AN ADAPTIVE NON LOCAL MAXIMUM LIKELIHOOD ESTIMATION METHOD FOR DENOISING MAGNETIC RESONANCE IMAGES

Jeny Rajan † *Johan Van Audekerke* ⋆ *Annemie Van der Linden*⋆
Marleen Verhoye⋆ *Jan Sijbers*†
† Vision Lab, Dept. of Physics, University of Antwerp, Belgium
⋆ Bio-Imaging Lab, Dept. of Biomedical Sciences, University of Antwerp, Belgium

ABSTRACT

Effective denoising is vital for proper analysis and accurate quantitative measurements from Magnetic Resonance (MR) images. Apart from following the general criteria for denoising, the algorithms that deal with MR images should also take into account the bias generated due to the Rician nature of the noise in the magnitude MR images. Maximum Likelihood (ML) estimation methods were proved to be very effective in denoising MR images. However, one drawback of the existing non local ML estimation method is the usage of a fixed sample size for ML estimation. As a result, optimal results cannot be achieved because of over or under smoothing. In this work, we propose an adaptive non local ML estimation method for denoising MR images in which the samples are selected in an adaptive way for the ML estimation of the true underlying signal. The method has been tested both on simulated and real data, showing its effectiveness.

Index Terms— MRI, Noise, NLML, Rician distribution

1. INTRODUCTION

The presence of noise remains one of the main causes of quality deterioration in MRI. The primary sources of noise in MR images are: (i) thermal noise from the body, (ii) quantization noise in the A/D devices, (iii) electronic noise, and (iv) thermal noise in the RF coil. Other than causing image degradation, the noise in the MR image also increases the mean signal intensity of image regions due to the usually performed magnitude reconstruction, which causes the data to be Rician distributed. As a result, conventional denoising algorithms, which assume the noise to be Gaussian, can introduce a bias in the denoised image. Even though the bias generated due to the Gaussian assumption may not be too large in structural MRI, it becomes more important in other kind of MR images where the SNR is relatively low, like diffusion weighted images [1].

This work was financially supported by the Inter-University Attraction Poles Program 6-38 of the Belgian Science Policy, and by the SBO-project QUANTIVIAM (060819) of the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen)

Several adaptive filtering techniques to improve the quality of the magnitude MR images have been proposed in the literature. Most of the methods proposed earlier can be mainly classified as either based on Partial Differential Equations (PDEs), wavelets or Non Local Means (NLM). Recently, ML based methods were proposed for the estimation of the true underlying signal given a noisy MR image [2, 3, 4]. From a statistical point of view, ML estimation method is known to yield optimal results and can be successfully applied for estimating the true signal from an MR image by incorporating the noise hypothesis provided that the underlying area is constant. The recently proposed non local maximum likelihood (NLML) estimation method [3] is an effective approach for denoising magnitude MR images.

In the NLML method, the true underlying intensity for each pixel is estimated using the ML estimation method applied on a set of non local (NL) pixels selected based on the intensity similarity of the pixel neighborhood. The number of NL pixels to be considered for ML estimation is fixed and is generally determined in a heuristic way. This fixed sample size can introduce under or over smoothing in the images. In this paper, we propose an adaptive way of selecting the sample size for the ML estimation and thus to improve the performance of the NLML method.

2. THEORY

The noise in the acquired raw complex MR data in the k - space is characterized by a zero mean Gaussian probability density function (PDF). The k - space data is then Fourier transformed to obtain the magnetization distribution. The noise distribution in the real and imaginary components will still be Gaussian due to the linearity and the orthogonality of the Fourier transform. However, due to the subsequent transform to a magnitude image, the data will no longer be Gaussian but Rician distributed. For an MR magnitude image defined on a discrete grid Ω, $M = \{m_i | i \in \Omega\}$, the probability distribution function of m_i, with underlying signal A, is: [5, 6]

$$p(m_i|A, \sigma_g) = \frac{m_i}{\sigma_g^2} e^{-\frac{m_i^2 + A^2}{2\sigma_g^2}} I_0\left(\frac{A m_i}{\sigma_g^2}\right) \epsilon(m_i) \qquad (1)$$

where $I_0(.)$ is the 0^{th} order modified Bessel function of the first kind, $\epsilon(.)$ is the Heaviside step function and σ_g^2 denotes the variance of the Gaussian noise in the complex MR data.

Proper estimation of the noise variance, σ_g^2, is important for effective denoising of MRI. In this work, we make use of the prior knowledge of the noise variance to select the sample size for ML estimation. Noise can be estimated from an MR image in a number of ways. A survey of these methods is given in [1].

2.1. Signal estimation using NLML method

Let $m_1, m_2, ..., m_n$ be n statistically independent observations within a region of constant signal intensity A. Then the joint pdf of the observations is:

$$p(\{m_i\}|A, \sigma_g) = \prod_{i=1}^{n} \frac{m_i}{\sigma_g^2} e^{-\frac{m_i^2+A^2}{2\sigma_g^2}} I_0\left(\frac{Am_i}{\sigma_g^2}\right) \tag{2}$$

The ML estimate of A can now be computed by maximizing the likelihood function $L(A)$ or equivalently $\ln L(A)$, with respect to A [2]:

$$\ln L = \sum_{i=1}^{n} \ln\left(\frac{m_i}{\sigma_g^2}\right) - \sum_{i=1}^{n} \frac{m_i^2 + A^2}{2\sigma_g^2} + \sum_{i=1}^{n} \ln I_0\left(\frac{Am_i}{\sigma_g^2}\right) \tag{3}$$

and

$$\widehat{A}_{\mathrm{ML}} = \arg\{\max_{A}(\ln L)\} \tag{4}$$

In a region of constant signal amplitude, the true underlying signal can be estimated from the noisy signal using Eq(4). The straightforward approach to denoise MR images using the above explained ML estimation method is to apply the method locally for each pixel with the assumption that the underlying area is constant for a small region. It is clear, however, that this assumption is not valid for regions with edges or fine structures and, as a result, these regions will get blurred.

One solution suggested to the above mentioned problem is to use non local (NL) pixels instead of local ones [3]. The NL pixels are selected based on the intensity similarity of the pixel neighborhood. The idea is that if two regions are similar, then their central pixels should have a similar meaning for the image and thus similar gray values [7]. Thus, the similarity between two pixels i and j can be computed by measuring the intensity distance d between neighborhood region W surrounding i and j [3]:

$$d_{i,j} = \|W_i - W_j\|. \tag{5}$$

The less the intensity distance, the similar the pixels. In NLML method, for each pixel i, the intensity distance d between i and all other N non local neighboring pixels are measured. The first k pixels are then selected for ML estimation after sorting the NL pixels in the increasing order of d. This method is highly efficient when compared to local ML estimation. However, one major concern with this approach is the proper selection of k. In the current NLML method, k is fixed and generally determined in a heuristic way. Low or high value of k can cause under or over smoothing. In the section below, we discuss how to select k in an adaptive manner.

2.2. Adaptive NLML estimation

The performance of the NLML method can be significantly improved by adaptively selecting the number of pixels k for the ML estimation. The optimal value for k yields the best performance of NLML. To compute k, we take advantage of the known value of the noise variance.

Let m_i be the pixel to be denoised and $\zeta = \{n_1, n_2, ..., n_N\}$ be the N independent NL neighbors of m_i. Compute the distance $D = \{d_1, d_2, ...d_N\}$ for each elements of ζ with m_i by applying Eq(5) and sort ζ in the increasing order of the distance. Let S be the list of sorted ζ. Estimate the underlying intensity $\widehat{A}_{S(j)}$ and noise variance $\widehat{\sigma^2}_{S(j)}$ from the first j elements of S by maximizing the log likelihood function $\ln L$ with respect to both $A_{S(j)}$ and $\sigma^2_{S(j)}$ and by varying j from 1 to N:

$$\{\widehat{A}_{S(j)}, \widehat{\sigma^2}_{S(j)}\} = \arg\{\max_{A_{S(j)}, \sigma^2_{S(j)}} (\ln L)\} \tag{6}$$

Now the true underlying intensity $\widehat{A}_i$ at i can be computed as

$$\widehat{A}_i = \widehat{A}_{S(k)} \tag{7}$$

where

$$k = \arg\{\min_{j} \text{abs}(\widehat{\sigma_g^2} - \widehat{\sigma^2}_{S(j)})\} \tag{8}$$

where $\widehat{\sigma_g^2}$ is the estimated noise variance in the image and $\widehat{\sigma^2}_{S(j)}$ is the noise variance estimated from the first j elements of S. The algorithm is summarized in algorithm 1.

One drawback of the proposed method is its increased time complexity. However, the computation time can be reduced either by using the method (correction factor based on SNR) mentioned in [9] or the skewness based method in [8] for noise estimation as a replacement to ML estimation given in Eq(6).

3. EXPERIMENTS AND RESULTS

To evaluate and compare the proposed adaptive NLML with the conventional NLML method, experiments were conducted on both synthetic and real MR images. For the experiments on the synthetic data, we used the standard MR image phantom of the brain obtained from the Brainweb database [10]. The phantom image was degraded with Rician noise for a wide range of noise levels and the denoising efficiency of both algorithms were evaluated based on the Peak Signal to Noise Ratio (PSNR), the mean Structural Similarity Index

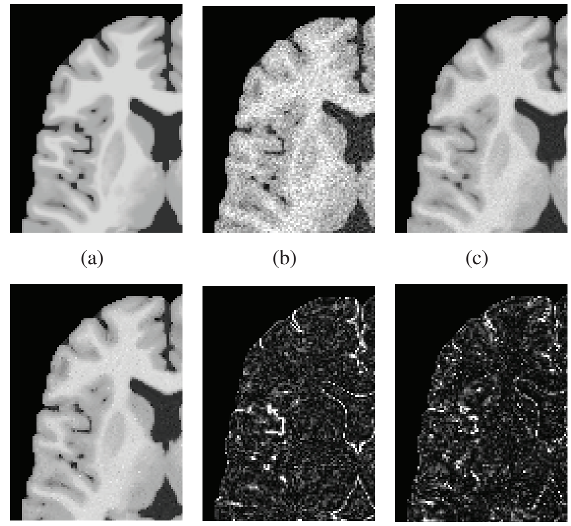
Algorithm 1 Algorithm for signal estimation using adaptive NLML

Estimate the noise variance $\widehat{\sigma_g^2}$ from the input magnitude image M { refer [8] for noise estimation} .
for every pixel $m(i)$ of M **do**
  Compute the similarity measure for N pixels in an extended neighborhood of $m(i)$) using a window size of $3 \times 3 \times 3$ and create a list D.
  S = Sort(D) {sort D in the increasing order of similarity measure }
  for $j = 1 \rightarrow N$ **do**
    $[\widehat{A}_{S(j)}, \widehat{\sigma^2}_{S(j)}]$ = MLestimate(*S(j)*) {ML estimation of $\widehat{A}_{S(j)}, \widehat{\sigma^2}_{S(j)}$ using the first *j* elements of *S* }
  end for
  $k = \arg\{\min_j \text{ abs}(\widehat{\sigma_g^2} - \widehat{\sigma^2}_{S(j)})\}$
  $\widehat{A}_i = \widehat{A}_{S(k)}$
end for

Matrix (SSIM)[11], Bhattacharrya coefficient (BC)[12] and the Mean Absolute Difference (MAD).

Fig. 1 shows the visual quality comparison of the image denoised with the NLML and the proposed improved version. This experiment was conducted on the brain image after corrupting the image by noise with $\sigma_g = 20$. It can be seen from the denoised images and from the corresponding residuals that the image denoised with the proposed adaptive NLML method is closer to the original one than the image denoised with NLML. Fig. 2 shows the quantitative analysis of both methods in terms of PSNR, mean SSIM, BC and MAD. Both filters were executed with the following parameters: search window size = $11 \times 11 \times 11$, neighborhood size = $3 \times 3 \times 3$. For high noise levels, a bigger neighborhood size is preferred. For NLML, the sample size k was chosen as 25 (as recommended in [3]). It can be observed from the plots that quality improvement can be achieved by incorporating the suggested improvement in the NLML method.

Next, for the experiments on the real data, we used the MR images of a Kiwi fruit. Two sets of Kiwi fruit images were reconstructed, one without averaging and the other by averaging 12 acquisitions respectively. Averaging was done in the complex k-space. These images were acquired on a 9.4T MR scanner with a slice thickness of 0.4 mm. Both the NLML and the adaptive NLML method were then applied over the image reconstructed without averaging and the resultant denoised image was compared with the image reconstructed by averaging 12 acquisitions. It can be observed from the images in Fig. 3 that the visual results are much better for the proposed adaptive NLML in terms of image contrast. Also the residual images shows that the adaptive NLML methods preserves the image structure much better than the NLML method. This experiment on real data clearly highlights the improved performance of the adaptive NLML method.



(a) (b) (c) (d) (e) (f)

Fig. 1. Visual quality comparison: (a) Ground truth (b) Ground truth corrupted with Rician noise of σ_g = 20 (c) denoised with NLML method (d) denoised with adaptive NLML method (e) NLML residuals (f) adaptive NLML residuals. The residuals are scaled in the range 0-20.

4. CONCLUSION

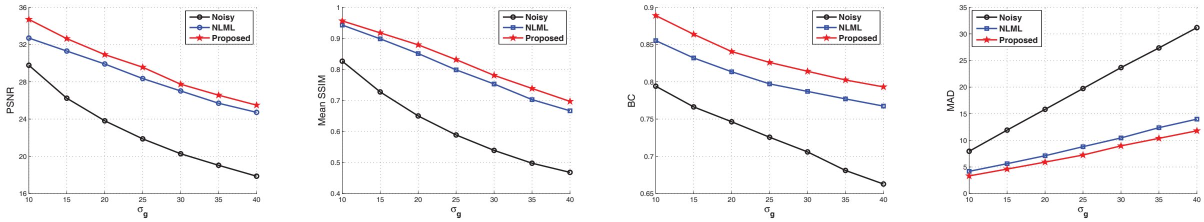
A method to improve the performance of the NLML method is proposed in this paper. The improvement is achieved by adaptively selecting the number of samples to be considered for the ML estimation. Through this approach, the over and under smoothing caused by the NLML can be reduced. Experiments have been carried out on simulated and real data sets. Quantitative analysis at various noise levels based on the similarity measures, PSNR, SSIM, BC and MAD shows that the proposed method is more effective than conventional NLML. Experiments were also performed on real MR images to prove the efficacy of the proposed method.

5. REFERENCES

[1] S. Aja-Fernández, A. Tristán, and C. Alberola-López, “Noise estimation in single and multiple coil magnetic resonance data based on statistical models,” *Magn Reson Imaging*, vol. 27, pp. 1397–1409, 2009.

[2] J. Sijbers and A. J. den Dekker, “Maximum likelihood estimation of signal amplitude and noise variance from MR data,” *Magn Reson Med*, vol. 51, no. 3, pp. 586–594, 2004.

[3] L. He and I. R. Greenshields, “A nonlocal maximum likelihood estimation method for rician noise reduction in MR images,” *IEEE Trans Med Imaging*, vol. 28, pp. 165–172, 2009.



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

Fig. 2. Quantitative analysis of the proposed method with NLML based on (a) PSNR (b) SSIM (c) BC and (d) MAD for image corrupted with Rician noise of σ_g varying from 10 to 40.

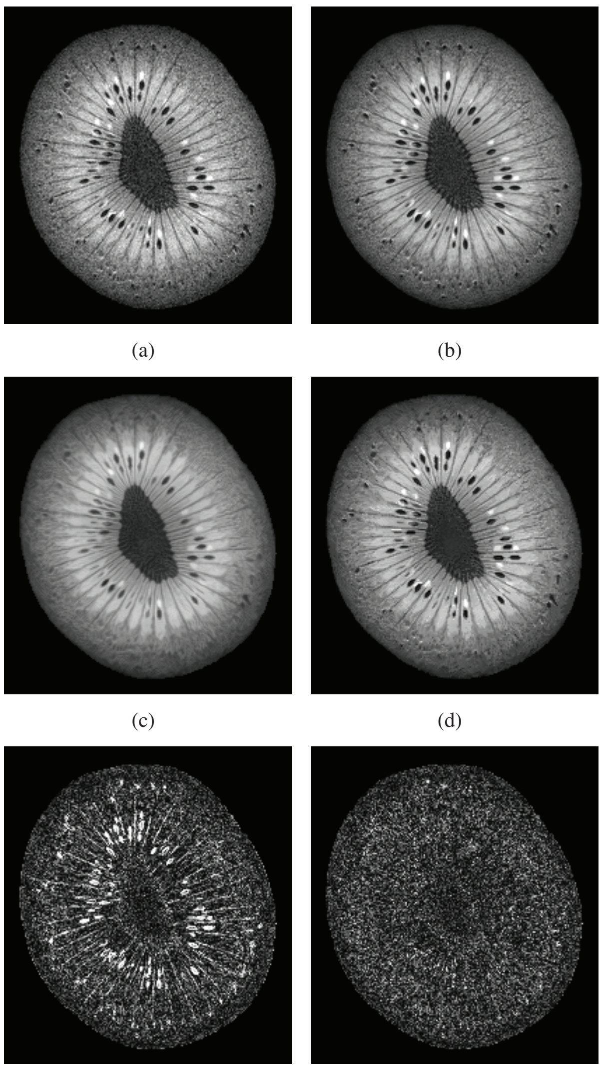


Fig. 3. Experiments on the MR image of a Kiwi fruit (a) and (b) Original images reconstructed with 1 and 12 averages respectively (c) denoised with NLML method (sample size k = 25) (d) denoised with the proposed adaptive NLML method (e) and (f) shows the corresponding residual images.

[4] J. Rajan, B.Jeurissen, M.Verhoye, J. V. Audekerke, and J. Sijbers, "Maximum likelihood estimation-based denoising of magnetic resonance images using restricted local neighborhoods," *Phys Med Biol*, vol. 56, pp. 5221–5234, 2011.

[5] S. O. Rice, "Mathematical analysis of random noise," *Bell Syst Tech*, vol. 23, pp. 282–332, 1944.

[6] M. A. Bernstein, D. M. Thomasson, and W. H. Perman, "Improved detectability in low signal-to-noise ratio magnetic resonance images by means of phase-corrected real construction," *Med Phys*, vol. 16, no. 5, pp. 813–817, 1989.

[7] S. Zimmer, S. Didas, and J. Weickert, "A rotationally invariant block matching strategy improving image denoising with non-local means," in *International Workshop on Local and Non-Local Approximation in Image Processing*, Switzerland, 2008, pp. 135–142.

[8] J. Rajan, D. Poot, J. Juntu, and J. Sijbers, "Noise measurement from magnitude MRI using local estimates of variance and skewness," *Phys Med Biol*, vol. 55, pp. N441–N449, 2010.

[9] 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, 2006.

[10] C. A. Cocosco, V. Kollokian, R.K.-S. Kwan, and A. C. Evans, "Brainweb: Online interface to a 3D MRI simulated brain database; http://www.bic.mni.mcgill.ca/brainweb/ ," *NeuroImage*, vol. 5, no. 4, pp. S425, 1997.

[11] Z. Wang, A.C. Bovik, H. R. Sheik, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans Imag Proc*, vol. 13, pp. 600–612, 2004.

[12] A. Bhattacharyaa, "On the measure of divergence between two statistical populations defined by their probability distributions," *Bull Calcutta Math Soc*, vol. 35, pp. 99–109, 1943.