

# A novel method for evaluating and improving the $^1\text{H}$ -MRSI glioma data quality\*

YUAN Kehong<sup>1,2</sup>, LU Hongyu<sup>2</sup>, BAO Shanglian<sup>2\*\*</sup>, CHEN Qiansheng<sup>1</sup>,  
LI Shaowu<sup>3</sup> and DUAN Chaijie<sup>2</sup>

(1. Key Laboratory of Pure and Applied Mathematics, School of Mathematical Sciences, Beijing 100871, China; 2. The Research Center for Tumor Diagnosis and Therapeutical Physics & the Beijing Key Laboratory of Medical Physics and Engineering, Peking University, Beijing 100871, China; 3. The Tian Tan Hospital, Beijing 100050, China)

Received November 19, 2004; revised January 18, 2005

**Abstract** Metabolic information obtained by proton magnetic resonance spectroscopic imaging ( $^1\text{H}$ -MRSI) has been approved to be a powerful tool to identify either benign or malignant glioma, as well as to confirm the tumor level. However,  $^1\text{H}$ -MRSI data are affected by various factors, such as the thermal noise, eddy currents, susceptibility artifacts, and rigid body motion. To get accurate quantitative metabolic information, the key problem is to assess the  $^1\text{H}$ -MRSI data quality. In this paper, we introduce a new evaluating system to filter the data, and a new method, called wavelet denoising method, to improve the data quality under the evaluating system. Experimental results on  $^1\text{H}$ -MRSI glioma data demonstrate that preprocessing is prerequisite and the proposed algorithm with evaluating system is effective.

**Keywords:**  $^1\text{H}$ -MRSI, evaluating system, glioma data quality, wavelet denoising

Proton magnetic resonance spectroscopic imaging ( $^1\text{H}$ -MRSI) has been intensively studied to quantitatively analyze glioma for clinician<sup>[1-3]</sup>, including defining the tumor boundary, tumor property and determining the tumor level. However, raw glioma data of  $^1\text{H}$ -MRSI usually involve various artifacts, such as the thermal noise, eddy currents, susceptibility artifacts, rigid body motion, physiological pulsation flow and hardware issues, which significantly affect the accuracy of the measured results.  $^1\text{H}$ -MRSI is different from magnetic resonance imaging (MRI) in which many artifacts are visible, and the data are just series of discrete values, so we cannot assert the data quality. Roland<sup>[4]</sup> proposed an extreme example to explain the pitfalls, and Knaap<sup>[5]</sup> found that quantitative analysis would reveal a 50% loss of all metabolites and hence yield the diagnosis of vanishing white matter disease. Therefore, evaluating and improving the data quality is necessary before quantitative analysis. Meanwhile, evaluating system should be credible and robust for different patients and different time<sup>[6-8]</sup>, the improving method should maintain intrinsic property.

In this paper, we introduce a novel method for

preprocessing using the wavelet analysis and an evaluating system which combines three numerical indexes of the spectrum data. The data set of 18 patients of glioma is used for our methodological study, which have been well diagnosed in Tiantan Hospital. The data of  $^1\text{H}$ -MRSI and the anatomic images have been acquired using GE 3.0T MRI facility, and raw data of 9 from 18 patients have been analyzed quantitatively by Quan et al.<sup>[3]</sup>. Our results are in good agreement with their results, which can verify the feasibility of our method.

## 1 Materials and methods

### 1.1 Clinical subjects

The human subjects involved in our study were 18 glioma patients including 10 female and 8 male subjects, aged from 30 to 56 years old. After the data acquisition in 3.0T MRI, the patients were treated by surgery. The positions of the biopsy were well decided by neural surgeons before the surgery according to the imaging data of MRSI and MRI, which included the region of the margin where cancer cells may be existent. The pathological conclusions were: 8 cases were at grade IV, 5 cases at grade III, and 5 cases at

\* Supported by National Natural Science Foundation of China (Grant Nos. 10275003 and 10175004), and the Natural Science Foundation of Beijing (Grand No. 3011002)

\*\* To whom correspondence should be addressed. E-mail: bao@pku.edu.cn

grade II. We selected 512 spectrum data of voxels for evaluating and denoising samples. All the images that the spectrum data came from were used by surgeries and clinician.

### 1.2 Magnetic resonance spectroscopy imaging

All glioma  $^1\text{H}$ -MRSI data were acquired with a 3.0T GE Signa scanner equipped with a quadrature head coil in Beijing Tiantan Hospital. Each examination began with a series of magnetic resonance imaging sequences, with parameters that TR is 2400 s, TE is 24 s, TI is 860, flip is  $90^\circ$ , image matrix is  $256 \times 256$ , FOV is 240 mm, NEX is 1, and slices is 5 mm. The IR volume was used as a reference for the multivoxel spectrum study. After IR imaging, the chemical shift imaging (CSI) with the point resolved spectroscopy (PRESS) was used to measure the spectra in the region containing both the lesion and the normal-appearing tissues. The parameters for CSI included TR 2000 s, TE 144 s, PHASE encode matrix  $256 \times 16$ , and slices 20 mm. Water suppression was achieved with additional chemical shift selective (CHESS) pulses.

### 1.3 Wavelet denoise with hard thresholding

In this subsection, we briefly introduce the wavelet decomposition and wavelet denoising method we adopted. For detailed information, see Refs. [10–12].

Let the model under consideration be  $X = D + \varepsilon$ , where  $X$  is the observation signal,  $D$  contains a deterministic signal, and  $\varepsilon$  is a Gaussian noise with zero mean. By Discrete Wavelet Transform (DWT),

$$\begin{aligned} w &= \text{DWT}(X) = \text{DWT}(D) + \text{DWT}(\varepsilon) \\ &= d + e, \end{aligned} \quad (1)$$

where  $w$  is the Discrete Wavelet Transform of observation signal  $X$ ,  $d$  is the transform coefficients of true single  $D$ ,  $e$  is the transform coefficients of noise  $\varepsilon$ . If  $\varepsilon$  is independent identical Gaussian distribution with a common variance  $\sigma^2$ , then after DWT, the  $e$ , wavelet coefficient of  $\varepsilon$ , is also independent identical Gaussian distribution with the common variance  $\sigma^2$ . In this scheme, the transform coefficient is in the vectors  $w_1, \dots, w_J$  and  $w$ . For denoising, only the coefficients in the  $w_k (1 \leq k \leq J)$  vectors are subjected to hard thresholding, but the elements of  $v_j$  are untouched, the equation for hard thresholding is

$$W_k^\delta = \begin{cases} W_k & |W_k| \geq \hat{\delta} \\ 0 & |W_k| < \hat{\delta} \end{cases} \quad (2)$$

where  $\hat{\delta}$  is the hard thresholding,  $W_k^\delta$  is the result of  $W_k$  subjected to hard thresholding  $\hat{\delta}$  Donoho<sup>[13]</sup> gave an excellent estimator of  $\hat{\delta}$

$$\hat{\delta} = \sigma \sqrt{2 \log N}, \quad (3)$$

where  $N$  is the number of measurements.

A practical procedure<sup>[11]</sup> is to estimate  $\sigma$  based upon the median absolute deviation (MAD) standard deviation estimate using just the  $\frac{N}{2}$  level  $j=1$  coefficients in  $w_1$ . By definition, this standard deviation estimator is

$$\hat{\sigma} = \frac{\text{median}\{|w_{1,0}|, |w_{1,1}|, \dots, |w_{1, \frac{N}{2}-1}|\}}{0.6745}. \quad (4)$$

The factor 0.6745 in the denominator rescales the numerator so that  $\sigma$  is also a suitable estimator for the standard deviation of the Gaussian white noise. The heuristic reason for calculating  $\sigma$  from the elements of  $w_1$  is that the smallest scale wavelet coefficients should be noise dominated, with the possible exception of the largest values. The MAD standard deviation estimate is designed to be robust against large deviations and hence should reflect the noise variance rather than the signal variance.

The hard thresholding algorithm consists of the following steps:

**Step 1:** Computing a level  $j (1 \leq j \leq J)$  partial DWT to obtain the coefficient vectors  $w_1, \dots, w_J$  and  $v_j$ , we have

$$w_{j,t} = d_{j,t} + e_{j,t}, \quad (5)$$

where  $j=1, \dots, J; t=0, \dots, N_j-1$ .

**Step 2:** Compute  $\hat{\sigma}$  using Eq. (4), because  $\sigma$  is unknown and  $\hat{\sigma}$  is its estimator.

**Step 3:** Specify the threshold level  $\hat{\delta}$  using Eq. (3).

**Step 4:** For  $w_{j,t}, j=1, \dots, J$  and  $t=0, \dots, N_j-1$ , apply the hard thresholding rule with  $w_{j,t}$  to obtain the thresholding coefficients  $w_{j,t}^\delta$ , which are then used to form  $w_j^\delta, j=1, \dots, J$ .

**Step 5:** Estimate  $D$  from  $D^\delta$  which is the inverse transform of  $w_1^\delta, \dots, w_J^\delta$  and  $v_j$ .

## 1.4 Evaluating system

Till now, no agreement has been reached on how to exactly define whether the spectrum is good or not. And indeed the quality criteria in MRSI do depend on many factors, for example, whether we deal with single voxel (SV) or multiple voxels of the spectroscopic imaging (SI) data, what kind of pulse sequences is used and what parameters are used for the data acquisition in the pulse sequence, and what product of MRI facility for what kind of imaged object both in organ classification and in the position in the organ. The quality criteria should be the signal to noise ratio (SNR) at a certain threshold or some individual metabolite concentration values, such as full-width at half-maximum peak height (FWHM)<sup>[4]</sup>, Cramer—Rao minimum variance bounds (CRMVB)<sup>[9]</sup>, etc.

We establish an evaluating system with three numerical indexes,<sup>[4–9]</sup> SNR, FWHM and CRMVB to judge whether the data is good or poor. The data is good if it satisfies three conditions: SNR is larger than 5, FWHM is between 0.07 ppm and 0.1 ppm, and CRMVB is higher than 50%. Here, SNR is often defined in the frequency domain as the height of the largest metabolite peak divided by the root-mean-square amplitude (SD) of the noise in a signal-artifact-free part of the spectrum; FWHM is the linewidth independent of the lineshape in the frequency domain, which can determine the resolution effective to discern spectral features; CRMVB is a statistical index without estimating the model or parameters with upper and lower confidence limits.

## 2 Results

We firstly process the glioma raw data with different wavelet bases, the minimum of residuum is based on Daubechies basis. Therefore, we employed it as the wavelet basis in the continuous study.

We selected 512 spectrum data from 18 patients as samples of our experiment, every spectrum data of those 512 data is in Region of interest (ROI) and bilateral spectrum. Figure 1 represents the raw data, denoised data and residual data of just one sample. After data denoising, the SNR increases 5.6 db, FWHM of NAA (N-acetylaspartate) changes from 0.064 to 0.078, and all the CRMVB are bigger than 50%. Based on our evaluating system, because FWHM of NAA is 0.064, lower than 0.07, this

voxel is poor. Here, we assume that if the MR Imaging of a spectrum data is good, the spectrum data is good. According to the statement and clinical subjects, all our samples are good and the sample should be also good. Figure 2 shows the image corresponding to the spectrum data of Fig. 1, the image is accepted by clinician. The data is changed to be good by wavelet denoising.

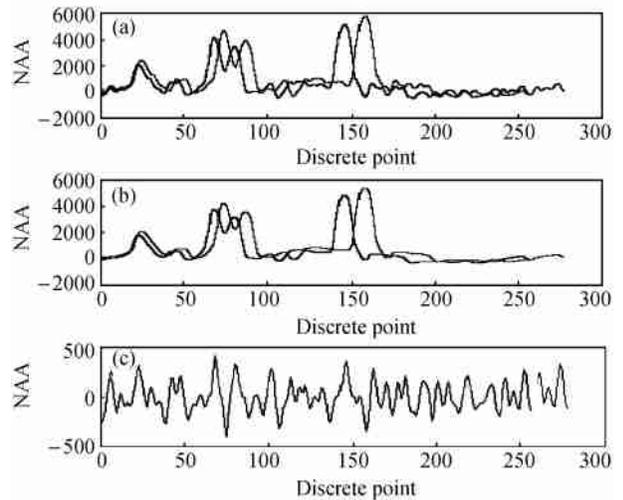


Fig. 1. Glioma <sup>1</sup>H-MRSI data of a sample. (1) Raw data; (b) denoised data; (c) residual data.

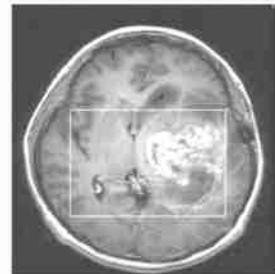


Fig. 2. MR Image corresponding to the spectrum data of Fig. 1.

We repeated the procedure for the 512 samples and computed the statistic results of SNR, FWHM and CRMVB, respectively. The AR (Accepted Ratio) was used for evaluation:

$$I_{AR} = \frac{N_{\text{good}}}{N_{\text{all}}} \times 100\%, \quad (6)$$

where  $I_{AR}$  is the AR index,  $N_{\text{good}}$  is the number of good voxels, and  $N_{\text{all}}$  is the number of all voxels.

The results given in Table 1 show that after data denoising, three indexes increase at different levels. The increments of SNR, FWHM and CRMVB are 4.23%, 10.35%, and 4.33%, respectively. All samples should be good, but after denoising, some samples are also poor. One of the reasons for this is

that the ROI is not exactly registered with the box of voxles in image, and the spectrum data of boundary voxels may be calculated with a part of the voxel. In addition, there are no samples found changing from good to poor by data denoising.

Table 1. Statistical results of 512 samples (%)

Data used	SNR	FWHM of NAA	CRM VB
Raw data	92.51	85.12	90.35
Denoised data	96.74	95.47	94.68

We selected 4 glioma patients in grade IV used by Quan<sup>[3]</sup> as samples and also adopted the z-scores

Table 2. Residual z-scores and metabolite ratios in 4 patients with glioma in grade IV

Data used	Cho/NAA z-scores		Cho/Cr z-scores	
	Average z-scores	Maximum z-scores	Average z-scores	Maximum z-scores
Raw data	7.30±1.83	15.81(9.95–25.97)	9.53±2.67	21.33(8.19–33.66)
Denoised data	7.29±1.42	15.76(10.02–22.45)	9.56±1.89	22.06(10.23–31.61)

### 3 Discussion and conclusion

From the above analysis, we can find that the evaluating system is important for asserting the data to satisfy requirements of quantitative analysis, and wavelet denoising is useful for improving the <sup>1</sup>H-MRSI glioma data quality, despite of the limited and special samples. However, for the method accepted by clinician, we should take various samples for test, for example, the blind data, different position, different tumor, or different MRI hardware. The evaluating system has been verified with different brain tumors, at different times for the same patient, etc.

In addition, we should investigate the better quantitative method for accurately segmenting the boundary of glioma, determining the area of glioma, increasing noninvasive robustness which are helpful for the imaging guided radiation therapy<sup>[14]</sup> in the future.

**Acknowledgement** The authors would like to thank Dr. Quan Hong, Wang Huiliang for giving their results for our comparison, and Dr. Huang Xinrui for some language corrections.

### References

1 Mcknight T. R. An automated technique for the quantitative assessment of 3D-MRSI data from patients with glioma. *Journal of Magnetic Resonance Imaging*, 2001, 13; 167–177.

method to calculate the two indexes; Cho/NAA z-scores and Cho/Cr z-scores. Table 2 shows the results of the two indexes using raw data and denoised data.

From the results of Table 2, we can find the data denoising process can improve the data quality, and the average of z-scores can be closer to the central value. Although the average values change a little, the classified results cannot be changed. These results are more acceptable for clinician, because they are based on the accepted data.

2 Pirzkall A. MR spectroscopy guided target delineation for high-grade glioma. *Int. J. Radiation Oncology Biol. Phys.*, 2001, 50 (4): 915–928.

3 Quan H. and Bao S. L. Diagnosis of glioma by multivoxel <sup>1</sup>H-MRSI. *Progress in Natural Science*, 2004, 14(9): 770–776.

4 Roland K. Issues of spectral quality in clinical <sup>1</sup>H-magnetic resonance spectroscopy and a gallery of artifacts. *NMR biomedicine*, 2004, 17(6): 361–381.

5 Knaap M. S., Barth P. G., Gabreels F. J. et al. A new leukoencephalopathy with vanishing white matter. *Neurology*, 1997, 48: 845–855.

6 Kubeka J., Burian M. and Hajek M. Quality control and routine stability check for single voxel <sup>1</sup>H *in vivo* quantitative magnetic resonance spectroscopy. *J. Magn. Reson. Anal.*, 1997, 3: 93–98.

7 Podo F., Henriksen O., Bovee W. M. et al. Absolute metabolite quantification by *in vivo* NMR spectroscopy: I. Introduction objectives and activities of a concerted action in biomedical research. *Magnetic Resonance Imaging*, 1998, 16: 1085–1092.

8 Certaines J. D. and Cathelineau G. Safety aspects and quality assessment in MRI and MRS: a challenge for health care systems in Europe. *J. MRI*, 2001, 13: 632–638.

9 Cavassila S., Daval S. and Huegen C. Cramér-Rao bounds: an evaluation tool for quantitation. *NMR Biomed.*, 2001, 14: 278–283.

10 Séphane M. Singularity detection and processing with wavelets. *IEEE Tran. Info. Theory*, 1992, 38(2): 617–643.

11 Donald B. P. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2000.

12 Séphane M. *A Wavelet Tour of Signal Processing*, 2nd ed. Singapore: Elsevier Ltd., 2003.

13 Donoho D. L. and Johnstone I. M. Adapting to unknown smoothness via wavelet shrinkage. *J. Am. Stat. Assoc.*, 1995, 90: 1200–1224.

14 Bao S. Imaging integration to satisfy the requirement of radiotherapy. *Australas. Phys. Eng. Sci. Med.*, 2004, 27(2): 91–93.