A new method for quantitative analysis of multiple scelerosis using MR images

Dongqing Chen[*a], Wei. Huang[a], C. Christodoulou[b], Lihong Li[a], Huayuan Qian[a], Lauren Krupp[b], and Z. Liang[a]

Depts. of Radiology[a] and Neurology[b], State University of New York, Stony Brook, NY

ABSTRACT

A method for quantitative analysis of multiple sclerosis (MS) was presented. An automatic self-adaptive image segmentation algorithm was first employed to classify voxels in multi-spectral magnetic resonance (MR) images. The segmentation results from multi-spectral MR images were then combined to obtain reliable results. The volumes of brain tissues and cerebral spinal fluid (CSF) were finally extracted. Since it is fully automated, the results of the segmentation algorithm are completely reproducible. The repeatability of the presented method was evaluated on volunteer data sets. The variation is less than 0.2% for the intra-cranial volume, the whole brain volume, the central CSF, the white matter (WM) and the gray matter (GM). The variation of 3% for the entire CSF is mainly due to the peripheral CSF part, which has more partial volume effect and is less important than the central one. Methods for minimizing this variation are under investigation. These measurements demonstrate the potential for study on whole brain atrophy and cerebral atrophy. Feasibility studies on 14 MS patients were performed. The results are promising.

Keywords: Adaptive Segmentation, Multi-spectral MRI, Quantative Analysis, Multipl Sclerosis, Brain Atrophy

1. INTRODUCTION

Multiple sclerosis (MS) is the most common inflammatory disease of the central neural system (CNS) and is the most frequent cause of nontraumatic neurologic disability in young and middle-aged adults [1]. Recently, surrogate markers that can be used as an endpoint of outcome measure in treatment trials are a major focus of experimental therapeutics and Food & Drug Adminstration (FDA) policy [2]. Magnetic resonance imaging (MRI) is an appealing surrogate in MS because there is increasing consensus that MRI changes reflect the pathology of the disease. Although there were still limitations for using MRI, a recent consensus of MS investigators called for an immediate, retrospective analysis of measures of atrophy on existing data sets and trials in progress to determine whether experimental therapeutic agents favorably influence this MRI measurement.

There were two approaches to measure brain atrophy in MRI. One approach uses registration and subtraction on a serial of MR images to measure the volumetric change [3,4]. Another segments the MR images and directly measures the segmented results [5,6]. The first approach highly relies on the registration method. Although subvoxel accuracy was reported with registration [4], the tissue volume, especially the lesion load, could not be automatically measured with this approach. The accuracy and consistency of image segmentation were the foundation of the second approach. Usually, good image quality and high spatial resolution are required for both approaches to achieve accurate measurements.

Fish et al [5] developed a knowledge-based three-dimensional (3D) method for segmenting brain MR images. It was utilized to measure whole brain atrophy in relapsing-remitting MS [7]. The evaluation of that method on both brain phantom images and volunteers' MR images was presented. However, what was not discussed in that method was that how to measure the lesion load and the central CSF. No repeatability evaluation was reported also. In this study, we prosented a new method of the second approach for quantitative analysis of MS and try to obtain the complete information about the entire brain tissues. The method was evaluated by volunteer data on its accuracy and repeatability and was applied to 14 MS patients for feasibility studies. The results showed that this method was a promising tool for quantitative analysis of MS.

2. METHODS

2.1) Overview. For quantitative analysis of MS, we tried to measure both brain atrophy and lesion load with MR images. Usually, T_2-weighted image is utilized to extract CSF in MS study [2], while the T_1-wighted image provides a better tissue

[*] For information, contact dchen@clio.rad.sunysb.edu, phone 1 631 444 9719; fax 1 631 444 6450; Department of Radiology, School of Medicine, State University of New York, Stony Brook, NY 11794.

contrast for brain tissues (WM and GM). However, the lesion might appear as the same intensity feature as GM in both T_1- and T_2-weighted images. In clinics, a FLAIR image is often used to detect the lesions since the lesion area is brighter than other tissues. Considering about these facts, multi-spectral MR images were utilized in this study. In order to achieve high consistency, we utilized several clinically available pulse sequences to acquire T_1-, T_2-weighted, and FLAIR images.

All acquired images were segmented with a robust, fully automatic, and self-adaptive segmentation algorithm [6,8]. First, the T_2-weighted image was segmented into 3 classes. The intra-cranial brain mask was delineated from this 3-class segmentation result. Then, the voxels inside this brain mask were classified into 2 classes. The total CSF was that class with higher average intensity. The central CSF volume was extracted from the total CSF volume by applying both mathematical morphology operations and region growing technologies. The volume of brain tissues, including both GM and WM, were obtained by subtracting the total CSF volume from the brain mask volume. The brain atrophy was then calculated by the extracted volumes. To obtain more information on the brain tissues, the brain tissues were segmented from the T_1-weighted image within the brain mask into two classes. Similarly the FlAIR image within the mask was also segmented into two classes. From those results, the volumes of GM, WM, and lesion load were obtained. Since some of the lesions might be classified as GM as errors, we subtracted the segmentation result of the FLAIR image from that of the T_1-weighted image to minimize the error.

2.2) MR scanning protocols. MRI scans were performed using a 1.5 T Marconi Edge whole-body scanner with the 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, TE = 5 ms, TR = 50 ms, 3 or 1.5 mm (3 mm for patients and 1.5 mm for volunteers) slice thickness, 24 cm FOV (field-of-view), 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 TE = 95 ms, TR = 4000 ms, and ETL = 136. 3D FLAIR images with CSF saturation were also acquired from each patient with the same slice location, slice number, slice thickness, and FOV. Usually, 48 to 52 slices (3 mm thick) were acquired for a single volumetric data depending on the brain size of the subject. The total MRI scanning time was around a half hour. Considering that the three scans were performed one-by-one in a sequence with the subject lying in the same position in the coil, we assume that these multi-spectral images are registered well in spatial location.

2.3) Interpolation. The slice thickness was 3 or 1.5 mm and larger than 0.975 mm of the voxel size in the X-Y plan. We added two more slices between any two consecutive original slices by interpolation. This might suppress partial volume effect and noise to some degrees. In this study, a linear interpolation was utilized. The interpolation procedure was applied to all T_1-, T_2-wieghted, and FLAIR images.

2.4) Segmentation algorithm. The segmentation algorithm classified voxels based on their intensity vector. The intensity vector for each voxel consists of intensities of itself and its neighbors. Principal components analysis was applied to the intensity vectors first to get feature vectors with lower dimension. A modified on-line vector quantization algorithm was then employed to classify the feature vectors [6,8]. Given a maximum class number to be created, the algorithm proceeds in a fully automatic manner. Since the algorithm is fully automatic, it is completely reproducible. For multi-spectral images, the same algorithm was utilized.

For the T_2-weighted image, the algorithm was applied twice. At the first time, it was applied with the maximum number equal to 3 to the entire image. The resulted 3 classes were the intra-cranial tissues and the scalp, the region outside of the FOV, and other background outside of the scalp. The largest connected region in the center of the image should be the intra-cranial brain mask. It was extracted by region growing techniques. In the brain mask, it might include some part of the eyes or other non-brain tissues. To separate the non-brain tissues, we applied mathematical morphology erosion three times. For each time, the eroded layer was labeled as a specific index, and then a region growing with the same seed was applied again to create an eroded brain mask. In this mask, the non-brain stuff should not appear. Finally, we dilated the eroded brain mask 3 times back to where we recovered the layer that was labeled previously, achieving the brain mask. At the second time of applying the algorithm to the T_2-weighted image, the maximum number of classes was equal to 2 within the brain mask. The created class with a higher average intensity was the region of CSF. Subtracting this CSF volume from the brain mask, the volume of brain tissues including both GM and WM was obtained. For T_1-weighted and FLAIR images, the algorithm was applied with the maximum class number of 2 within the brain tissue volume. Some lesion area might be classified as gray matter in error. By subtracting the FLAIR segmentation from that of the T_2-weighted image, the error region were corrected. Then the total lesion load was measured from the subtraction results.

2.5) Extract central CSF. The total CSF consists of central CSF and peripheral CSF. It is a single connected volume. Some criteria must be set to automatically separate the central CSF from the peripheral one. The central CSF is defined as the fluid within the ventricles. The ventricular system connects to the peripheral CSF through several narrow points. This makes it possible to separate the central CSF from the peripheral one by eroding the total CSF volume in a light manner, then growing out the central CSF with several seed points and recovering the volume with dilation operation. However, the connection between the fourth ventricle and the main body of lateral ventricles is narrow, results in missing of the fourth ventricle. In this study, we ignored the fourth ventricle and counted it as peripheral CSF.

2.6) Definition of atrophy. The followings are the definitions that were used in this study.

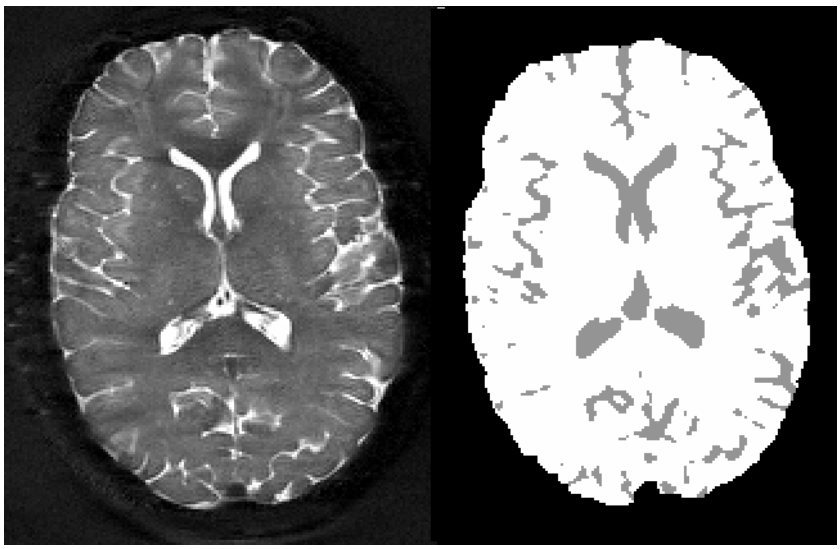
1) Intra-cranial volume (IV) = GM + WM + Total CSF;
2) Central atrophy (CA-I) = Central CSF / (GM + WM);
3) Central atrophy (CA-II) = Central CSF / Intra-cranial volume;
4) Central atrophy (CA-III) = Central CSF / (GM + WM + Peripheral CSF);
5) Total atrophy (TA) = Total CSF / Intra-cranial volume;
6) Brain parenchymal fraction (BPF) = (GM + WM) / Intra-cranial volume;

2.7) Evaluation method. The repeatability and accuracy was evaluated with the volunteer data. For each volunteer, two scans were acquired in different days (6 days apart). The atrophy and volume measurements were calculated for data from both days. The relative errors between those results from different days were calculated with the following formula.

Relative error = | (measurement 1) – (measurement 2) | / (measurement 1) * 100%.

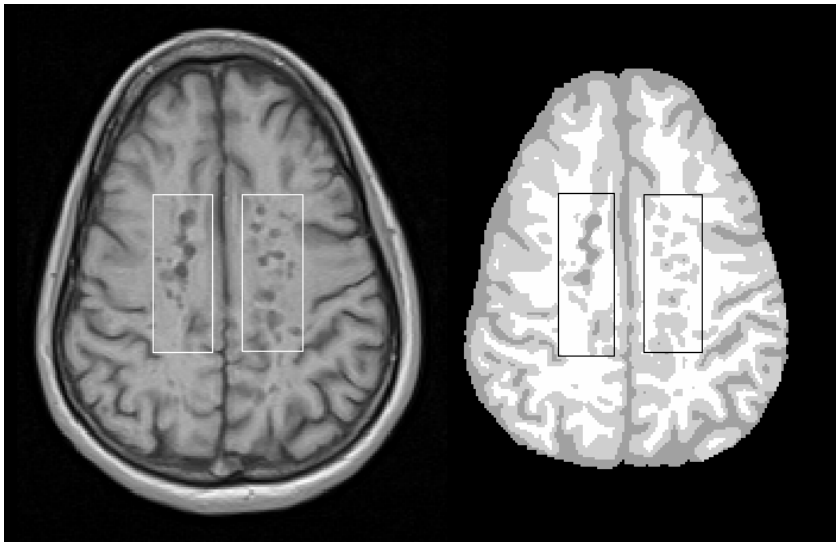
3. RESULTS AND DISCUSSIONS

Two healthy volunteers (age 33, one male and one female) and fourteen MS patients (age range 20-55 yrs, mean ± SD = 46 ± 9 yrs) were recruited in this study. The patients were clinically described as relapsing remitting (n = 9) and secondary progressive (n = 5) MS patients. For all the patients, 3 mm slice-thickness images were acquired. For the volunteers, both 3 and 1.5 mm images were acquired. Figure 1(a) shows a slice of the original T_2-weighted image, which shows good contrast between CSF and brain tissues. Figure 1(b) depicts the extracted brain mask and the segmented CSF regions in the same slice. The segmented results are promising by visual inspection. Figure 2(a) shows a slice of the original T_1-weighted image, which shows good contrast between brain (WM and GM) tissues. Figure 2(b) depicts the segmented regions of GM, WM, and lesions. The segmented results are again seen to be promising by visual inspection. Figures 3(a) and 3(b) display the central CSF from two different views in a three dimensions. Table 1 lists the measurements from 1.5 mm slice thickness volunteers' data sets. The relative errors between the two scans of six-day apart are also listed to show the repeatability of our approach. Table 2 lists the measurements from patients' data sets.



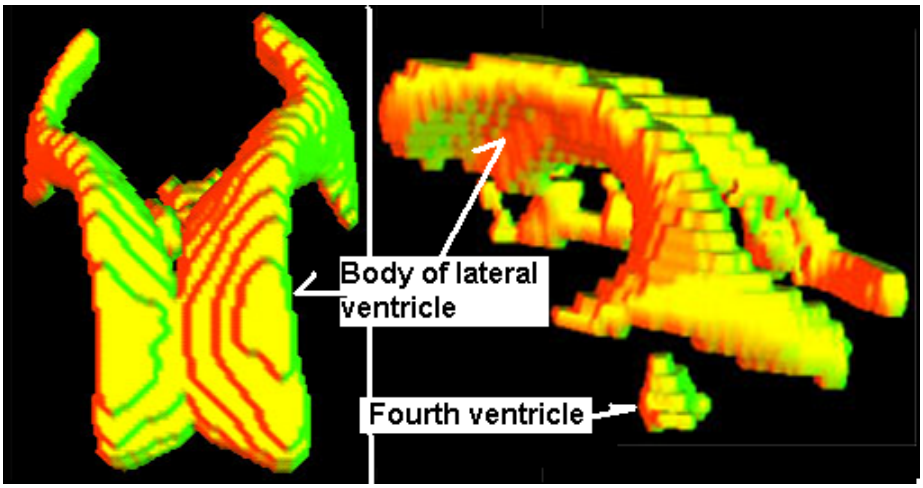
a b

Figure 1 (a) shows a slice of original T_2-weighted image and (b) depicts the extracted brain mask and the CSF regions of the same slice.



a b

Figure 2 (a) shows a slice of original T_1-weighted image and (b) depicts the segmented regions of GM, WM, and lesions.



a b

Figure 3 (a) and (b) display the central CSF from two different views in three dimensions.

From Table 1, we see that the volume and atrophy measurements have high repeatability except for the total CSF. The reason for larger error in total CSF measurement may be due to the inter-slice intensity inhomogeneity and the different slicing locations between the two scans of six-day apart. Usually, the slices near the top of the head create more errors in the CSF regions. The average tissue intensity in those slices was lower than those of the inter-slices. The solution to fix this problem is to apply an inhomogeneity correction prior to the segmentation. This is an on-going project of our team [9,10]. From the clinical point of view, the volume of central CSF is more significant than the volume of total CSF. The central CSF is completely enclosed within the middle slices. Hence, even without considering the correction, our current method can still provide very useful results for quantitative analysis of MS. With the correction, the accuracy should be improved. The inhomogeneity artifacts appeared more worse in the FLAIR images. If we want better quality in FLAIR images, we have to increase the scan time. This results in more patient motion and less patient load. A possible solution to improve the accuracy in lesion measurement is to detect the lesion location by the FLAIR image and to measure the lesion load using the segmentation result from the T_1-weighted image instead of subtracting the two spectral images.

The partial volume effect was another factor that affects the accuracy of measurements, especially for lesion load. In Figure 2, some lesions in the rectangle area were error-labeled as gray matter. To some extent, this error could be corrected by integrating the segmentation result of the FLAIR image. By the integration, we could know the locations of the lesion

regions. Utilizing the morphology and region growing techniques mentioned previously, these regions could be separated from the GM regions. Then, all the voxels connected to a lesion region could be counted as a lesion if it was error-labeled as GM. However, the partial volume effect might be a variable in scans acquired in different days. That may affect the repeatability of the measurements. A thinner slice thickness and interpolation are needed. We are currently continuing the investigation on this problem. For example, by zero padding or extrapolation to fill in the high frequency space in the Fourier domain, we can interpolate images in any size without losing the information contained in the images [11].

Table 1. Evaluation result of volunteers' data sets (1.5 mm slice thickness, volume in mm^3)

	Sex	Tot. CSF	Ctr. CSF	GM&WM	IV	CA-I	CA-II	CA-III	TA	BFP
V-1(day 1)	F	175.010	8.730	1199.129	1374.14	0.007281	0.006354	0.006394	0.127	0.873
V-1(day 2)	F	181.422	8.713	1201.151	1382.573	0.007254	0.006302	0.006342	0.131	0.869
Rel. error V1		3.5%	0.197%	0.16%	0.6%					
V-2 (day 1)	M	136.558	5.388	1140.426	1276.984	0.004725	0.004219	0.004237	0.107	0.893
V-2 (day 2)	M	134.631	5.352	1139.927	1274.558	0.004695	0.004199	0.004217	0.106	0.894
Rel. error V2		1.43%	0.067%	0.044%	0.19%					

Table 2. Results for patients' data sets (3 mm slice thickness, volume in mm^3)

Subject	Age/Sex	IV	Total CSF	GM	WM	Ctr CSF
1	58F	1205.231	177.735	391.681	635.813	26.636
2	52F	1174.676	123.374	400.293	651.008	9.210
3	52F	1488.390	230.391	611.903	646.096	8.448
4	46F	1348.900	73.859	422.104	852.936	9.054
5	55F	1422.106	203.757	454.472	852.936	11.416
6	55F	1268.499	96.473	410.566	761.464	7.699
7	54M	1513.031	142.192	599.357	771.480	43.089
8	44F	1311.936	159.858	389.200	762.876	19.453
9	40F	1461.662	194.974	457.626	809.061	46.532
10	39F	1563.285	241.671	507.497	814.121	35.226
11	45F	1301.589	114.673	429.170	757.745	12.466
12	46F	1068.229	172.235	472.468	423.525	18.559
13	20F	1160.156	86.803	406.165	667.187	1.964
14	46F	1345.633	150.292	458.016	737.324	8.139

CONCLUSION

The presented method for quantitative analysis of MS was demonstrated with a very high repeatability by volunteer studies, see Table 1. Experimental results on 14 patients were very satisfactory in correlation to clinical findings, see Table 2. A fully automated segmentation algorithm and the followed extraction strategy are necessary to guarantee the reproducibility, as ours. To achieve over 97% repeatability for adequate quantifying the normal variation and therefore for reliable MS studies, the image slice thickness must not be greater than 1.5 mm to reduce the partial volume effect within the normal variation. Suitable interpolation is needed if the slice thickness is different from the in-plan pixel size. Since the image quality decreases with thinner slices, interpolation shall play a role to balance the image resolution and noise. Fourier transform-based interpolation by zero padding or extrapolation to fill in the high frequency region is a choice [11]. Correction for the inter- and intra-slice inhomogeneity is also needed to guarantee the desired repeatability [9,10]. It has been recognized that simultaneous segmentation of multi-spectral MR images can provide improved classification [12]. Extension of this computationally efficient segmentation algorithm for simultaneously classifying the multi-spectral MR images is under progress. This presented quantitative analysis method can be a promising means for MRI-based neurological diagnosis and treatment planning and evaluation in a time period.

ACKNOWLEDGEMENTS

This study was supported by the NIH grant #CA82402 of the National Cancer Institute, and partially by MD OnLine Inc and Viatronix Inc.

REFERENCES

1. S. L. Hauser, *Multiple Sclerosis and Other Demyelinating Diseases*, **1**, New York: McGraw-Hill, p. 2287-2295, 1994

2. W. J. Jagust and J. H. Noseworthy, "Brain atrophy as a surrogate marker in MS: faster, simpler, better?" *Neurology*, **54**(3), p. 782-783, 2000
3. S. J. Nelson, B. S. Nalbandian, E. Proctor, and D. B. Vigneron, "Registration of images from sequential MR studies of the brain," *Journal of Magnetic Resonance Imaging*, **4**(3), p. 877-883, 1994
4. J. V. Hajnal, N. Saeed, E. J. Soar, A. Oatrifge, I. R. Young, and G. M. Bydder, "A registration and interpolation procedure for subvoxel matching of serially acquired MR images," *Journal of Computer Assisted Tomography*, **19**(2), p. 289-296, 1995
5. E. Fisher, R. M. Cothren, J. A. Tkach, T. J. Masaryk, and J. F. Cornhill, "Knowledge-based 3D segmentation of MR images for quantative MS lesion tracking," *SPIE Medical Imaging*, **3034**, p. 599-610, 1997
6. D. Chen, L. Li, and Z. Liang, "A self-adaptive vector quantization algorithm for MR image segmentation," *Proc. ISMRM 7th Scientific Meeting and Exhibition*, Philadelphia, May 1999, p. 122
7. R. Rudick, E. Fisher, J. Lee, J. Simon, and L. Jacobs, "Use of the brain parenchymal fraction to measure whole brain atrophy in relapsing-remitting MS," *Neurology*, **53**(11), p. 1698-1704, 1999
8. D. Chen, Z. Liang, M. Wax, L. Li, B. Li, and A. Kaufman, "A novel approach to extract colon lumen from CT images for virtual colonoscopy," *IEEE Transactions on Medical Imaging*, **10**(12), to appear, 2000
9. Z. Liang, D. Wang, J. Ye, H. Li, C. Roque, and D. Harrington, "Correction of RF inhomogeneity for automatic segmentation from multi-spectral MR images," *Proc. ISMRM 5th Scientific Meeting and Exhibition*, New York, May 1997, p. 2042
10. D. Chen, L. Li and Z. Liang (2000), "A renormalization method for inhomogeneity correction of MRI data," *Proc. ISMRM 8th Scientific Meeting and Exhibition*, Denver, May 2000, p. 1542
11. Z. Liang, J. Ye, J. Cheng, and D. Harrington, "The inversion of exponential Radon transform for quantitative brain SPECT," *Physics in Medicine and Biology*, **41**(6), p. 1227-1232, 1996
12. Z. Liang, J. MacFall, and D. Harrington, "Parameter estimation and tissue segmentation from multi-spectral MR images," *IEEE Transactions on Medical Imaging*, **13**(3), p. 441-449, 1994