The Maximum Spacing Noise Estimation in Single-coil Background MRI Data

Tomasz Pieciak

AGH University of Science and Technology, Krakow, Poland
THE MAXIMUM SPACING NOISE ESTIMATION IN SINGLE-COIL BACKGROUND MRI DATA

Tomasz Pieciak

AGH University of Science and Technology, Krakow, Poland

ABSTRACT

This paper presents new noise level estimation technique in single-coil background MRI data based on the maximum spacing estimation (MSP) principle. We derive new MSP estimator for Rayleigh distribution using the Kullback-Leibler divergence approximation, which in comparison with maximum likelihood approach, is based on spacings between successive order statistics. Moreover, we generalize the MSP estimator into higher order spacings and derive subsequent MSP estimators inferring from different statistical distances, i.e., J-divergence, Rényi divergence and Vajda’s measure.

We validate our approach both on synthetic brain and real cardiac MRI data in comparison with literature reports. The experimental results show that the MSP estimator has a comparatively low bias and reach Cramér-Rao lower bound on the variance. Finally, we assess properties of MSP estimators derived from different statistical distances with each other.

Index Terms— Noise estimation, maximum spacing estimation (MSP), magnetic resonance imaging (MRI)

1. INTRODUCTION

The magnetic resonance imaging (MRI) datasets are affected by inherent noise as a consequence of different factors, e.g., eddy currents, inhomogeneity of magnetic field, and Johnson-Nyquist (thermal) noise coming from gradient coils [1, 2]. For that reason, it leads to numerous ambiguities in quantitative and qualitative MRI data analysis, e.g., diffusion tensor estimation from diffusion-weighted imaging (DWI) data [3, 4], diffusion tensor tractography algorithms [5], and so on. Moreover, the noise in MRI datasets has a negative impact not only on image segmentation and registration procedures results, but also influence on computational anatomy models accuracies [6]. Thus, an appropriate noise estimation technique is necessary, of which results can be exploited in further image processing stages such as image denoising.

It is known that single-coil MRI signal can be described as a complex Gaussian system after two-dimensional inverse discrete Fourier transform of the raw data from $k$-space [7, 8]

$$S(x) = A(x) + \eta(\sigma^2), \ x \in \Xi \subset \mathbb{Z}^2$$

where $A(x)$ is noise-free complex signal and $\eta(\sigma^2)$ is additive noise component defined as $\eta(\sigma^2) = \eta_R(\sigma^2) + i\eta_I(\sigma^2)$, where $\eta_R, \eta_I \sim \mathcal{N}(0, \sigma^2)$. Therefore, the MRI data magnitude $m$ is the envelope of the signal (1), which in fact, is Rician distributed with probability density function (PDF) [9]

$$f(m; \nu, \sigma) = \frac{m}{\sigma^2} \exp \left(-\frac{m^2 + \nu^2}{2\sigma^2}\right) I_0 \left(\frac{m\nu}{\sigma^2}\right), \ m \geq 0$$

where $\nu$ denotes noise-free signal amplitude and $I_0$ is the modified Bessel function of the first kind and zeroth order. When the signal-to-noise ratio decreases to zero, i.e., $\frac{\nu}{\sigma} \to 0$, the Rician PDF (2) boils down to the Rayleigh PDF given by

$$f(m; \sigma) = \frac{m}{\sigma^2} \exp \left(-\frac{m^2}{2\sigma^2}\right), \ m \geq 0$$

Up till now, plenty of noise estimation techniques in single-coil systems were proposed including least-squares fitting the Rayleigh PDF to the background data histogram by Brummer et al. [10], its extension to maximum likelihood (ML) principle derived by Sijbers et al. [11] and Parzen window density estimation suggested by Chang et al. [12]. In [13], the scheme based on the second-order Rician moments of double acquisition was proposed. However, it suffers from ghosting artifacts as well as requires prior image registration procedure. More recent techniques include noise estimation based on local statistics, i.e., Ajé’s methods [14], Brummer’s and Sijbers’s methods generalizations by Aja et al. [8], variance and skewness estimation in foreground magnitude MRI data [15] as well as the variance-stabilization approach [16].

The common drawbacks of cited methods, whether they estimate noise in a global manner (prior background segmentation is needed) or in local neighbourhoods, are large bias and variance. In this paper, however, we focus on the former case and derive new background MRI data noise estimators based on the maximum spacing estimation (MSP) principle. Proposed MSP approach is characterized by a relative low bias and reaches the Cramér-Rao lower bound on the variance for all $\sigma$ values. Surprisingly, the MSP estimation technique has not been used in signal and image processing tasks yet, thus, it still remains unfamiliar in this community.

The paper is organized as follows. In the next section new MSP estimator based on Kullback-Leibler (KL) divergence

This work is founded by the AGH University of Science and Technology as a research project No. 11.11.120.612.
In view of the cumulative distribution function (CDF) properties that \( \lim_{m \to -\infty} F_\xi (m; \sigma) = 0 \) and \( \lim_{m \to \infty} F_\xi (m; \sigma) = 1 \), we simplify \( \psi_1 \) and \( \psi_{n+1} \) cases as

\[
\psi_1(m; \sigma) = F_\xi (m_{(1)}; \sigma) \\
\psi_{n+1}(m; \sigma) = 1 - F_\xi (m_{(n)}; \sigma)
\]

### 2. THE MAXIMUM SPACING ESTIMATION

The MSP principle has been proposed independently by Cheng [17] and Ranneby [18] as a general method for parameters estimation of univariate continuous distributions. In comparison with ML technique, the MSP approach is based on spacings between successive order statistics and can produce consistent and asymptotically efficient estimators.

Let \( f \) and \( g \) be probability density functions and suppose that \( F = \{f(m; \sigma); \sigma \in \Theta\} \) is a parametric family of PDFs. Then, the KL divergence between \( f \) and \( g \) is given by [19]

\[
\int_{\infty}^{\infty} g(m) \log \left( \frac{g(m)}{f(m; \sigma)} \right) \, dm
\]  

(4)

Suppose that \( \mathbf{m} = \{m(1), \ldots, m(n)\} \) are realizations of the random variable \( \xi \) with Rayleigh PDF (3), which in fact, are intensities in the background of the MRI image. Now assume that \( \mathbf{m} \) is the order statistics, i.e.,

\[
-\infty = m(0) \leq m(1) \leq \ldots \leq m(n) \leq m(n+1) = \infty
\]  

(5)

Using the mean value theorem, it can be shown that minimizing (4) is equivalent to maximizing the expression [18, 20]

\[
\hat{\sigma}_{\text{MSP}} = \arg \max_{\sigma \in \Theta} Q_{\text{MSP}}(\sigma; \mathbf{m})
\]  

(6)

with the cost function \( Q_{\text{MSP}} \) and spacings \( \psi_i \) defined as

\[
Q_{\text{MSP}}(\sigma; \mathbf{m}) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log \left( (n+1)\psi_i(m; \sigma) \right)
\]  

(7)

\[
\psi_i(m; \sigma) = \int_{m(i-1)}^{m(i)} f(m; \sigma) \, dm, \quad i = 1, \ldots, n+1
\]  

(8)

In view of the cumulative distribution function (CDF) properties that \( \lim_{m \to -\infty} F_\xi (m; \sigma) = 0 \) and \( \lim_{m \to \infty} F_\xi (m; \sigma) = 1 \), we simplify \( \psi_1 \) and \( \psi_{n+1} \) cases as

\[
\psi_1(m; \sigma) = F_\xi (m_{(1)}; \sigma) \\
\psi_{n+1}(m; \sigma) = 1 - F_\xi (m_{(n)}; \sigma)
\]

#### 2.1. Rayleigh distribution

The CDF of the Rayleigh distributed real-valued random variable \( \xi \) is defined by the formula \( F_\xi (m; \sigma) = 1 - \exp \left( -\frac{m^2}{2\sigma^2} \right) \). Given (8-10), the appropriate first order spacings \( \psi_i \) in the cost function (7) we express in the following way

\[
\psi_i(m; \sigma) = \exp \left( -\frac{m_i^2}{2\sigma^2} \right) - \exp \left( -\frac{m_{i-1}^2}{2\sigma^2} \right)
\]  

(11)

\[
\psi_1(m; \sigma) = 1 - \exp \left( -\frac{m_{(1)}^2}{2\sigma^2} \right)
\]  

(12)

\[
\psi_{n+1}(m; \sigma) = \exp \left( -\frac{m_{(n)}^2}{2\sigma^2} \right)
\]  

(13)

Therefore, the \( Q_{\text{MSP}} \) cost function (7) takes the form

\[
Q_{\text{MSP}}(\sigma; \mathbf{m}) = \frac{1}{n+1} \left( \sum_{i=2}^{n} \log (\alpha_{i-1} - \alpha_i) \right)
\]  

(14)

\[+ \log (1 - \alpha_1 - \frac{m_{(n)}^2}{2\sigma^2}) + \log(n+1)\]

where \( \alpha_i = \exp \left( -\frac{m_i^2}{2\sigma^2} \right) \). For the extremum test of (14), see Appendix A. The derivative of \( Q_{\text{MSP}} \) with respect to \( \sigma \) is given then by

\[
\frac{dQ_{\text{MSP}}}{d\sigma} = \frac{1}{\sigma^3 (n+1)} \left( \sum_{i=2}^{n} \frac{\alpha_{i-1} m_{(i-1)}^2 - \alpha_i m_{(i)}^2}{\alpha_{i-1} - \alpha_i} \right)
\]  

(15)

\[+ \frac{\alpha_1}{\alpha_1 - 1} m_{(n)}^2 + m_{(n)}^2 \]
In numerical calculations, however, we have to adjust log functions in (14) and denominators in the expression (15) by a small correction factor $\delta > 0$ to avoid impermissible numerical calculations, i.e., when $m_{(i)} = m_{(i-1)}$ or $m_{(i)} = 0$.

### 2.2. Higher order spacings

Higher order spacings, which, in fact, also approximate Kullback-Leibler divergence, can be defined [20]. The spacings of the order $k$ $(1 < k < n)$ are given as

$$
\psi_{i}^{k}(m; \sigma) = F_{\xi}(m_{(i+k-1)}; \sigma) - F_{\xi}(m_{(i-1)}; \sigma)
$$

(16)

where $i = 1, 2, \ldots, n-k+2$. The spacings (16) define new $Q_{\text{MSP}}^{k}$ estimator of the order $k$. For $k = 1$, the spacings $\psi_{i}^{1}$ reduce to elementary ones defined by (8).

### 2.3. Different statistical distances approximation

Interestingly, MSP estimators for different than KL divergence information measures can be obtained. We derived new estimators of the order $k$ given J-divergence (symmetrized KL divergence), Rényi divergence and Vajda’s measure (see Table 1). The $i$-th component of the cost function is a concave function defined as $(1 - \beta \psi_{i}) \log(\beta \psi_{i})$, $\text{sgn}(1 - p) \beta^{p} \psi_{i}^{p}$ and $-1 + \beta \psi_{i}$ for J-divergence, Rényi divergence and Vajda’s measure, respectively. The $\beta$ parameter is defined as $\beta = \frac{n+1}{k}$. However, we note that the derivative of the cost function derived from the Vajda’s measure does not exist for $\alpha_{i-1} - \alpha_{i+k-1} = \frac{1}{\beta}$, $1 - \alpha_{k} = \frac{1}{\beta}$ and $\alpha_{n-k+1} = \frac{1}{\beta}$, thus, gradient-based optimization techniques cannot be used.

### 3. NUMERICAL EXPERIMENTS

To verify our noise estimation technique in a numerical way, we carried out experiments based on the synthetic BrainWeb T1-weighted MRI phantom with 1mm slice thickness and 0% field intensity non-uniformity as well as real DWI data of the fixed canine heart (for the description of the cardiac data, see [22, 23]). We compared our method (denoted as MSP-Rayleigh) with Aja’s method [14], eq. 18), Brummer’s method [10], Brummer-Aja’s method (8), eq. 20), Chang’s method [12], maximum likelihood approach (ML-Rayleigh) [14], eq. 6), Sijbers’s method [11], and Sijbers-Aja’s method [8], eq. 24). For local methods, i.e., Aja’s, Brummer-Aja’s, and Sijbers-Aja’s, estimation on entire images were conducted. For all other approaches, which, in fact, are non-local, we used prior background segmentation provided by BrainWeb phantom as well as Sauvola & Pietikäinen [24] method for real data. We used gradient descent optimization method to find the minimum of $Q_{\text{MSP}}^{k}$ functions, except $Q_{\text{MSP}}^{2}$ derived from the Vajda’s measure, where the Nelder-Mead method was arranged.

We performed Monte Carlo (MC) simulations with noise

---

1The MATLAB source code of the proposed method can be downloaded at http://www.mathworks.com/matlabcentral/fileexchange/45236
level between $\sigma = 5$ and $\sigma = 50$ with $\sigma_{\text{step}} = 5$ step, and 1000 iterations for each $\sigma$ value (see Fig. 1a–f and corresponding Peak Signal-to-Noise Ratios (PSNRs)). We also calculated Fisher’s information of the likelihood function for the Rayleigh distribution, thus, the Cramér-Rao lower bound (CRLB) for an unbiased estimator is $\text{CRLB}_\theta(\sigma) = \sigma^2 / 4n$.

Figures 2a and 2b show mean bias and variance of the MSP-Rayleigh noise estimator in transverse plane of the brain phantom in comparison with literature reports over 1000 trials. For low noise level MSP-Rayleigh is overtaken by cited methods, nevertheless, for higher noise level our approach shows the best performance among others. Moreover, the MSP-Rayleigh estimator has the lowest variance among others and reach CRLB bound for all $\sigma$ values. In Fig. 2c noise estimates for best methods over ten successive slices in coronal plane and actual noise level $\sigma = 25$ are also compared. It is clearly shown that our method remains rather small bias and has a low dispersion. We call attention to the ML estimation, though has similar results to MSP, for higher noise level ($\sigma \geq 20$) departs from our proposal in terms of bias (Fig. 2a).

Second experiment concerns noise level estimation on very long averaged real canine cardiac DWI data without ground truth\(^\text{3}\) (Fig. 1g). Obtained results also confirm that ML and MSP are comparable as well as other techniques are characterized by a positive bias (Table 2). However, we note that, if the image background is improperly segmented, i.e., includes skull regions, ML and MSP methods fail.

In the last experiment we also conducted MC simulations, as in experiment one, to assess $Q_{\text{MSP}}$ estimators using higher order spacings (16) and different statistical distances, which are presented in Table 1. Figures 3a and 3b show that we can obtain even more accurate estimates including higher order spacings. We call attention to the estimator inferred from Rényi divergence approximation, although has a greater bias and variance in comparison with KL and J-divergence, it does not require special treatment with non-unique samples.

\(^\text{3}\)We are thankful to Drs. Patrick A. Helm and Raimond L. Winslow at the Center for Cardiovascular Bioinformatics and Modeling and Dr. Elliot McVeigh at the National Institute of Health for provision of data.

### Table 2. Estimated noise level value $\hat{\sigma}$ based on the entire real cardiac DWI dataset without ground truth

<table>
<thead>
<tr>
<th>Method</th>
<th>Prior segmentation</th>
<th>Entire image</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aja’s method [14]</td>
<td>–</td>
<td>11.163</td>
</tr>
<tr>
<td>Brummer’s method [10]</td>
<td>11.798</td>
<td>–</td>
</tr>
<tr>
<td>Sibiers-Aja’s method [8]</td>
<td>–</td>
<td>12.045</td>
</tr>
<tr>
<td>MSP-Rayleigh (KL divergence)</td>
<td>10.195</td>
<td>–</td>
</tr>
</tbody>
</table>

4. CONCLUSIONS

In this paper we proposed and verified in a numerical way novel noise level estimation method in single-coil MRI background data based on the MSP principle. We brought the MSP technique into the image processing field and derived new estimator for Rayleigh distributed data, which had not been considered in the literature before. Moreover, we generalized the MSP approach and derived cost functions of the order $k$ for $J$-divergence, Rényi divergence and Vajda’s measure as well. Finally, for Kullback-Leibler divergence approximation our noise estimation approach is characterized by a relative low bias and meets the Cramér-Rao lower bound on the variance.

The MSP estimation technique can be naturally extended to the Rician distribution including parallel SENSE reconstructed data as well as parallel GRAPPA reconstructed data which follow (non)central chi distribution [8]. Nevertheless, that is beyond the scope of this paper and will be investigated in the further research. We draw attention to the fact that MSP technique can be exploited in other image processing tasks like image segmentation and registration procedures. Furthermore, MSP estimation principle is a promising tool for wider image processing applications, e.g., parameter estimation in mixture models.

5. APPENDIX A

The second derivative test for the (14) at a critical point $\hat{\sigma}_{\text{MSP}}$

$$\left. \frac{d^2 Q_{\text{MSP}}}{d\sigma^2} \right|_{\sigma = \hat{\sigma}_{\text{MSP}}} = - \frac{3}{\hat{\sigma}^2_{\text{MSP}}} \frac{dQ_{\text{MSP}}}{d\sigma} \bigg|_{\sigma = \hat{\sigma}_{\text{MSP}}} = \frac{1}{\hat{\sigma}^6_{\text{MSP}}(n+1)} \left[ \sum_{i=2}^{n} \frac{\alpha_{i-1}\alpha_i}{\alpha_i(\alpha_i - 1)^2} \left( m_{(i-1)}^2 - m_i^2 \right) \right]$$

In result, $\frac{d^2 Q_{\text{MSP}}}{d\sigma^2} \bigg|_{\sigma = \hat{\sigma}_{\text{MSP}}} < 0$, thus, the critical point $\hat{\sigma}_{\text{MSP}}$ is a relative maximum.
6. REFERENCES


