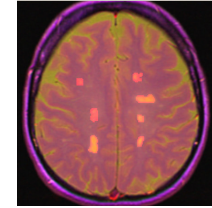


A Method for Fully Automated Measurement of Neurological Structures in MRI

Edward Ashton, Jonathan Riek, Larry Molinelli, Michel Berg*, Kevin Parker**
 VirtualScopics, LLC, Pittsford, NY; *Dept. of Neurology, University of Rochester Medical Center; **Dept. of Electrical Engineering, University of Rochester



BACKGROUND: Accurate identification and measurement of various neurological structures is a vital tool both for surgical planning and for evaluation of disease progression and patient response to therapy for numerous diseases. Measurement of white matter lesion burden is a primary method for evaluation of multiple sclerosis (MS) progression (1). Hippocampal volume is an important endpoint for diagnosing and monitoring both intractable temporal lobe epilepsy and Alzheimer's disease (2,3). Current standard methods for obtaining these data points are largely manual and subjective, and are therefore both error-prone and subject to inter- and intra-operator variability. Ashton *et al.* and Hsu *et al.* have presented semi-automated methods for the identification and measurement of the hippocampus (4,5). In addition, several groups have recently evaluated the speed, precision and accuracy of automated and semi-automated novel MS lesion identification and quantification techniques (6,7,8,9,10). However, there is as yet no generally accepted solution to either of these problems.

In this work, a method for automating the detection and measurement of various neurological structures is described and evaluated. The test cases used involved the measurement of multiple sclerosis lesions in axial MRI series and the measurement of hippocampus volumes in coronal MRI series, but the technique is a general one, and should be useful for measuring most normal and abnormal neurological structures. The results from both experiments were compared to computer-aided manual structure identification and measurement.

MEASUREMENT TECHNIQUES: One manual and one automated measurement technique were evaluated in this study. All measurements were carried out using software developed by VirtualScopics, a biomedical image analysis firm based in Rochester, NY. Manual measurements were carried out by expert observers, who were required to trace the boundaries of each structure of interest in each experimental data set using a computer mouse. Once the tracing was completed, volume measurements for each structure were calculated automatically.

Automated measurements were made using a trained statistical classifier. Training was carried out using a software package which took as its inputs a baseline data set with accompanying classification map, as well as an arbitrary number of training data sets, also with accompanying classification maps. Each training data set was registered to the baseline data set using a warp-based automated registration algorithm. Means and covariance matrices were then calculated for each class represented in the classification maps. In addition, the prior probability of each class occurring at each voxel in the training data was calculated. Once all training data had been processed, the results were written out to intermediate files. The operation of this algorithm is shown in Figure 1.

The classification software package took as its inputs the data set under consideration, the baseline data set used during training, and the outputs of the training process. At initiation the baseline data set was registered to the data set under consideration, and the resulting warp field was then used to register the prior probabilities matrix to the data set under consideration. The data, prior probabilities matrix, means, and covariance matrices were then passed to an iterative maximum likelihood classification algorithm. The convergence criterion was global maximization of the maximum likelihood discriminant function, given by:

$$g_i(x) = \ln(p(\omega_i)) - \frac{1}{2} \ln |R_i| - \frac{1}{2} (x - m_i)^T R_i^{-1} (x - m_i) \quad (1)$$

where R is the class covariance matrix, m is the class mean, x is the signature of the voxel under consideration, and $p(\omega_i)$ is the prior probability that the voxel under consideration is a member of class i . A flowchart for this process is given in Figure 2.

ABSTRACT: A method for fully automating the measurement of various neurological structures in MRI is presented. This technique uses an atlas-based trained maximum likelihood classifier. The classifier requires a map of prior probabilities, which is obtained by registering a large number of previously classified data sets to the atlas and calculating the resulting probability that each represented tissue type or structure will appear at each voxel in the data set. Classification is then carried out using the standard maximum likelihood discriminant function, assuming normal statistics.

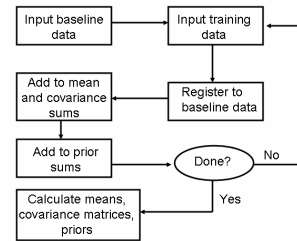


Figure 1: Flowchart showing the operation of the statistical training software used in this project

	Auto. Volume	Man. Volume	Abs. Error	Pct. Error	Sigma Error
Subj. 1R	3.314	2.977	0.337	11.3	10.9
Subj. 1L	3.089	2.769	0.32	11.5	0.56
Subj. 2R	3.714	3.771	-0.057	-1.5	-0.08
Subj. 2L	3.539	3.678	-0.139	-3.7	-0.18
Subj. 3R	3.411	3.16175	0.24925	7.8	0.29
Subj. 3L	4.006	2.989	1.017	34.0	1.64
Subj. 4R	4.973	3.30225	1.67075	50.5	2.72
Subj. 4L	4.535	3.166	1.369	43.2	1.72
Subj. 5R	3.998	3.14675	0.85125	27.0	0.87
Subj. 5L	4.815	2.86475	1.95025	68.0	2.72
Subj. 6R	3.225	3.47525	-0.2502	-7.2	-0.33
Subj. 6L	3.447	3.6565	-0.2095	-5.7	-0.34
Subj. 7R	3.036	2.80775	0.22825	8.1	0.32
Subj. 7L	3.089	2.60475	0.48425	18.5	1.06
Subj. 8R	4.323	3.66525	0.65775	17.9	0.77
Subj. 8L	4.829	3.52375	1.30525	37.0	2.12
Subj. 9R	3.839	3.44025	0.39875	11.5	0.34
Subj. 9L	3.408	3.26675	0.14125	4.3	0.16
Subj. 10R	3.751	3.376	0.375	11.1	0.56
Subj. 10L	3.804	3.64675	0.15725	4.3	0.49

Table1: Results of automated hippocampus measurement experiment. Volumes are given in cubic centimeters. Note that the automated measurements are within one standard deviation of the manual ones in 13/20 cases, and within two standard deviations in an additional 3 cases.

	Auto. Volume	Man. Volume	Abs. Error	Pct. Error	Sigma Error
Subj. 1	22.4	21.2	1.2	5.6	0.3
Subj. 2	25.7	24.1	1.6	6.6	0.57
Subj. 3	11.4	9.3	2.1	22.5	1.16

Table 2: Results of automated white matter lesion measurement experiment. Volumes are given in cubic centimeters. Note that automated measurements are within one standard deviation of manual ones in 2/3 cases.

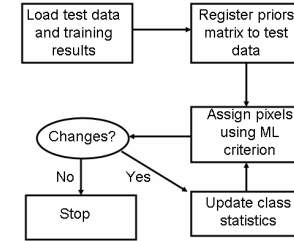


Figure 2: A flowchart describing the maximum likelihood classification scheme used in this work.

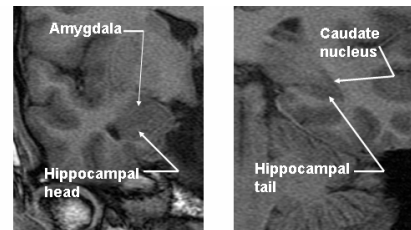


Figure 3: (a) Separation of the right hippocampal head from the basal nucleus of the amygdala. (b) Separation of the left hippocampal tail from the tail of the caudate nucleus.



Figure 4: In this case a mis-alignment of the anatomical atlas with Subject 4 has caused the identification of the hippocampus to include a portion of the collateral sulcus, resulting in an over-estimation of the total hippocampal volume.

EXPERIMENTAL PROCEDURE: This study was intended to determine the ability of the algorithms described in Section 2 to identify and measure both normal and abnormal neurological structures. We therefore made use of two experimental data sets. The first set was intended to allow measurement of abnormal structures. This data set consisted of T1, T2, and proton density weighted MRI scans for three multiple sclerosis patients, with 2 to 4 repeats for each patient taken over the course of 4 to 8 days. The second experimental data set was intended to allow the measurement of the hippocampus. This data set consisted of 10 coronal T1 weighted MRI studies taken from normal volunteers.

The experiments carried out during this study were intended to assess several parameters. First, we wished to determine the level of precision achievable in the volumetric measurement of both normal and abnormal neurological structures using manual tracing. The second parameter of interest was absolute accuracy of the automated measurement system. Because this system is fully automated, intra- and inter-observer variability were assumed to be zero. The results of the automated analyses were compared to the mean of the inter-observer variability measurements, with bias calculated in terms of volume, percent of manual measurement, and volume divided by the appropriate standard deviation.

RESULTS: The results of the automated hippocampus measurement experiment are given in Table 1. Note that comparison is made to the manual results, with error calculated in terms of manual standard deviations from the inter-observer mean value. Results for the automated white matter lesion measurement experiment are given in Table 2. Error is again calculated in terms of standard deviations from the inter-observer manual mean.

There were two primary sources of error in these measurements. The first, which related primarily to the hippocampal measurements, was ambiguity in the exact structural borders, particularly at the hippocampal head and tail. This problem affected both the automated and manual measurements, and is illustrated in Figure 3.

The second significant source of error also applied primarily to the hippocampal measurements. An examination of the 5 cases with significant deviations shows that in each of these cases the error is an over-measurement stemming from an error in registering the anatomical atlas volume to the test data. An example of this sort of error is given in Figure 4. Errors of this sort should be correctable through a refinement of the registration algorithm used in this work.

REFERENCES:

1. M. Filippi, V. Douzet, H. McFarland, D. Miller, R. Grossman. "Role of magnetic resonance imaging in the diagnosis and monitoring of multiple sclerosis: Consensus report of the white matter study group." *J Magn Reson Imaging* 15, pp. 499 - 504, 2002.
2. C. Jack, F. Shattough *et al.*, "Temporal lobe seizures: Lateralization with MR volume measurements of the hippocampal formation." *Neuroradiology* 47, pp. 423 - 429, 1998.
3. C. Jack, R. Peterson *et al.*, "MR-based hippocampal volumetry in the diagnosis of Alzheimer's disease." *Neurology* 42, pp. 183 - 188, 1992.
4. E. Ashton, K. Parker, M. Berg, C. Chen. "A novel volumetric feature extraction technique, with applications to MR images." *IEEE Trans Medical Imaging* 16(4), pp. 365 - 371, 1997.
5. Y. Ito, N. Schuff *et al.*, "Comparison of automated and manual MRI volumetry of hippocampus in normal aging and dementia." *Annals of MRI* 14, pp. 305 - 310, 2002.
6. A. Brunetti, A. Laobina, M. Quarantelli, M. Tedeschi, E. Ciurmello, A. Cowelli, and M. Salvatore. "Automated segmentation and measurement of global white matter lesion volume in patients with multiple sclerosis." *J. of MRI* 12(6), pp. 799 - 807, 2000.
7. M. Rovaris, M. Ingolese, R. van Schijndel, M. Sormani, M. Rodighiero, G. Comi, and M. Filippi. "Sensitivity and reproducibility of volume change measurements of different brain portions on magnetic resonance imaging in patients with multiple sclerosis." *J. of Neurology* 247(12), pp. 960 - 965, 2000.
8. P. Dastidar, T. Helenius, T. Lehtimäki, M. Ukkonen, J. Peltonen, J. Peltola, T. Erila, E. Lassonen, and I. Elvonen. "Volumes of brain atrophy and plaques correlated with neurological disability in secondary progressive multiple sclerosis." *J. of the Neurological Sciences* 165(1), pp. 36 - 42, 1999.
9. M. Ballester, A. Zuercher, and M. Brady. "Segmentation and measurement of brain structures in MRI including confidence bounds." *Medical Image Analysis* 4(3), pp. 199 - 200, 2000.
10. K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester, and P. Sootens. "Automated segmentation of multiple sclerosis lesions by model outlier detection." *IEEE Trans. Medical Imaging* 20(8), pp. 677 - 688, 2001.