An automatic method for estimating noise-induced signal variance in magnitude-reconstructed magnetic resonance images

Lin-Ching Chang[*a], Gustavo K. Rohde[ab], Carlo Pierpaoli[a]
[a]NICHD, National Institutes of Health, Bethesda, Maryland, USA.
[b]Applied Mathematics and Scientific Computation Program,
University of Maryland, College Park, Maryland, USA.

ABSTRACT

Signal intensity in magnetic resonance images (MRIs) is affected by random noise. Assessing noise-induced signal variance is important for controlling image quality. Knowledge of signal variance is required for correctly computing the chi-square value, a measure of goodness of fit, when fitting signal data to estimate quantitative parameters such as T1 and T2 relaxation times or diffusion tensor elements. Signal variance can be estimated from measurements of the noise variance in an object- and ghost-free region of the image background. However, identifying a large homogeneous region automatically is problematic. In this paper, a novel, fully automated approach for estimating the noise-induced signal variance in magnitude-reconstructed MRIs is proposed. This approach is based on the histogram analysis of the image signal intensity, explicitly by extracting the peak of the underlining Rayleigh distribution that would characterize the distribution of the background noise. The peak is extracted using a nonparametric univariate density estimation like the Parzen window density estimation; the corresponding peak position is shown here to be the expected signal variance in the object. The proposed method does not depend on prior foreground segmentation, and only one image with a small amount of background is required when the signal-to-noise ratio (SNR) is greater than three. This method is applicable to magnitude-reconstructed MRIs, though diffusion tensor (DT)-MRI is used here to demonstrate the approach.

Keywords: signal variance, noise variance, magnitude-reconstructed MRI, density estimation.

1. INTRODUCTION

Magnetic resonance (MR) images are inevitably corrupted by random noise introduced during data acquisition. Estimation of the noise-induced signal variance in images is important for many reasons including its use as a measure of image quality assessment and MR system analysis. Knowledge of signal variance is also required for many image processing tasks such as image registration and filtering, and for functional MRI analysis. Moreover, it is needed to correctly compute the chi-square value, a measure of goodness of fit, when fitting signal data to estimate quantitative parameters such as T1 and T2 relaxation times or diffusion tensor elements.

Noise variance σ^2_{noise} is commonly computed by selecting a signal-free background region manually. Signal variance σ^2 can be estimated by using the appropriate bias correction factor described by Henkelman et.al. [1, 2], i.e., $\sigma^2 = \sigma^2_{noise} / 0.655$. Automated object segmentation can be used to extract the background of an image; however, the background of MR images, particularly those acquired using echo planar imaging techniques, are rarely free from ghosting and other artifacts. Using the whole background region in estimating the noise variance could yield a biased signal variance. In addition, the results of an extracted background region often depend on the availability of accurate and reliable segmentation procedures.

An approach takes into account that the noise measurement of magnitude-reconstructed MR image measurement can be characterized by a Rayleigh distribution [3]. The mean and variance of the noise in the background of magnitude-reconstructed images are estimated by fitting the Rayleigh distribution to an estimated probability density function of

pixel intensities. However, this implementation is susceptible to local optima encountered during the non-linear data fitting procedures. Furthermore, in a typical MR image, a bimodal or multi-modal distribution (see Fig. 1) containing a mixture of background and foreground values can be observed. Determining where to fit presumable distributions to the intensity histogram is not a trivial task
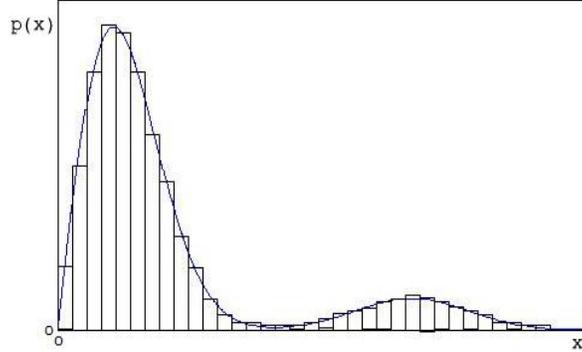

Fig. 1. Bimodal distribution of a typical MR image

Another approach for estimating the signal variance in MR images is the so-called double acquisition method [4]. This approach, however, requires the use of two identical images (with the exception of noise) which are impractical to obtain due to physiological motion that may occur between acquisitions. Moreover, artifacts such as ghosting, magnetic susceptibility induced distortions, etc., may not be identical during two separate acquisitions.

A novel method for estimating the noise-induced signal variance in magnitude-reconstructed MR images is proposed to overcome the shortcomings associated with the above methods. This proposed method is completely automatic and does not depend on prior foreground segmentation. It requires only one image with a small amount of background as long as the signal-to-noise ratio (SNR) is greater than three, though more images can be used to reduce the bias in the estimation.

2. METHODS

Suppose A_r and A_i are the real and imaginary data corrupted by Gaussian distributed noise with zero mean and standard deviation σ. The probability density function (*pdf*) of the magnitude reconstructed data, $A = \sqrt{A_r^2 + A_i^2}$, is given by a Rician distribution (E1) where M is the magnitude image intensity and β_0 is the zeroth-order Bessel function of the first kind. At low SNR, this distribution approaches the Rayleigh distribution (E2) with mean $\overline{M} = \sigma\sqrt{\pi/2}$ and variance $\sigma_{noise}^2 = (2-\pi/2)\sigma^2$. At high SNR, the noise distribution approaches a Gaussian distribution (E3) with mean $\overline{M} = \sqrt{A^2+\sigma^2}$ and variance $\sigma_{noise}^2 = \sigma^2$.

$$p(M) = \frac{M}{\sigma^2} e^{-(M^2+A^2)/(2\sigma^2)} \beta_0\left(\frac{A\cdot M}{\sigma^2}\right) \qquad \text{(E1)}$$

$$p(M) = \frac{M}{\sigma^2} e^{-M^2/(2\sigma^2)} \qquad \text{(E2)}$$

$$p(M) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(M-\sqrt{A^2+\sigma^2})^2/(2\sigma^2)} \qquad \text{(E3)}$$

Note that the value of M for which $p(M)$ in (E2) is maximum equals σ. This finding can be easily checked by differentiating (E2) with respect to M, setting the result equal to zero, and solving for M.

$$\frac{\partial p(M)}{\partial M} = \frac{1}{\sigma^2} e^{-M^2/(2\sigma^2)} + \frac{M}{\sigma^2} (\frac{-2M}{2\sigma^2}) e^{-M^2/(2\sigma^2)} = 0$$

$$(1 - \frac{M^2}{\sigma^2}) \frac{1}{\sigma^2} e^{-M^2/(2\sigma^2)} = 0 \quad \rightarrow \quad \sigma = M$$

This information can be used to extract the standard deviation of the signal intensity σ in the images by simply identifying the peak of the noise distribution. The corresponding peak location is thus equal to σ.

2.1 Density Estimation Algorithm

2.1.1 Frequency histograms

A frequency histogram is used to summarize the intensity distribution of a univariate pixel data set (see Fig. 1). Given a location x_0 and a bin width h, the frequency histogram can be written as

$$f(x_0) = \frac{1}{nh} \times \text{ number of data points that fall in the interval } [x_0 - h/2, x_0 + h/2),$$

where n is the number of total data points.

Choosing the number of bins or, equivalently, the width of each bin is important. If too many bins are used in constructing the histogram, the width of each bin, h, becomes smaller and the resultant histogram will show lots of variation. In contrast, if the width of each bin is too large, the result may not show well important features of the data such as bimodality. For normal distributed data, several proposed rules can ideally construct the corresponding frequency histogram that represents the underlying probability density function [5-7]. However, noise in MR data is known to be Rician distributed and a bimodal or multi-modal distribution is usually assumed. We aim to develop an algorithm to compute an optimal bin width for this application. It uses the Sturges rule, $log_2(n)+1$, as the initial number of bins [5] to compute the number of data points in a bin that contains the first peak of the histogram, say n_p. The number of bins (or bin width) is changed dynamically until n_p reaches a predefined number, say n_{peak}. (i.e., the number of bins in the next iteration will be used in the current iteration by multiplying a default ratio, n_{peak} / n_p). The choice of n_{peak} is empirically set to 1% of the total data points for the diffusion tensor (DT)-MR data in our study.

2.1.2 Parzen density estimation

Apart from the histogram binning, kernel estimator is also commonly studied. The advantage of kernel probability density estimation over the histogram binning is their smoothness. We use Parzen density estimation, a nonparametric density estimator, to estimate the peak of the noise distribution. The general form of this kernel estimator can be written as follows,

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^{n} K(\frac{x - x_i}{h}), \text{ where } \int_{-\infty}^{\infty} K(t)dt = 1, \text{ and } K(t) \geq 0, \forall t.$$

The function $K(t)$ is the kernel or basis function. The parameter h is the bandwidth, also called the window width or smooth parameter. We chose a Gaussian kernel as it is easy to manipulate and derive. The kernel estimator with a standard Gaussian basis function can be viewed as the sums of many small Gaussian bumps. The kernel size or window width is important and sometimes adapted to each application. The window width can be computed by minimizing the mean square error between the true and estimated density. In our simulation, the window width, h, is set to $1.06 \times s \times n^{-1/5}$ as suggested by Silverman [10], where s is the sample standard deviation and n is the sample size. The peak can be computed using an optimization algorithm such as the golden section search or Brent's method in one dimension [8]. We converted the problem of maximization to minimization by placing a negative sign on the function and used an IDL (Interactive Data Language, Research System Inc.) built-in function, the DFPMIN procedure, to find the peak.

2.2 Computational Consideration

To estimate the signal variance in a 3D MRI volume, each slice is used individually as a subset to get the estimate of signal variance σ_i^2. The minimum σ_i^2 is then selected as the signal variance for the whole volume. By selecting the minimum σ_i^2, we can obtain a value from the least "object signal" contaminated slice. In our DT-MRI data set that contains multiple volumes, the slices coming from different volumes, but with the same slice location, are used to form a subset for the estimation. For the Parzen window approach, the minimum estimated value is taken as the signal variance for the entire data set. For the histogram approach, however, the quartile value is taken instead. (See the Discussion Section for an explanation for this implementation.)

3. DATA

3.1 Monte Carlo Simulation

A Monte Carlo simulation is performed to validate the proposed method. Since the requirement of our approach is a single image with certain signal-free regions (i.e., image background that contains only thermal noise), we are interested in knowing how much background is needed to accurately estimate the noise-induced signal variance in the object under different SNR. To simulate images with different SNRs we created synthetic images containing objects of different sizes and added Gaussian distributed noise in quadrature.

Artifacts in the background of an MR image usually cause signal intensities to be higher than their normal values. To simulate this situation, we randomly corrupted some background points by increasing their intensity values by 50%. Our objective is to test the accuracy and robustness of the proposed method when different level of artifacts appeared in the image background under different SNR.

3.2 MR data acquisition

In order to assess the performance of the proposed method in real MR data, we acquired the brain DT-MR data sets from seven healthy volunteers. An image with RF turned off was acquired for each data set before subject scanning to provide a baseline for noise variance estimation. Each volunteer was scanned three times on different dates under the same setting to test the variability on noise assessment. Images were acquired on a CNV LX 1.5 GE MRI System (General Electric, Milwaukee, WI) with a diffusion-weighted, spin-echo, single-shot, EPI sequence with $2 \times 2 \times 2$ mm^3 resolution and 72 to 80 slices to cover the whole brain.

4. RESULTS

Figure 2 shows the simulation result of the proposed method on the amount of background required to properly estimate the peak of the noise distribution. If the error of the estimated signal standard deviation is set within 10%, 65% of the background in an image is required when SNR =3; 22% is required when SNR=4; and only a small amount of background is needed when SNR $\geq$ 5. In general, more background provides a better result. If the background is less than required, for example, less than 60% when SNR=3, the estimated signal standard deviation is somewhat over-estimated. This result is understandable because the noise distribution is contaminated by the "object signal" and the mixture of noise and object signal will always cause the estimated peak to shift to the right.
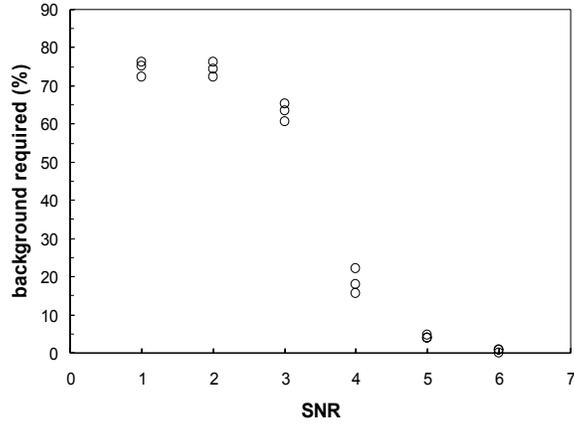
Fig. 2 The background required for the kernel estimator to properly estimate
the signal standard deviation ($0.9\sigma < \sigma_{estimated} < 1.1\sigma$) from an image (512*512).

The simulation also shows that the proposed method is robust to artifacts if the artifacts in the image background cause the pixel intensities to be higher than their normal value. Figure 3 shows the error on estimated signal standard deviation when the artifact has a pixel intensity 50% higher than its normal value. Object size in this simulation is set at 30% of the image. Note that the error for the estimated signal standard deviation is within 15% even when the percentage of artifacts in the background reaches 60%.
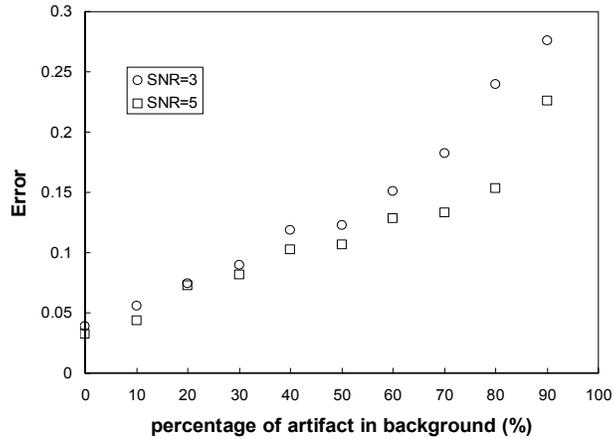

Fig. 3 The error of estimated signal standard deviation with different
percentages of artifacts in the background. Object size is 30% of the image.
Artifacts' pixel intensity is 50% higher than normal.

For the DT-MR data images in our study, the SNR for all of 21 data sets was about 20. The true signal standard deviation for each data set was computed using the noise-only image (RF turned off) to provide a baseline value for comparison. Figure 4 shows the estimated noise-induced signal standard deviation using (1) the Parzen window approach, (2) the histogram approach, and (3) the conventional method where we manually select 30 signal-free regions from the background and compute the average of the estimated parameter. The result shows all three methods yield a reasonable estimation of the signal standard deviation with the Parzen window approach having the best result overall.
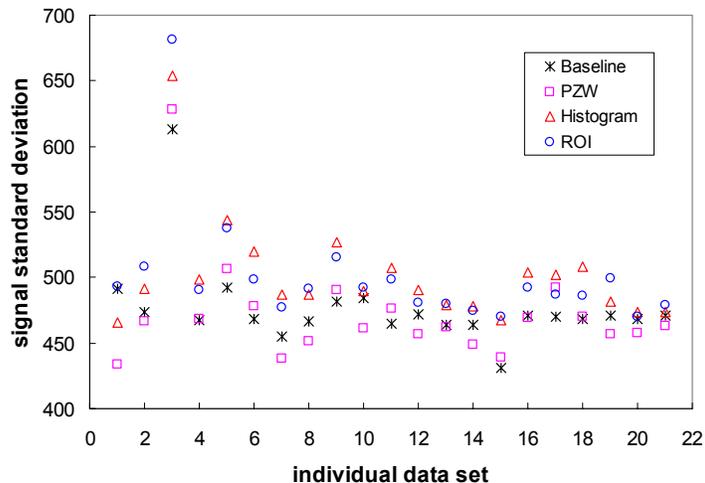
Fig. 4 Comparison of the baseline and estimated signal standard deviation using the Parzen window approach, the frequency histogram approach and the conventional method. For the conventional method, we manually select 30 signal-free regions in the image background.

Although all three methods gave similar results for estimating the signal variance (see Fig. 4), the deviation of the estimated parameter using different slices varies. The Parzen window approach has the lowest average deviation which is about 23.24 (less than 5% of error). The histogram approach has an average deviation of about 47.19 (less than 10% of error). The conventional method has an average deviation of 25.63 if the background regions are selected from the bottom few slices that result the value in Fig 4. The bottom few slices in our DT-MR data set is the slices which contain large amounts of background. If the background regions are select from each slice, then the deviation is larger and estimated signal variances are varied.

5. DISCUSSION

Most statistical packages use Sturges rule, $log_2(n)+1$, as the number of bins [5]. However, the Sturges rule leads to an oversmoothed histogram especially for large samples (for n greater than 200) and underestimates the peak [9]. Alternative rules for constructing a histogram, specified in terms of bin width, h, include Scott's rule [6], $h = 3.5\hat{\sigma} n^{-1/3}$, and Freedman and Diaconis's rule [7], $h = 2Rn^{-1/3}$, where $\hat{\sigma}$ is the sample standard deviation and R is the interquartile range of the sample. Although these rules are based on constructing a histogram from normal data and lead to an oversmoothed histogram for large samples [9], we used the Sturges rule to provide an initial guess (other rules can be used) with the final number of bins depending on a predefined empirical choice (as described in the Methods section). In addition, as we are interested in finding the location of the peak instead of the peak's amplitude, the oversmoothed histogram has less effect on the location of the peak.

For the kernel estimator, the exact choice of basis function is not very important as long as it is smooth and bell-shaped [10]. The choice of a smoothing parameter, h, is important, but there is no firm rule governing the choice. The best choice of window width, h, may require further experimentation to balance the desired smoothness of the estimator with the right amount of oversmoothing to avoid overlooking real variation in the density. As with the histogram approach, a number objective methods remove some of this subjective choice, such as the rule we used, $h = 1.06 \times s \times n^{-1/5}$ or $h = 1.06 \times \min(s, R/1.34) \times n^{-1/5}$, where s is the sample standard deviation and R is the interquartile range from the sample [10].

The result in Figure 4 used the slices from each volume as a subset and a set of signal standard deviation { σ_i | i is the slice number} was obtained. For large SNR (> 5), the standard deviation of σ_i should be small since the object-signal

distribution is less contaminated by the noise (i.e., the noise and the object-signal distributions are better separated). For a small SNR, the high intensity part of noise distribution may mix with the low intensity part of object-signal distribution such that the peak of the noise distribution may be shifted to the right. The minimum of σ_i should thus be considered as the signal standard deviation for the entire data set. For the histogram approach, the bin width was set empirically and when the variation of the histogram is large, the local maximum might be identified as the peak of noise distribution. Consequently, using the quartile of σ_i is recommended.

Because of its accuracy, the Parzen window method is the best choice for estimating signal variance, but it requires a long computation time. To reduce this length, any slice in a volume can be used for estimation because the deviation of estimated σ_i using different slices is small. However, because more background in general will provide a better result, we suggest using the first and the last slices for estimation, taking the one with the smallest estimated value.

6. CONCLUSION

An automatic method for estimating the signal variance in magnitude-reconstructed MRI is presented. This method works on only one image to estimate the signal variance, does not require any user interaction as no destination of background pixels is needed, and does not require prior brain segmentation. The proposed method is less subjective to artifacts than the conventional manual object-free background selection. Also, the proposed method yields an estimated signal variance with lower variability than that resulting from manual background selection. The proposed method is useful in creating a pipeline in processing massive MR data.

ACKNOWLEDGEMENT

REFERENCE

1   R. M. Henkelman, "Measurement of Signal Intensities in the Presence of Noise in MR Images," *Medical Physics*, vol. 12, pp. 232-233, 1985.
2   H. Gudbjartsson and S. Patz, "The Rician Distribution of Noisy MRI Data," *Magnetic Resonance in Medicine*, vol. 34, pp. 910-914, 1995.
3   M. E. Brummer, R. M. Mersereau, R. L. Eisner, and R. R. J. Lewine, "Automatic Detection of Brain Contours in MRI Data Sets," *IEEE Transactions on Medical Imaging*, vol. 12, pp. 153-166, 1993.
4   J. Sijbers, A. J. den Dekker, J. Van Audekerke, M. Verhoye, and D. Van Dyck, "Estimation of the Noise in Magnitude MR Images," *Magnetic Resonance Imaging*, vol. 16, pp. 87-90, 1998.
5   H. Sturges, "The Choice of a Class-interval.," *J. Amer. Statist. Assoc.*, vol. 21, pp. 65-66, 1926.
6   D. W. Scott, "Optimal and Data-Based Histograms," *Biometrika*, vol. 66, pp. 605-610, 1979.
7   D. Freedman and P. Diaconis, "On the Histogram as a Density Estimator - L2 Theory," *Zeitschrift Fur Wahrscheinlichkeitstheorie Und Verwandte Gebiete*, vol. 57, pp. 453-476, 1981.
8   W. H. Press, Teukolsky S. A., Vetterling, W. T., Flannery B. P., *Numerical Recipes in C: The Art of Scientific Computing, 2nd edition*. New York: Cambridge University Press, 1992.
9   D. W. Scott, *Multivariate Density Estimation: Theory, Practice and Visualization*. New York: John Wiley & Sons, 1992.
10   B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. New York: Chapman and Hall, 1986.