiener	0.9092	0.9839	57.9197	0.8146	0.9076	161.8120
Wavelet	0.8723	0.9889	51.8368	0.7695	0.9441	132.4157
LMMSE (Overestimation)	0.8107	0.8206	229.0185	0.7465	0.7077	376.9618
LMMSE (Manual noise)	0.9168	**0.9921**	53.9731	0.8346	**0.9613**	130.5361
LMMSE (Automatic noise)	0.9177	**0.9921**	53.6904	0.8389	0.9606	128.1376
RLMMSE (8 steps)	0.9270	0.9917	**51.8197**	**0.8597**	0.9502	**122.5699**
RLMMSE (50 steps)	**0.9298**	0.9915	51.8487	0.8540	0.9429	129.5132
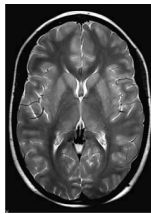
Fig. 16. Original MRI image from [14].
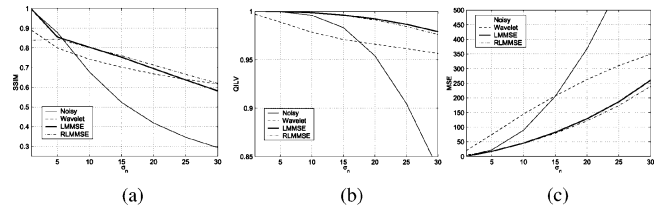
Fig. 17. Comparition between LMMSE and Wavelet-based schemes. Different values of σ_n have been considered. (a) SSIM, (b) QILV, and (c) MSE.

proposed schemes. In addition, the filter has a parameter K that has to be manually tuned to achieve the best results.

Finally, the RLMMSE results show a very good performance, when compared with the other schemes. There is a good balance between noise cleaning (SSIM index) and edge and structural information preservation (see QILV values). In addition, the filter shows great numerical stability: after 50 iterations the results are similar to those after eight iterations, indicating that the filter reaches a steady state.

One main advantage of the LMMSE filter (and to some extent for the RLMMSE filter) is that the solution can be computed in one single step (or a few steps for the RLMMSE filter), making it computationally efficient for large data sets as frequently encountered in DWI. This is in contrast to the EM and ML schemes, as well as to the approach by Koay and Basser, where the solution is found by numerical optimization and, thus, iteratively.

A further comparison has been done with the Wavelet based filter. The brain image used in [14] has been used as ground truth, see Fig. 16. The image is artificially corrupted with Rician noise with different values of σ_n; from 1 to 30. The noisy image is filtered using the wavelet filter (with $K = 2$), the LMMSE scheme and a RLMMSE with 5 iterations and noise estimation using (18). In every case, a 5×5 window has been used. The average of the quality measures of 100 experiments for each σ_n value is depicted in Fig. 17. Results show a better performance of both LMMSE schemes.

A next synthetic experiment is done using a 3-D ball, where Rician noise has been included (see Fig. 18). The image is filtered using a 3-D version of the filter and a $5 \times 5 \times 3$ window. As no zero background is available, the noise estimation is to be done following the method in (31). After ten iterations, the result is the image in Fig. 18 (right).
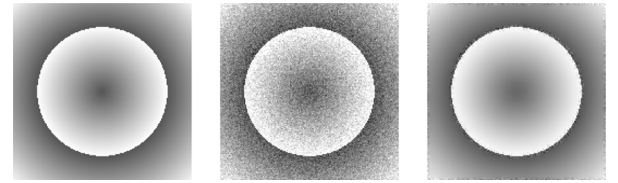

Fig. 18. One slice of a 3-D volume. Original image (left), image with Rician noise (center) and filtered after ten iterations (right).
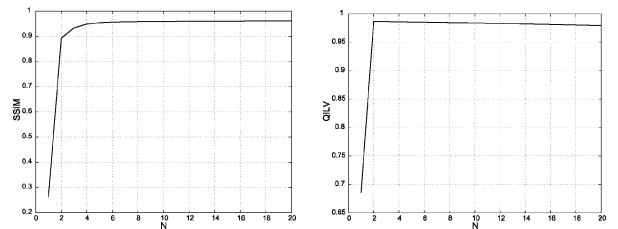

Fig. 19. SSIM index (left) and QILV index (right) of the 3-D noisy ball. There is a noise cleaning as well as the structural information is maintained.

SSIM and QILV have been used again as quality indices. Nineteen iterations of the 3-D RLMMSE filter have been done, with adaptive noise estimation; results are in Fig. 19. After four or five iterations, the filter reaches the optimum value. If we keep on filtering, the effect will be just a slight blurring of the image.

B. Estimation Using the Background

The methods for noise estimation in MRI described so far have the advantage that they do not need a mask to be segmented in order to estimate the noise. So, they can be easily used in any automatic procedure. This is one of the advantages of these methods over the traditional ones described before.
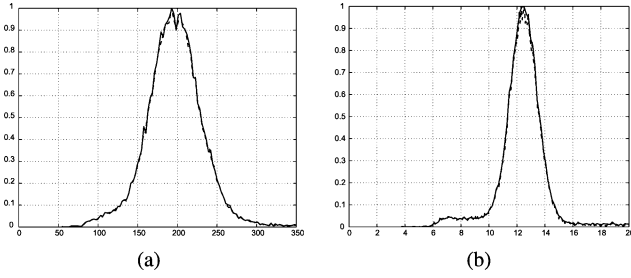
Fig. 20. (a) Distribution of the local second-order moment (using 7×7 windows) of the image without noise (solid) and with rice noise ($\sigma_n = 10$): Background pixels isolated (dash-dotted) and whole image (dashed). (b) Distribution of the local mean of the noisy image with $\sigma_n = 10$ using 7×7 windows: Background pixels isolated (dash-dotted) and whole image (dashed).
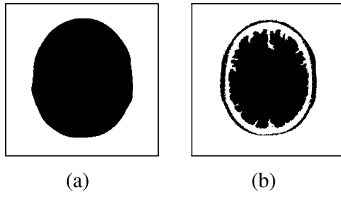

Fig. 21. Background masks used in the experiments.

If the background of the image may be isolated, the estimation can be done only over these background pixels. However, this is not a great advantage. In Fig. 20(a), the distribution of the local second-order moment is depicted, both for the whole noisy image and for the background of the image. Both distributions have almost the same shape, and, what is more important, their maxima are located at the same point. Very similar is the behavior of the local mean as shown in Fig. 20(b).

In order to understand the influence of a proper background segmentation over the noise estimation, a new experiment has been done. Local moments are calculated using 7×7 windows. For the noisy image, $\sigma_n = 10$, and 50 simulations have been done. First, we have used the mask in Fig. 21(a) to segment the background pixels. This mask totally separates the background from the signal, and it has been manually built form the original image. The second mask is in Fig. 21(b). It has been built from noisy data using automatic thresholding. Some of the tissue has been assigned to background. This is a common error in automatic segmentation. We will compare four different estimators:

1) the classical unbiased estimator for σ_n^2 of (6);
2) the classical unbiased estimator for σ_n of (8);
3) the new σ_n^2 estimator of (17), based on second-order moment distribution;
4) the new σ_n estimator of (18), based on local mean distribution.

Results are in Fig. 22(a) for the first mask and in Fig. 22(b) for the second one.

When the background pixels are perfectly identified, the performance of the estimators is quite similar. All the estimations are around $\widehat{\sigma_n} = 10$. So, in this case, the new methods do not present any advantage over the classical ones. However, in the second experiment, when some of the pixels of the brain are wrongly assigned to the background, the classical methods do not estimate properly, but the new ones do.
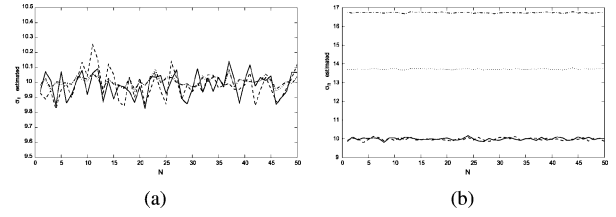

Fig. 22. Estimation of σ_n. (a) First background mask has been used: The new σ_n^2 estimator of (17) based on the local second-order moment (solid), new σ_n estimator of (17) based on the local mean (dashed), classical unbiased estimator of (6) for σ_n^2 (dash-dotted) and classical unbiased estimator of (8) for σ_n (dotted). (a) Using first background mask. (b) Using second background mask.
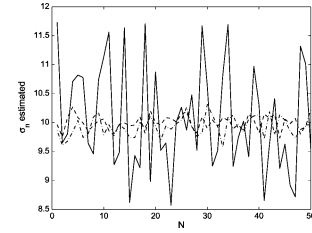

Fig. 23. Estimation of σ_n using the second background mask. Method based on the mode of the histogram (solid), new σ_n^2 estimator of (17) based on the local second-order moment (dashed), and new σ_n estimator of (17) based on the local mean (dash-dotted).

The noise estimation method based on the mode of the histogram proposed in [24] is also robust when a wrong segmentation of the background is done. However, as it is shown in Fig. 23 the estimator has a larger variance than the methods proposed in this paper.

C. Filtering MRI

To further verify the hypotheses proposed in the previous sections, and to have a visual idea of the behavior of the filter and noise estimators, a new experiment is carried out based on real data. A coronal slice from a 3-D MRI volume[2] has been selected.

This original image, see Fig. 24(a), exhibits noise. The image is filtered using the LMMSE filter and for the noise estimation the mode of the sample mean is used (18). Result is shown in Fig. 24(f). Some other noise-removal techniques have been used to compare with: EM estimation, the method proposed by Koay and Basser, the wavelet domain filter and the Wiener filter, as the latter showed a good performance in the former experiments. In all the cases adaptive noise estimation is performed using (18) with a 5×5 window. Results are given in Fig. 24. We have also considered once more the possibility of making the filter recursive. Results in Fig. 24(g) (eight iterations) shows the good behavior of this technique. Even after 100 iterations, Fig. 24(h), due to a good noise estimation, the filter is not blurring the edges. Again, as expected, the visual results are much better for the LMMSE-based schemes, as well as the wavelet based filter,

[2]Scanned in a 1.5-Tesla GE Echospeed system. Scanning sequence: Maximum gradient amplitudes: 40 mT/M. Six images with four high (750 s/mm^2), and two with low (5 s/mm^2) diffusion weighting. Rectangular FOV (field of view) 220×165 mm. 128×96 scan matrix (256×192 image matrix). 4-mm slice thickness, 1-mm interslice distance. Receiver bandwidth 6 kHz. TE (echo time) 70 ms; TR (repetition time) 80 ms (effective TR 2500 ms). Scan time 60 s/slice.

TABLE II
DIFFERENT ESTIMATORS OF σ_n^2 AND σ_n FOR RICIAN NOISE PROPOSED IN THE PAPER

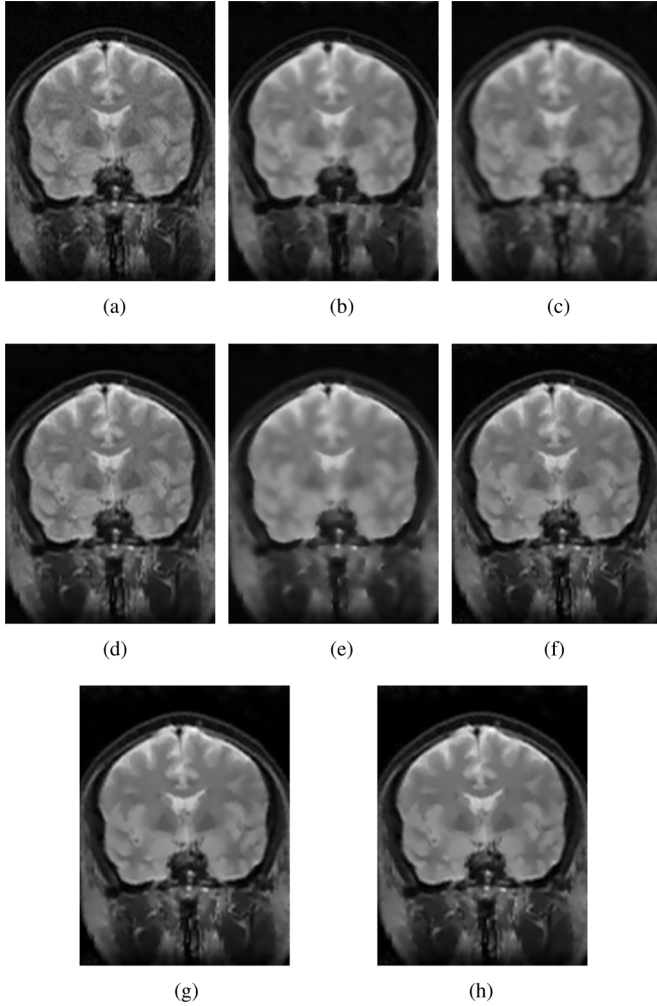| Use	Estimator	Local Statistics	Black background	Eq.
MRI	$\widehat{\sigma_n^2} = \frac{1}{2}\text{mode}(\widehat{\mu_{2ij}})$	$\widehat{\mu_{2ij}} = \frac{1}{\|\eta_{ij}\|-1}\sum_{p\in\eta_{ij}} I_p^2$	Needed	(17)
MRI	$\widehat{\sigma_n} = \sqrt{\frac{2}{\pi}}\text{mode}(\widehat{\mu_{1ij}})$	$\widehat{\mu_{1ij}} = \frac{1}{\|\eta_{ij}\|}\sum_{p\in\eta_{ij}} I_p$	Needed	(18)
MRI	$\widehat{\sigma_n^2} = \frac{2}{4-\pi}\text{mode}(\widehat{\sigma_{Iij}^2})$	$\widehat{\sigma_{Iij}^2} = \frac{1}{\|\eta_{ij}\|-1}\sum_{p\in\eta_{ij}}\left(I_p - \widehat{\mu_{1ij}}\right)^2$	Needed	(24)
Any image	$\widehat{\sigma_n^2} = \text{mode}(\widehat{\sigma_{Iij}^2})$	$\widehat{\sigma_{Iij}^2} = \frac{1}{\|\eta_{ij}\|-1}\sum_{p\in\eta_{ij}}\left(I_p - \widehat{\mu_{1ij}}\right)^2$	Not needed	(31)

Fig. 24. Coronal slice from a 3-D acquisition. (a) Original image. (b) EM method. (c) Koay's method. (d) Wavelet domain filter. (e) 2-D Wiener filter. (f) LMMSE estimator, adaptive noise estimation using the local mean and a 5×5 window. (g) RLMMSE estimator, adaptive noise estimation using the local mean and a 5×5 window (8 iterations). (h) RLMMSE estimator, adaptive noise estimation using the local mean and a 5×5 window (100 iterations).

though the latter still presents a slightly noisy pattern inside the tissues. The behavior shown for the other schemes is consistent with the synthetic experiments: noise is attenuated at the cost of blurring the image.

VI. CONCLUSION

A new filtering method based on the LMMSE estimator for Rician distributed images has been introduced, together with several noise estimation methods for that model. The filtering method has proved to be suitable for restoration in this kind of images, as it keeps the structure of the original image unaltered while suppresses most of the noise. Its performance is directly related with the goodness of the noise estimation method employed. If a good dynamic estimator is chosen, the filter may also be used recursively, showing a very good performance in noise cleaning. In addition, unlike other existing schemes also based on Rician models, the fact that a closed-form expression for the LMMSE method has been derived makes the filtering process computationally far more efficient and easier to implement.

We have also presented four different noise estimation methods, all of them based on the mode of some local statistic. A survey of these methods is in Table II. The use of the mode of the sample distribution of some local statistics (second-order moment, mean and variance) makes the estimator less dependent of parameters such as the size of the estimation window, the uniformity of the background, and outliers.

The experiments done with synthetic and real images show that the combination of the LMMSE filter with the noise estimation techniques here proposed may be a very useful tool for future MRI restoration.

APPENDIX A
ABOUT (20)

Parameter b is defined $b = (\sigma^2/N)[(2N-1)!!]^{1/N}$, where $n!! = n(n-2)!!$, for $n \geq 2$. We can write [31]

$$[(2N-1)!!]^{1/N} = \left[\frac{2^N\Gamma\left(N+\frac{1}{2}\right)}{\sqrt{\pi}}\right]^{1/N}$$
$$= \left[\frac{2^{1-N}\Gamma(2N)}{\Gamma(N)}\right]^{1/N}$$

and using the Stirling's approximation [31]

$$\Gamma(z) \approx e^{-z}z^{z-1/2}(2\pi)^{1/2}\left[1 + \frac{1}{12z} + \frac{1}{288z^2} + \cdots\right]$$
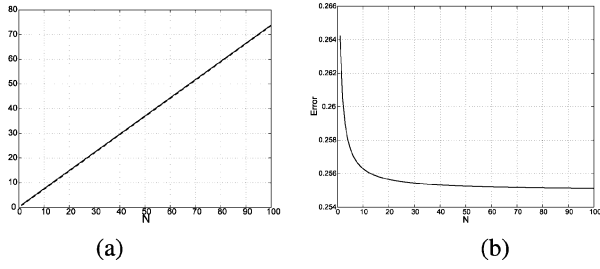
Fig. 25. (a) Function $[(2N-1)!!]^{1/N}$ (solid) and its approximation with function $2N/e$ (dashed). (b) $\mathrm{Error} = [(2N-1)!!]^{1/N} - 2N/e$.

Fig. 27. Normalized distribution of local variances of the images in Fig. 26, displayed in the same order. A 5×5 window has been used.
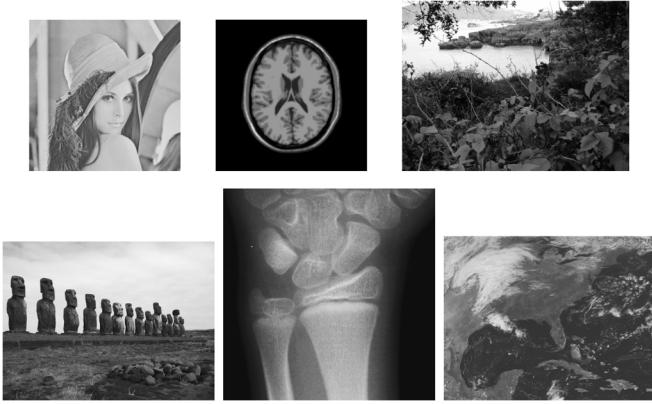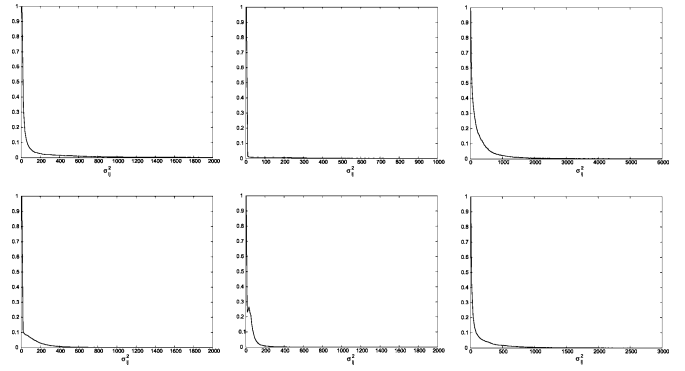

Fig. 26. Images to be analyzed. MRI from Brainweb [28]. Radiography from Hospital Río Carrión, Palencia, Spain. Aerial picture from GOES Project Science (http://goes.gsfc.nasa.gov).
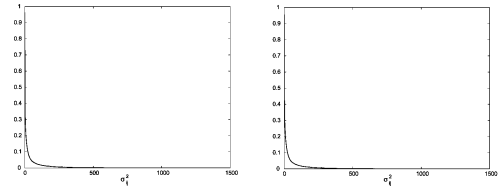

Fig. 28. Mean of the normalized distribution of local variances of the images from LIVE database [35]. A 5×5 window has been used. Grayscale images (left) and RGB images (right). Each color component have been evaluated separately.

for $z \to \infty$ and $|\arg z| < \pi$, we obtain

$$[(2N-1)!!]^{1/N}$$
$$\approx \left[2^{1-N} \frac{e^{-2N}(2N)^{2N-1/2}(2\pi)^{1/2}\left[1 + \frac{1}{24N} + \cdots\right]}{e^{-N}(N)^{N-1/2}(2\pi)^{1/2}\left[1 + \frac{1}{12N} + \cdots\right]} \right]^{1/N}$$

and as $N \gg$ we can approximate $[(2N-1)!!]^{1/N} \approx (2N/e)$ and $b \approx \sigma^2(2/e)$. In Fig. 25 the function $[(2N-1)!!]^{1/N}$ is depicted together with its approximation $2N/e$ for $N \in [1, 100]$.

APPENDIX B
LOCAL VARIANCE DISTRIBUTION IN NONTEXTURED IMAGES

One of the noise estimators proposed is based on the assumption that the distribution of the local variances in a non-noisy image follows a distribution whose maximum is set on zero. To evaluate this assumption, the histogram of local variances of different kind of images has been evaluated. The images are shown in Fig. 26. Medical, meteorological and natural images have been used. The sampling local variance is calculated using 5×5 windows. The results are shown in Fig. 27. Finally, the mean of the distribution of 29 images from LIVE database [35] is presented in Fig. 28. Two cases have been considered: first, grayscale images and, second, RGB images where each color component have been evaluated separately.

APPENDIX C
PDF OF THE SAMPLE LOCAL VARIANCE FOR RAYLEIGH RANDOM VARIABLES

Let $R_i(\sigma^2)$, $i = \{1, \ldots, N\}$ be a set of random variables with Rayleigh distribution. The (biased) sample variance is defined as

$$\widehat{\mathrm{Var}}(R_i) = \frac{1}{N}\sum_{i=1}^{N}(R_i)^2 - \left(\frac{1}{N}\sum_{i=1}^{N}R_i\right)^2.$$

We can define

$$S_1 = \frac{1}{N}\sum_{i=1}^{N}R_i^2(\sigma^2) \sim \chi_{2N}\left(\frac{\sigma^2}{N}\right)$$

being χ_{2N} a Chi-Square distribution, and

$$S_2 = \frac{1}{N}\sum_{i=1}^{N}R_i(\sigma^2).$$

The sample variance can be defined as

$$V = \widehat{\mathrm{Var}}(R_i) = S_1 - (S_2)^2.$$

The PDF of S_2 can be approximated [30] by

$$p_{S_2}(x) = \frac{x^{2N-1}N^N e^{-x^2 N/2b}}{2^{N-1}b^N \Gamma(N)}.$$

According to the fundamental theorem [29] the PDF of $(S_2)^2$ is

$$p_{(S_2)^2}(x) = \frac{x^{N-1}N^N e^{-xN/2b}}{2^N b^N \Gamma(N)}$$

$$(S_2)^2 \sim \chi_{2N}\left(\frac{b}{N}\right).$$

We can approximate (see Appendix A) $b \approx \sigma^2(2/e)$, so

$$(S_2)^2 \sim \chi_{2N}\left(\frac{\sigma^2}{N}\frac{2}{e}\right).$$

The distribution of the sample variance is then the distribution of the difference of two Chi-Square random variables. According to [22] this distribution, for $x \geq 0$ and the same (even) number of degrees of freedom is

$$p_X(x) = \frac{1}{2\sigma_1^2}e^{-\frac{x}{2\sigma_1^2}}\frac{1}{(N-1)!}\left(\frac{\sigma_1^2}{\sigma_1^2+\sigma_2^2}\right)^N$$

$$\times \sum_{k=0}^{N-1}\frac{(2N-2-k)!}{k!(N-1-k)!}$$

$$\times \left(\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right)^{N-1-k}\left(\frac{x}{2\sigma_1^2}\right)^k.$$

Making $\sigma_1^2 = (\sigma^2/N)$, $\sigma_2^2 = (\sigma^2/N)(2/e)$, we can write

$$p_V(x) = C_V e^{-\frac{xN}{2\sigma^2}}\sum_{k=0}^{N-1}\frac{(2N-2-k)!}{k!(N-1-k)!}\left[\left(1+\frac{e}{2}\right)\frac{xN}{2\sigma^2}\right]^k$$

being

$$C_V = \frac{N}{2\sigma^2}\frac{1}{(N-1)!}\frac{(2/e)^{N-1}}{(1+2/e)^{2N-1}}.$$

APPENDIX D
DIFFERENCE OF NONCENTRAL CHI-SQUARE RANDOM VARIABLES

Let X_1 and X_2 be two random variables with noncentral Chi-square distributions, such as

$$X_1 \sim \chi_{K_1}^2(x\alpha_1; \lambda_1) \quad X_2 \sim \chi_{K_2}^2(x\alpha_2; \lambda_2)$$

with PDFs

$$p_i(x) = f_i(x\alpha_i; K_i, \lambda_i) = \frac{e^{-\frac{x\alpha_i+\lambda_i}{2}}}{2}$$

$$\times \left(\frac{x\alpha_1}{\lambda_i}\right)^{\frac{K_i-2}{4}}I_{\frac{K_i}{2}-2}\left(\sqrt{x\alpha_i\lambda_i}\right)u(x) \quad (32)$$

being $I_n(z)$ the modified Bessel function of the first kind, which can be rewritten using ascending series [31] as

$$I_n(z) = \sum_{k=0}^{\infty}\frac{\left(\frac{1}{2}z\right)^{n+2k}}{k!\Gamma(n+k+1)}.$$

Equation (32) can be accordingly rewritten as

$$p_i(x) = e^{-\frac{x\alpha_i+\lambda_i}{2}}\sum_{m=0}^{\infty}\left(\frac{1}{2}\right)^{\frac{K_i}{2}+2m}\frac{(x\alpha_1)^{\frac{K_i-2}{2}+2m}\lambda_1^m}{m!\Gamma\left(\frac{K_i}{2}+m\right)}u(x).$$

Let us define the variable V as $V = X_1 - X_2$. For $x \geq 0$, we can write

$$p_V(x) = \int_0^{\infty}p_1(x+y)p_2(y)\,dy$$

If using Newton's generalized binomial theorem, the final PDF can be written as

$$p_V(x) = e^{-\frac{x\alpha_1+\lambda_1+\lambda_2}{2}}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{p=0}^{\infty}\left(\frac{1}{2}\right)^{n+m+p+1}\lambda_1^m\lambda_2^n$$

$$\times \binom{\frac{K_1}{2}-1+m}{p}\frac{\alpha_1^{K_1/2-1+m}\alpha_2^{K_2/2-1+n}}{(\alpha_1+\alpha_2)^{\frac{K_1+K_2}{2}-1+m+n-p}}$$

$$\times \frac{\Gamma\left(\frac{K_1+K_2}{2}-1+m+n-p\right)}{m!n!\Gamma\left(\frac{K_1}{2}+m\right)\Gamma\left(\frac{K_2}{2}+n\right)}x^p. \quad (33)$$

APPENDIX E
SAMPLE VARIANCE IN RICIAN DISTRIBUTED DATA

Let R_i, $i = \{1,\ldots,N\}$ be a set of random variables with Rician distribution and

$$S_1 = \frac{1}{N}\sum_{i=1}^{N}R_i^2 \quad S_2 = \frac{1}{N}\sum_{i=1}^{N}R_i.$$

The sample variance can be defined as $V = S_1 - (S_2)^2$. The PDF of S_1 is [21]

$$p_{S_1}(x) = M_1 x^{\frac{N-1}{2}}e^{-xN/2\sigma^2}I_{N-1}\left(\frac{\sqrt{xN}A}{\sigma^2}\right)$$

with M_1 a constant, and $A^2 = \sum_i A_i^2$. The sum of Rician distribution is, as in the Rayleigh case, a classical problem in communications. We can use the approximation in [21]. From here, the PDF of $(S_2)^2$ would be

$$p_{S_2^2}(x) = M_2 x^{\frac{N-1}{2}}e^{-xN/2c_2^2}I_{N-1}\left(\frac{\sqrt{xN}b}{c_1c_2}\right)$$

with M_2 a constant, and c_1, c_2, and b parameters related with N and the Signal to Noise ratio. Both PDF may be expressed as noncentral Chi-Square distributions

$$S_1 \sim \chi_{2N}^2\left(\frac{xN}{\sigma^2}, \frac{A^2}{\sigma^2}\right) \quad S_2^2 \sim \chi_{2N}^2\left(\frac{xN}{c_2^2}, \frac{b^2}{c_1^2}\right).$$

The subtraction of two noncentral Chi-Square random variables has been studied in Appendix D. Accordingly, the PDF of the sample variance will be like the one in (33). However, this long equation may have some problems when trying to numerically determine its maximum, specially when the argument of the Modified Bessel function is not small.
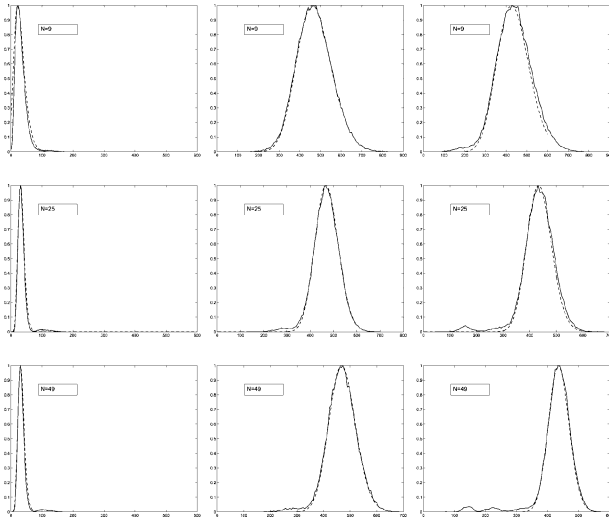
Fig. 29. Rician distributed data. Theoretical PDF versus real data distribution. Left: Local Sample variance. Center: Local second-order moment. Right: Square local mean. Solid line: data distribution. Dashed line: theoretical PDF.

Experimentally, a good approximation has been obtained, without any dependence on the parameters c_1, c_2, A, and b

$$p(t) \approx C_T e^{-t} \sum_{m=0}^{\frac{2N-1}{4}} \binom{\frac{2N-1}{4}}{m} \times t^m (1+N)^m \Gamma(2N - 1/2 - m) \quad (34)$$

with $t = (xN/(2\sigma_n^2))$.

To show the behavior of the approximation, some data has been generated following a Rician model. A constant image with value 20 has been corrupted with Rician noise with $\sigma_n = 6$. The distribution of some local statistics have been calculated using 3×3, 5×5 and 7×7 square windows. The results are shown in Fig. 29. In the solid line, the actual data distribution is depicted, and the theoretical distribution is in dashed line.

ACKNOWLEDGMENT

REFERENCES

[1] D. Drumheller, "General expressions for Rician density and distribution functions," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 2, pp. 580–588, Apr. 1993.
[2] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magn. Reson. Med.*, vol. 34, pp. 910–914, Aug. 1995.
[3] S. Basu, T. Fletcher, and R. Whitaker, "Rician noise removal in diffusion tensor MRI," in *Proc. MICCAI*, 2006, vol. 1, pp. 117–125.
[4] R. Henkelrnan, "Measurement of signal intensities in the presence of noise in MR images," *Med. Phys.*, vol. 12, no. 2, Mar. 1985.
[5] G. McGibney and M. Smith, "Un unbiased signal-to-noise ratio measure for magnetic resonance images," *Med. Phys.*, vol. 20, no. 4, pp. 1077–1078, Jul. 1993.
[6] J. Sijbers, A. J. den Dekker, P. Scheunders, and D. Van Dyck, "Maximum-likelihood estimation of Rician distribution parameters," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 357–361, Jun. 1998.
[7] J. Sijbers, A. J. den Dekker, D. Van Dyck, and E. Raman, "Estimation of signal and noise from Rician distributed dala," in *Proc. Int. Conf. Signal Processing Communications*, Las Palmas de Gran Canaria, Spain, Feb. 1998, pp. 140–142.
[8] J. Sijbers and A. J. den Dekker, "Maximum likelihood estimation of signal amplitude and noise variance form MR data," *Magn. Reson. Imag.*, vol. 51, pp. 586–594, Feb. 2004.
[9] L. Jiang and W. Yang, "Adaptive magnetic resonance image denoting using mixture model and wavelet shrinkage," presented at the 7th Digital Image Computing: Techniques and Applications, Sydney, Australia, Dec. 2003.
[10] M. D. DeVore, A. D. Lanterman, and J. A. O'Sullivan, "ATR performance of a Rician model for SAR images," in *Proc. SPIE, ISSV 4050*, Orlando, FL, Apr. 2000, pp. 34–37.
[11] T. Marzetta, "EM algorithm for estimating the parameters of multivariate complex Rician density for polarimetric SAR," in *Proc. ICASSP*, May 1995, vol. 5, pp. 3651–3654.
[12] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache, "Clinical DT-MRI estimation, smoothing and fiber tracking with log-euclidean metrics," in *Proc. IEEE Int. Symp. Biomedical Imaging*, Apr. 2006, pp. 786–789.
[13] R. Nowak, "Wavelet-based Rician noise removal for magnetic resonance imaging," *IEEE Trans. Image Process.*, vol. 8, no. 10, pp. 1408–1419, Oct. 1999.
[14] A. Pižurica, W. Philips, I. Lemahieu, and M. Acheroy, "A versatile Wavelet domain noise filtration technique for medical imaging," *IEEE Trans. Med. Imag.*, vol. 22, no. 3, pp. 323–331, Mar. 2003.
[15] M. Lysaker, A. Lundervold, and X. Tai, "Noise removal using fourth order partial differential equations with applications to medical magnetic resonance imaging in space and time," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1579–1590, Dec. 2003.
[16] T. McGraw, B. C. Vemuri, Y. Chen, M. Rao, and T. Mareci, "DT-MRI denoising and neuronal fiber tracking," *Med. Imag. Anal.*, vol. 8, pp. 95–111, Jun. 2004.
[17] C. B. Ann, Y. C. Song, and D. J. Park, "Adaptive template filtering for signal-to-noise ratio enhancement in magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 18, no. 6, pp. 549–556, Jun. 1999.
[18] S. Awate and R. W. RT, "Nonparametric neighborhood statistics for MRI denoising," in *Proc. Int. Conf. Inf. Process. Med. Imaging*, 2005, vol. 3565, pp. 677–688.
[19] C. G. Koay and P. J. Basser, "Analytically exact correction scheme for signal extraction from noisy magnitude MR signals," *J. Magn. Reson.*, vol. 179, pp. 317–322, Apr. 2006.
[20] M. Martin-Fernández, C. Alberola-López, J. Ruiz-Alzola, and C.-F. Westin, "Sequential anisotropic Wiener filtering applied to 3-D MRI data," *Magn. Reson. Imag.*, vol. 25, pp. 278–292, Feb. 2007.
[21] J. Hu and N. C. Beaulieu, "Accurate closed-form approximations to Ricean sum distributions and densities," *IEEE Commun. Lett.*, vol. 9, no. 2, pp. 133–135, Feb. 2005.
[22] M. K. Simon, *Probability Distributions Involving Gaussian Random Variables*. Norwell, MA: Kluwer, 2002.
[23] J. Sijbers, A. J. den Dekker, J. Van Audekerke, M. Verhoye, and D. Van Dyck, "Estimation of the noise in magnitude MR images," *Magn. Reson. Imag.*, vol. 16, no. 1, pp. 87–90, Jan. 1998.
[24] J. Sijbers, A. J. den Dekker, D. Poot, M. Verhoye, N. Van Camp, and A. Van der Linden, "Robust estimation of the noise variance from background MR data," presented at the SPIE Medical Imaging: Image Processing, Mar. 2006.
[25] J. Sijbers, D. Poot, A. J. den Dekker, and W. Pintjenst, "Automatic estimation of the noise variance from the histogram of a magnetic resonance image," *Phys. Med. Biol.*, vol. 52, pp. 1335–1348, Feb. 2007.
[26] R. Salvador, A. Peña, D. K. Menon, T. Carpenter, J. Pickard, and E. T. Bullmore, "Formal characterization and extension of the linearized diffusion tensor model," *Human Brain Map.*, vol. 24, pp. 144–155, Feb. 2005.
[27] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[28] D. Collins, A. Zijdenbos, V. Killokian, J. Sled, N. Kabani, C. Holmes, and A. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Trans. Med. Imag.*, vol. 17, no. 3, pp. 463–468, Mar. 1998.
[29] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. Boston, MA: McGraw-Hill, 1991.
[30] N. C. Beaulieu, "An infinite series for the computation of the complementary probability distribution function of a sum of independent ramdom variables and its application to the sum of Rayleigh random variables," *IEEE Trans. Commun.*, vol. 38, no. 9, pp. 1463–1474, Sep. 1990.

[31] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed.   New York: Dover, 1964.

[32] J. S. Lim, *Two Dimensional Signal and Image Processing*.   Englewood Cliffs, NJ: Prentice-Hall, 1990.

[33] Z. Wang, A. C. Bovik, H. 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, Apr. 2004.

[34] S. Aja-Fernández, R. San-José-Estépar, C. Alberola-López, and C. Westin, "Image quality assesment based on local variance," in *Proc. 28th IEEE EMBS*, New York, Sep. 2006, pp. 4815–4818.

[35] Live Database. Laboratory for Image and Video Engineering (LIVE) [Online]. Available: http://live.ece.utexas.edu/research/quality/subjective.htm The Univ. Texas, Austin

Carlos Alberola-López received the engineering of Telecommunication degree and the Ph.D. degree from the Politechnical University of Madrid, Madrid, Spain, in 1992 and 1996, respectively.

He is a Professor at the ETSI Telecomunicación, University, Valladolid, Spain. In 1997, he held a visiting scientist position at Thayer School of Engineering, Dartmouth College, NH. His research interests are statistical and fuzzy methods for signal and image processing applications. He is the Head of the Laboratory of Image Processing (LPI), University of Valladolid. He is a reviewer for several scientific journals and he is a consultant of the Spanish Government for the evaluation of research proposals.

Santiago Aja-Fernández received the Ingeniero de Telecomunicación and the Ph.D. degrees from the University of Valladolid, Spain, in 1999 and 2003, respectively.

He is an Assistant Professor with the E.T.S.I. Telecomunicación, the University of Valladolid, where he is also with the Laboratory of Image Processing (LPI). His research interests include medical image analysis and processing.

Dr. Aja-Fernández was awarded with a Fulbright Scholarship for a one-year stay as a Research Fellow with the LMI, Brigham and Women's Hospital, Boston, MA.

Carl-Fredrik Westin (M'99) received the M.Sc. degree in applied physics and electrical engineering in 1988, the Lic.Techn. degree on the topic of feature extraction from a tensor image description in 1991, and the Ph.D. degree in computer vision in 1994 from Linköping University, Linköping, Sweden.

He joined the Computer Vision Laboratory in the Department of Electrical Engineering, Linköping University, the same year where he did research on color, information representation, image flow, frequency estimation, filtering of uncertain and irregularly sampled data and tensor operators in image analysis. In 1996, he joined Brigham and Women's Hospital and Harvard Medical School, Boston, MA, where he became the Director of the Laboratory of Mathematics in Imaging (LMI) in the Department of Radiology in 2001, and a Research Affiliate of the Artificial Intelligence Laboratory at the Massachusetts Institute of Technology, Cambridge.

Dr. Westin was awarded the SAAB-SCANIA prize for his work in computer vision in 1991.