Volumetric Analysis of Multiple Sclerosis Using Multispectral MR Images: Method and Validation

Lihong Li, *Student Member, IEEE*, Hongbing Lu, *Member IEEE*, Xiang Li, Wei Huang, Alina Tudorica, Chris Christodoulou, Lauren Krupp, and Zhengrong Liang, *Member, IEEE*

***Abstract* --- We present a fully automatic mixture-based algorithm for segmentation of brain tissues (white and gray matters – WM and GM), cerebral spinal fluid (CSF) and brain lesion to quantitatively analyze multiple sclerosis. The method performs intensity-based tissue classification using multispectral magnetic resonance (MR) images based on a stochastic model. With the existence of white Gaussian noise and spatially invariant blurring in acquired MR images, a Karhunen-Loeve (K-L) domain Wiener filter is applied for an accurate noise reduction and resolution restoration on blurred and noisy images to minimize the partial volume effect (PVE), which is a major limiting factor for the quantitative analysis. Following that, we utilize a Markov random field Gibbs model to integrate the local spatial information into the established expectation-maximization model-fitting algorithm. Each voxel is then classified by a mixture-based maximum *a posterior* (MAP) criterion, indicating its probabilities of belonging to each class, i.e., each voxel is labeled as a mixel with different tissue percentages, leading to further minimization of the PVE. The volumes of WM, GM and CSF are extracted from the mixture-based segmentation and the corresponding brain atrophies are computed. In this study, we have investigated the accuracy and repeatability of the algorithm with inclusion of noise analysis and point spread function for image resolution enhancement. Experimental results on both phantom and healthy volunteer studies are presented.**

I. INTRODUCTION

Multiple sclerosis (MS) is a chronic, immune-mediated, demyelinating disease that affects the central nervous system [1]. Approximately 80 to 100 per 100,000 people have MS in the United States, making it the most frequent cause of disability in early to middle adulthood other than trauma [2]. Quantitative analysis of white matter (WM), gray matter (GM), cerebral spinal fluid (CSF), as well as brain lesion provides a means for establishing a parametric index for neuropsychological evaluation of MS burden [3]. In addition to the parametric index, the ability of accurate quantification in a longitudinal study using a systematic segmentation approach provides a more objective measure in determination of the effect of therapy. Magnetic resonance imaging (MRI) technique is an appealing surrogate in the MS study as MRI changes reflect the pathology of the disease [4]. With current fast MRI techniques, a set of three or more (multispectral) images can be acquired rapidly as T_1 and T_2 relaxation time and proton density weighted, as well as fluid attenuated inversion recovery (FLAIR), respectively. Because these images are spatially registered over the three-dimensional (3D) space, information extracted by means of image processing from multispectral images is obviously more valuable than that extracted from each image individually [5]. Furthermore, segmenting each voxel as a mixel with different tissue percentages is theoretically attractive and clinical desired to minimize the partial volume effect (PVE). Taking into account the PVE, we proposed a mixture-based maximum *a posterior* probability expectation-maximization (MAP-EM) segmentation framework for quantification of those tissue components inside each voxel using the multispectral MR images [6]. In this study, we evaluated the accuracy and repeatability of the algorithm with inclusion of noise analysis and point spread function (PSF) for image resolution enhancement. Experimental results on phantom and healthy volunteer studies were presented.

II. METHODS

A. MRI Acquisition Protocols

MRI sessions were performed using a 1.5 Tesla Marconi Edge whole-body scanner with a body coil as the transmitter and a birdcage head coil as the receiver. A 3D SPGR sequence was employed to acquire T_1-weighted axial images covering the whole brain with 30^o flip angle, T_E = 5 ms, T_R = 30 ms, 1.5 mm slice thickness, 24 cm field-of-view (FOV), and 256x256 matrix size. A 3D EXPRESS sequence with fat suppression was used to collect T_2-weighted axial images with the same acquisition location and parameters, except for T_E = 95 ms, T_R = 4000 ms, and ETL = 136. A FLAIR image with CSF saturation was also acquired from each patient with the same location and FOV in the sessions. The total MRI scanning time was less than 40 minutes. These three scans were performed sequentially with the subject lying in the same position in the coil. The multispectral images were registered well in the spatial domain.

This work was supported in part by the National Institutes of Health (grant #CA82402) and the National Multiple Sclerosis Society (grant #RG3042-A-2).

L. Li is with the Radiology and Electrical Engineering Departments, State University of New York, Stony Brook, NY 11794 USA (telephone: 631-444-7921, e-mail: lli@clio.rad.sunysb.edu).

H. Lu, X. Li, W. Huang and A. Tudorica are with the Radiology Department, State University of New York, Stony Brook, NY 11794 USA.

C. Christodoulou and L. Krupp are with the Neurology Department, State University of New York, Stony Brook, NY 11794 USA.

Z. Liang is with the Radiology and Computer Science Departments, State University of New York, Stony Brook, NY 11794 USA.

B. Fourier Domain Interpolation for Isotropic Voxel Size

MR images are acquired with a square pixel size but with variable thickness, resulting in a non-cubic voxel size. In our study, the slice thickness of the constructed image is 1.5 mm, which is larger than the pixel size (0.9375x0.9375 mm^2) in the X-Y plane. Therefore, a Fourier domain-based interpolation by zero padding was first applied to the data sets for constructing an equal distance for local intensity vector selection in the 3D spatial domain. This procedure was applied to all multispectral images.

C. K-L Wiener Restoration of Blurred and Noisy Images

MR image datasets include a spatially invariant blurring, which can be characterized by the PSF. Deblurring is desired to improve the quantitative analysis, but usually at the cost of noise. Therefore, noise suppression must be considered. MRI noise can be modeled as white Gaussian, embedded in the in-phase and quadrature components of the received signal [7].

For data contaminated with white Gaussian noise and spatially invariant blurring, Wiener restoration is an optimal solution based on the minimum least-square criterion for a linear system. With further consideration of the correlation among the multispectral images, we employed the Karhunen-Loeve (K-L) transform strategy for an improved noise treatment [8]. Then the K-L domain Wiener restoration has the form of

$$\hat{\breve{Y}}_i(u,v) = \frac{H_c(u,v)\breve{Y}_i(u,v)}{|H_c(u,v)| + \sigma_n^2 / \Phi_{\breve{Y}}(u,v)} \qquad (1)$$

where $\Phi_{\breve{Y}}$ is the power spectrum of K-L transformed image $\breve{y}_i$ at slice *i* and (u,v) denote the 2D Fourier transform (FT) coordinates. The system matrix H_c describes the blurring kernel in each individual slice. By taking both the spatial correlation and noise/blurring properties into account, the K-L domain Wiener filtering provides a theoretically-based restoration on blurred and noisy image.

D. Mixture-based Segmentation Algorithm

Conventionally, an image is classified by labels and each voxel is assigned a label, which reflects the main property of that voxel. As the voxel size increases, quantitative analysis error appears due to the PVE among different tissues. Therefore, we proposed an optimal solution for the PVE, i.e., a mixture-based segmentation for quantification of those tissue components inside each voxel.

We first applied a fully automated fast online vector quantization segmentation [9] to extract the intra-cranial volume (ICV) as a brain mask from the T_2-weighted image. Following that, we applied the self-adaptive vector quantization segmentation again to assign labels to those tissue types within the ICV. From these labels, we computed the initial parameter estimations for our iterative mixture-based segmentation algorithm, as described below.

The algorithm iteratively estimates the model parameters through the EM strategy and segments the voxels by MAP in an interleaved manner between labels and mixtures, converging to a solution where the model parameters and voxel labels are stabilized within a specified criterion.

Let $Y_i = \{Y_{il}\}_{l=1}^{L}$ be the density vector of *L*-channel MR images at location *i* over the 3D image array of *I* voxels. Assume that the images consist of *K* classes (or tissue types) and each class *k* is characterized by a Gaussian parameter vector $\theta_k(\mu_k, \nu_k)$, i.e., the mean and variance. Let $p_k(Y_i \mid \theta_k)$ be the probability distribution of those voxel densities that are associated with class *k*. We further assume that the multispectral MR images are statistically independent. The likelihood for each voxel Y_i, falling into *K* distinct classes, is then described by a finite multivariate functional as

$$g(Y_i \mid X, \Theta) = \sum_{k=1}^{K} p(k \mid X_{N_i})\, p_k(Y_i \mid \theta_k) \qquad (2)$$

where $p(k \mid X_{N_i})$ is the locally-dependent probability [10] of voxel label $X_i = k$. By the EM strategy, we have, at each iteration *n*,

$$\mu_{kl}^{(n+1)} = \frac{\sum_i Z_{ik}^{(n)} Y_{il}}{\sum_i Z_{ik}^{(n)}} \qquad (3)$$

$$\nu_{kl}^{(n+1)} = \frac{\sum_i Z_{ik}^{(n)} (Y_{il} - \mu_{kl}^{(n+1)})^2}{\sum_i Z_{ik}^{(n)}} \qquad (4)$$

where Z_{ik} is the conditional probability that voxel Y_i belongs to class *k*, which represents the tissue percentages within that voxel, i.e., it is the task of our mixture-based quantitative analysis,

$$Z_{ik}^{(n)} = \frac{p^{(n)}(k \mid X_{N_i}) \prod_{l=1}^{L} p(Y_{il} \mid \theta_k^{(n)})}{\sum_j [p^{(n)}(j \mid X_{Ni}) \prod_{l=1}^{L} p(Y_{il} \mid \theta_j^{(n)})]} \qquad (5)$$

If ignore the local dependence of the voxel labels, $p(k \mid X_{N_i})$ becomes the weights of the finite multivariate function [5], [11]. Then the estimation of Z_{ik} becomes the maximum likelihood (ML) solution.

There are several ways to include the local dependence by specifying a function of $p(k \mid X_{N_i})$. A Markov random field (MRF) model is commonly used to reflect the neighborhood information [12]-[13]. An energy function *U(X)* is constructed to specify the degree of penalty imposed on the neighbors, and can be defined, in three dimensions, as

$$U(\mathrm{X}) = \sum_{i=1}^{N} \left\{ \sum_{r \in c_i^1} [1 - \delta(X_i - X_r)] + \sum_{s \in c_i^2} [1 - \delta(X_i - X_s)] / \sqrt{2} \right\} \qquad (6)$$

where $\delta(0)=1, \delta(\neq 0)=0$. The index r runs over the 6 first-order neighbors and s runs over the 12 second-order neighbors. The assignment of labels over the voxel array is performed by the MAP criterion.

E. Extraction of the Central Cerebral Atrophy (CCA)

Besides of quantitative analysis on brain tissues of WM and GM, one neuroimaging measure of cerebral injury was also used in this study -- cerebral atrophy. Cerebral atrophy has been shown to correlate with a variety of cognitive functions, including those functions most commonly disturbed. It may represent the final cumulative effect of different types of MS induced lesions [14]. In the segmented mixture image or mixels, by choosing a central CSF seed, the central CSF volume was delineated by both morphology and region growing technologies. In some datasets where the central CSF and peripheral CSF are connected, these connections could induce an overestimation of the central CSF. Then a semi-automated tool was developed to assist the physicians to cut appropriately the connections between the central and peripheral CSFs. In clinic studies, lesion appears in patient datasets with the same density feature as CSF in both T_1-weighted and T_2-weighted images. Therefore, the Flair image was acquired for detecting lesions. In the Flair image, lesion density appears brighter than those of other tissues. The segmented lesion was subtracted from the segmented CSF result to minimize the error for atrophy measure. In this study, we proposed the following definitions for quantitative validation.

(1) ICV = WM + GM + Total CSF;
(2) WM fraction (WMF) = WM / ICV;
(3) GM fraction (GMF) = GM / ICV;
(4) Total atrophy (TA) = Total CSF / ICV;
(5) CCA = Central CSF / ICV;
(6) Peripheral atrophy (PA) = Peripheral CSF / ICV.

F. Evaluation Methods

We evaluated the accuracy and repeatability of the algorithm through both phantom and volunteer studies. The accuracy is assessed using both the mathematical brain phantom images downloaded from the Connell Brain Imaging Center of McGill University and a physical brain phantom, where the gold standards are known. The repeatability is assessed by performing our MAP-EM segmentation scheme on a series of images from both the phantoms and healthy volunteers, where the gold standards are the assumption that the phantoms and the brains of the healthy volunteers remain the same in a time period. Two volunteers (age 33, one male and one female) were recruited in this evaluation study. Four scans were performed within one year for each volunteer. The repeatability was assessed by calculating the following formula:

Repeatability = standard deviation (stdev) / mean

III. EXPERIMENTAL RESULTS

A. Visual Evaluation

Fig. 1(a) shows a slice of the T_2-weighted image of a volunteer, which has a good contrast between CSF and brain tissues. Fig. 1(b) shows the extracted brain mask corresponding to the ICV. Fig. 2 shows the original T_1- and T_2-weighted images and the corresponding MAP-EM segmented labels indicating WM, GM and CSF (quantitative measures on mixtures is reported below). Fig. 3 shows 3D surface rendering of the extracted central CSF from one subject as viewed from two different angles.

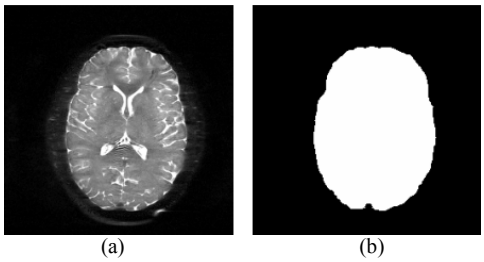


Fig. 1. (a) A slice of the T_2-weighted image. (b) The extracted brain mask corresponding to the ICV.

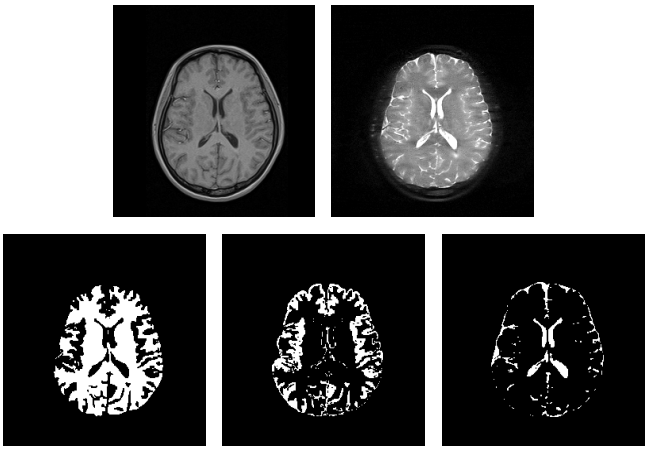


Fig. 2. The T_1-, T_2-weighted MR images (top) and the extracted WM, GM, and CSF (bottom).

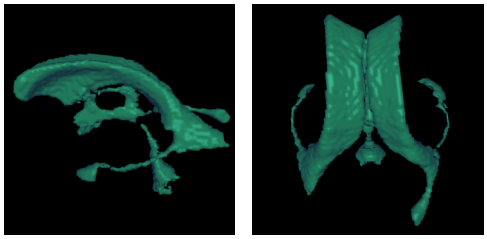


Fig. 3. The 3D surface rendering of the extracted central CSF from one subject as viewed from two different angles.

The total computing time of this algorithm for the results was approximately eighteen minutes for a set of multispectral images with 256x256x160 size on a PC/800Mhz Pentium III.

B. Quantitative Evaluation

B.1. Mathematical phantom studies

The presented segmentation scheme is fully automated except for the extraction of central CSF. Therefore, quantitative estimation of WM, GM and total CSF is 100% reproducible.

For the accuracy measure by the phantom images or the anatomical models from the Connell Brain Imaging Center of McGill University, the total volumes of CSF, GM, and WM are given as 371.9, 902.9, and 674.8 (cm^3), respectively. Therefore, the ICV volume is 1949.6 (cm^3). We used the Brainweb MRI simulator to generate a series of six scans/images based on the mathematical brain phantom with 1 mm cubic voxel size at different noise levels. The non-uniformity effect of image density in this study was set to 0%. Fig. 4 demonstrates the accuracy of the mixture-based quantitative estimation of brain tissues at different noise levels, where the region to cover a mixture component is given by the corresponding labeled area. Fig. 5 shows the repeatability test on the phantom studies at different noise levels. The accuracy and repeatability are satisfactory.

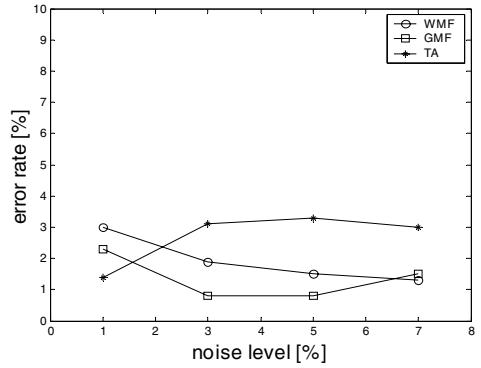


Fig. 4. Error rate of the mixture-based quantitative estimation of brain tissues at different noise levels.

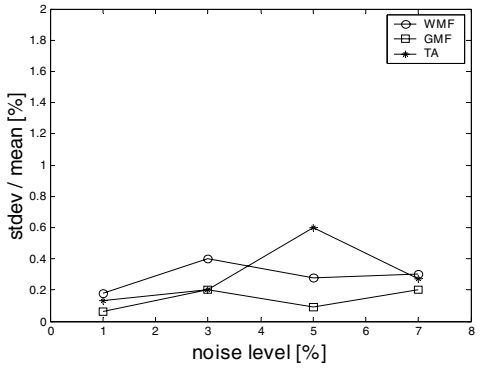
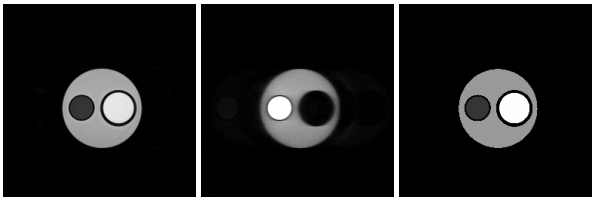


Fig. 5. Repeatability test on the phantom images at different noise levels.

B.2. Physical phantom studies

A brain phantom was constructed by plastic materials with three compartments of WM, GM, and CSF, filled with different concentrations of Gadolinium (Gd) contrast solution. From density brightest to darkest in the T_1-weighted image of Fig. 6(a), it contains 6 mM, 1 mM, 0.1 mM Gd solutions, respectively. From size big to small, the diameters of the cylinders are 97.5 mm, 37.5 mm (with wall thickness 3 mm), and 28.5 mm (with wall thickness 1.8 mm), respectively. The physical phantom was scanned by our 1.5 Tesla whole body MRI scanner using the protocol as described before. Five scans were performed in five weeks, i.e., one scan per week. Fig. 6 shows one slice of a volume scan from the physical phantom datasets (T_1-, T_2-weighted images) and the corresponding segmented image. Fig. 7 demonstrates the accuracy on the physical phantom studies. Fig. 8 shows the repeatability on the phantom series scans. The error rate is less than 3.3% and the repeatability is below 0.1%.



(a) (b) (c)

Fig. 6. One of the image planes from the physical phantom dataset. (a) T_1-weighted image. (b) T_2-weighted image. (c) Segmented image.

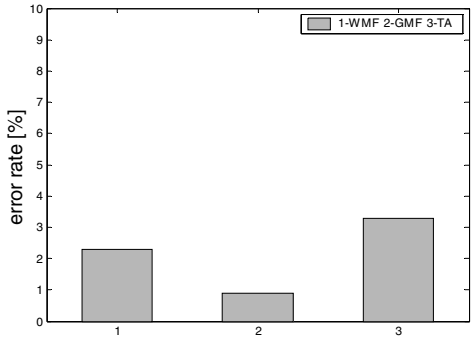


Fig. 7. Quantitative error rate of estimating different volumes on physical phantom studies.

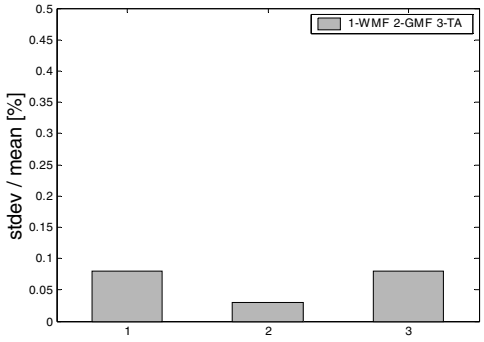


Fig. 8. Repeatability test on physical phantom studies.

B.3. Volunteer studies

For healthy volunteers, only repeatability test was assessed. Fig. 9 shows the repeatability test on the two volunteer studies each having a series of four scans. The CCA measurements have achieved a high repeatability with less than 1.9%, indicating the feasibility of our segmentation scheme for neurological diagnosis and follow up evaluation on MS. The atrophies on WM, GM, total CSF and peripheral CSF have an average variation of around 2.5%.

It should be noted that for the phantom studies, the gold standards are known, so both the accuracy and repeatability tests do not have the problem of image registration. However, for the volunteer studies, the registration of the series images can be a major problem for quantitative measures. In our quantitative analysis framework, we assume that the skull of an adult remains the same in a time period. Based on this assumption, the ICVs of each subject in the series scans become the relative references for those scans. Therefore, the ratios on each relative reference, respectively, have the spatial invariant property, eliminating the registration problem.

In Fig. 9, the variation of the ICV causes a noticeable change for the brain tissues, but the ratios minimize the variation.

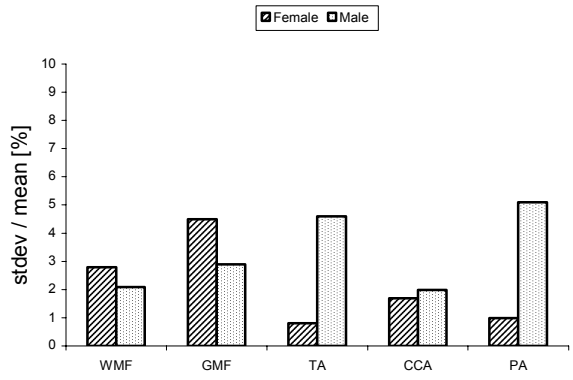


Fig. 9. Repeatability test of volumetric analysis on volunteer studies.

IV. DISCUSSION AND CONCLUSIONS

We validated our mixture-based MAP-EM segmentation algorithm with inclusion of noise reduction as well as PSF analysis for resolution enhancement. The initial estimate for the iterative MAP-EM update is given by the self-adaptive quantization scheme [9]. The iterative convergence is guaranteed by the MRF model. Therefore, the algorithm is fully automated with 100% reproducibility on quantitative estimation of WM, GM, CSF as well as brain atrophy. Experiments on accuracy and repeatability demonstrate the feasibility of this algorithm for quantitative analysis of brain MR images. Further research on improving the analysis with inclusion of inter- and intra-slice in-homogeneity correction is under progress [15]. Preliminary studies on patient datasets showed a strong correlation with clinical findings [16].

ACKNOWLEDGMENT

The authors wish to thank radiologists, Robert Peyster, Clemente Roque, and Patricia Roche for their comments on this project.

REFERENCES

[1] S. L. Hauser, "Multiple sclerosis and other demyelinating diseases," 1, New York: McGraw-Hill, pp. 2287-2295, 1994.

[2] J. F. Kurtzke, M. Wallin, "Multiple Sclerosis: diagnosis, medical management, and rehabilitation," *Epidemiology*, pp. 49-71, In J.S. Burks, K. P. Johnson (Eds), New York, 2000.

[3] R. A. Rudick, E. Fisher, J. C. Lee, J. Simon, L. Jacobs, and the MS Collaborative Research Group, "Use of brain parenchymal fraction to measure whole brain atrophy in relapsing remitting MS," *Neurology,* vol. 53, pp. 1698-1704, 1999.

[4] D. Chen, W. Huang, C. Christodoulou, L. Li, H. Qian, L. Krupp, and Z. Liang, "A new method for quantitative analysis of multiple sclerosis using MR images," *SPIE Medical Imaging*, vol. 4321, pp 375-380, 2001.

[5] Z. Liang, J.R. MacFall, D. P. Harrington, "Parameter estimation and tissue segmentation from multispectral MR images," *IEEE Trans. Med. Imag.*, vol. 13, pp. 441-449, 1994.

[6] L. Li, X. Li, W. Huang, C. Christodoulou, D. Chen, A. Tudorica, P. Djuric, L. Krupp, and Z. Liang, "A novel mixture-based segmentation algorithm for quantitative analysis of multiple sclerosis using multispectral MR images," *Proc. Intl. Soc. Mag. Reson. Med.*, 1: 345*,* 2002.

[7] D. L. Parker and G. T. Gullberg, "Signal-to-noise efficiency in magnetic resonance imaging," *Medical Physics*, vol. 17, pp. 250-257, 1990.

[8] H. Lu, X. Li, I. Hsiao, and Z. Liang "Analytical noise treatment for low-dose CT projection data by penalized weighted least-square smoothing in the K-L domain," *SPIE Medical Imaging,* vol. 4682, pp. 146-152, 2002.

[9] L. Li, D. Chen, H. Lu, Z. Liang, "Segmentation of MR brain images: a self-adaptive online vector quantization approach," *SPIE Medical Imaging*, vol. 4322, pp. 1431-1438, 2001.

[10] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Trans. Medical Imaging*, vol. 20, pp. 45-57, 2001.

[11] Z. Liang, R. J. Jaszczak, and R. E. Coleman, "Parameter estimation of finite mixtures using the EM algorithm and information criteria with application to medical image processing," *IEEE Trans. Nuclear Science*, vol. 39, pp. 1126-1133, 1992.

[12] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, R. Kikinis, and H. Muller-Gartner, "Markov random field segmentation of brain MR images," *IEEE Trans. Medical Imaging*, vol. 16, pp. 876-886, 1997.

[13] R. Leahy, T. Hebert, and R. Lee, "Applications of Markov random fields in medical imaging," *Information Processing in Medical Imaging*, pp. 1-14, 1991.

[14] W. J. Jagust, "Brain atrophy as a surrogate marker in MS: Faster, simpler, better?" *Neurology*, vol. 54, pp 782-783, 2000.

[15] D. Chen, L. Li, D. Yoon, J. H. Lee, and Z. Liang, "A renormalization method for inhomogeneity correction of MR images," *SPIE Medical Imaging,* vol. 4322, pp. 939-942, 2001.

[16] X. Li, L. Li, W. Huang, C. Christodoulou, C. Roque, P. Roche, R. Peyster, L. Krupp, and Z. Liang, "A Volumetric Analysis on Central Cerebral Atrophy for Multiple Sclerosis Studies with Cognitive Impairment," to be presented at RSNA, 2002.