Data Driven Brain Tumor Segmentation in MRI Using Probabilistic Reasoning over Space and Time

Jeffrey Solomon[1,3], John A. Butman[2], and Arun Sood[1]

[1] Center for Image Analysis, George Mason University,
{jsolomon,jbutmana}@cc.nih.gov
asood@cs.gmu.edu
[2]Diagnostic Radiology Department, Clinical Center, National Institutes of Health,
[3] Medical Numerics Inc.

Abstract. A semi-automated method for brain tumor segmentation and volume tracking has been developed. This method uses a pipeline approach to process MRI images. The pipeline process involves the following steps: 1) automatic alignment of initial and subsequent MRI scans for a given patient, 2) automatic de-skulling of all brain images, 3) automatic segmentation of the brain tumors using probabilistic reasoning over space and time with a semi-automatic correction of segmentation results on the first time point only and, 4) brain tumor tracking, providing a report of tumor volume change. To validate the procedure, we evaluated contrast enhanced MRI images from five brain tumor patients, each scanned at three times, several months apart. This data was processed and estimated tumor volume results show good agreement with manual tracing of 3D lesions over time.

1 Introduction

The best accepted measure of brain tumor viability is interval change in tumor size and decisions on efficacy of clinical treatments in a given patient and in clinical trials are most commonly based on this measure. Commonly used measurement methods to assess changes in lesion or tumor size over time include (1) greatest diameter [1], (2) greatest diameter multiplied by greatest perpendicular diameter [2], (3) manual lesion tracing, [3] (4) computer-based automated techniques to segment the lesion and calculate volume [4].

The one and two dimensional measures are accurate only when lesions are spherical or elliptical in nature. While this assumption is often invalid, these techniques are used in clinical situations because of their simplicity [3]. However, the use of such measurements on serial studies can be problematic as it is difficult to define precisely the same location for measurement of subsequent studies. As stated by Sorenson [3] and others, 3D segmentation techniques can improve lesion measurement accuracy relative to 1D and 2D methods. Although manual 3D techniques are more accurate, they are not commonly used clinically largely due to limitations of time involvement and inter-rater variability. As medical imaging devices provide higher spatial resolution, more image slices are created for each study. This requires processing of a significant number of images to determine lesion

C. Barillot, D.R. Haynor, and P. Hellier (Eds.): MICCAI 2004, LNCS 3216, pp. 301–309, 2004.

volume. There is a need to develop automated techniques that can detect lesions (e.g. tumors) in medical images and track their progression in size, shape and intensity.

An interesting class of intensity based 3D segmentation techniques uses Bayesian analysis for tissue classification. These methods assume that the image histogram can be modeled as a sum of Gaussian distributions, known as a Gaussian mixture model. The probability distribution for this model is:

$$P(x_j) = \sum_{i=1}^{k} P(C_i)P(x_j \mid C_i) \tag{1}$$

where $P(x_j)$ is the probability of pixel j having intensity x_j, $P(C_i)$ is the probability of class C_i in the image, k is the number of classes and $P(x_j|C_i)$ is the probability of obtaining pixel intensity x_j given that the tissue class is C_i. Using the approaches of others [5], $P(x_j|C_i)$ is assumed to follow a Gaussian probability distribution for each tissue type, C_i , in the brain. Bayes' Rule states that the probability of pixel j being in tissue class C_i given its intensity x_j is given by:

$$P(C_i \mid x_j) = \frac{P(x_j \mid C_i)P(C_i)}{P(x_j)} \tag{2}$$

The expectation-maximization (E-M) algorithm has been used to iteratively estimate the parameters of the Gaussians in the mixture model in order to classify various tissue classes [5]. To segment and track the progression of diseased tissue, e.g. multiple sclerosis lesions or brain tumors, in MR images of the brain, segmentation is performed independently on each 3D image of the 4D time series (three spatial dimensions plus time). Temporal information does not influence the segmentation.

Since there is some relation between images at different times, temporal information has the potential to improve segmentation of lesions in MR images. Perhaps, the simplest temporal technique is image differencing as explored in a paper by Hajnal et al [6]. This difference image only represents change in intensity not shape. If a lesion does not change much over time, it will not be detected by this system.

Detecting change by estimating the displacement field was proposed by Rey et al [7]. This automatic method performed well on the task of lesion detection but did not perform as well with segmentation accuracy. Gerig et al explored the use of pixel-based temporal analysis in lesion detection [8]. In this work, rigid registration and global intensity normalization (to account for global pixel intensity differences across scanning sessions) were performed on a series of MR images of patients with multiple sclerosis. This method was shown to be sensitive to registration accuracy producing false positives. In order to reduce the number of false positives resulting from registration errors, a spatio-temporal model was developed [9]. While this technique has some limitations, the work is significant in the author's realization that methods exclusively concentrating on the spatial or temporal aspects cannot be expected to provide optimal results.

It is hypothesized that using both temporal and spatial properties of the 4D image set will improve the automatic segmentation of lesions in the time series compared with techniques that independently detect lesions from one scan to the next or focus only on areas of change. Solomon and Sood [10] have demonstrated positive results comparing the E-M algorithm with a combined E-M+HMM (Hidden Markov Model) approach which incorporates probabilistic reasoning over space and time to segment

4D lesions in MR images. This work was validated on simulated data generated with Gaussian mixture model.

This paper describes the validation of this technique on sample data from 5 patients whose brain tumor volume was assessed both by manual tracings and by automated 4D segmentation. In addition, a practical pipeline approach to the automatic processing of the 4D MRI data sets was developed that minimizes user interaction.

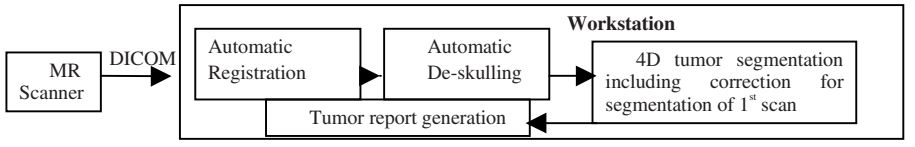

Fig. 1. MR images are sent from the scanner to a workstation via the DICOM protocol. The MR images are automatically registered using FLIRT. Next the brains are automatically skull stripped. The 4D data is segmented using probabilistic reasoning over space and time via the E-M+HMM method. As part of this technique, the segmentation of the 1st time point is manually corrected by removing tissue incorrectly classified as tumor by the E-M algorithm (e.g. enhancing vasculature). The remaining volumes in the time series are segmented automatically without user interaction. Finally a report is generated presenting the tumor volumes over time.

2 Methods

A pipeline approach was used to pre-process the series of MR images for each patient.

2.1 Registration Engine

A software program termed the "registration engine" was used to automatically align all MR images in the time series for each patient. This engine automatically searches for new DICOM formatted images on the workstation. If a new data set exists, the engine determines if a previous study exists for this patient. If not, the data set is saved as a reference scan for that patient. If a previous scan does exists, its reference is automatically retrieved and registered via the FLIRT [11] linear registration program with a 6 parameter rigid registration using correlation as the cost function and sinc interpolation to minimize effects of resampling the image data. The results of the registration are stored for further processing.

2.2 Automatic De-skulling

Once all desired images are registered, a process known as de-skulling is performed to exclude all extracranial tissue (e.g. scalp fat, head and neck musculature, bone marrow) from the MR image and minimize the number of tissue classes for further segmentation. The normal intracranial contents include predominantly brain tissue (gray and white matter), cerebrospinal fluid (CSF), and vasculature. In the pathologic cases, brain tumor serves as an additional tissue class. Automated programs have been developed for de-skulling, but do not behave as well on data with pathology, such as

the brain tumors in this study, as they do on normal brains. In addition to a change in morphology due to the tumors, these images are acquired with contrast agents making vessels appear hyper-intense. These vessels create bridges between tissues in the brain and the skull, thus making the automated delineation of the brain more difficult. We have used a template based approach to automatically de-skull brain MRIs in this group. This process is accomplished by registering the first time point in the MR series for a given patient to the MNI template brain. The transformation matrix derived from this 12-parameter affine registration [11] was saved using the MEDx software package (Medical Numerics, Inc. Sterling, VA). The matrix is inverted and applied to a previously de-skulled brain mask image, available from the SPM99 software package (http://www.fil.ion.ucl.ac.uk/spm/) which shares the same coordinate system as the MNI template. The pixel values in the brain mask image represent probability of brain tissue. This image was thresholded such that all voxels above a 60% probability of brain tissue were set to 1. This transformed brain mask is multiplied by all registered MR images in the time series, resulting in de-skulled time series of data ready for tumor segmentation.

2.3 4-D Lesion Segmentation Using Probabilistic Reasoning over Space and Time

Segmentation of aligned de-skulled images is performed using probabilistic reasoning in space (via the expectation-maximization) algorithm and time (using the hidden Markov model). Initialized by a K-means clustering, the E-M algorithm is applied to each volume in the time series. This results in estimates of the Gaussian parameters for the pixel distributions of all tissue classes (including tumor) for each volume. This information is used by a hidden Markov model which makes use of previous and subsequent scans to estimate the classification of a voxel in the image.

Change over time is modeled as a series of snapshots, each describing the state of the world at a particular time with a set of random variables (some observable and some not). If we assume that the entire state is determined by a single discrete unobservable random variable, we have a hidden Markov model (HMM).

If we consider the state of each voxel in our 4D medical image data to be lesion status, with possible values of lesion and non-lesion, then the problem of segmenting and tracking lesions over time can be modeled as a HMM. The state of the system (lesion/non-lesion), represented in this model as X_t, is not observable. What is observable is the pixel intensity value. This "evidence" variable will be represented as E_t. We will also make a 1st order Markov assumption (the current state can be determined completely from the previous state). The equation below represents this 1st order Markov state to state "Transition model".

$$P(\overline{X}_t \mid \overline{X}_{t-1}) = P(\overline{X}_t \mid \overline{X}_{0:t-1}) \qquad \text{Transition Model} \qquad (3)$$

The evidence variable E_t is dependent on only the current state. This dependence is known as the Sensor Model.

$$P(E_t \mid \overline{X}_t) = P(E_t \mid \overline{X}_{0:t-1}) \qquad \text{Sensor Model} \qquad (4)$$

Each voxel in the 4D image will independently follow the HMM.

In a process known as *filtering*, one computes the posterior distribution over the current state given all evidence to date $P(X_t|E_{1:t})$. In our case, if the state X_t equals

lesion or non-lesion and E_t is equal to the pixel's intensity, then $P(X_t|E_{1:t})$ provides the tissue classification given the pixel's intensity value. It can be shown that there is a recursive formula to express this posterior distribution:

$$P(\overline{X}_{t+1} \mid E_{1:t+1}) = \alpha P(E_{t+1} \mid \overline{X}_{t+1}) \sum_{x_t} P(\overline{X}_{t+1} \mid \overline{X}_t) P(\overline{X}_t \mid E_{1:t}) \tag{5}$$

The current state is projected forward from *t* to *t+1* then updated using new evidence E_{t+1}. Another process known as *smoothing* is used to compute past states given evidence up to the present. In other words, you may re-estimate the state at time points 1-4 given that you have estimates up to time point 5. This will theoretically improve the estimates of these states by using more information gathered. The following expression represents the probability of state X_k ($1 <= k < t$), given all evidence acquired so far.

$$P(\overline{X}_k \mid E_{1:t}) = \alpha P(\overline{X}_k \mid E_{1:k}) P(E_{k+1:t} \mid \overline{X}_k) \tag{6}$$

The right side of this equation contains a normalization constant, α, the familiar filtering expression from equation (5) and a backward expression, $P(E_{k+1:t}|X_k)$. This backward expression is given by the recursion:

$$P(E_{k+1:t} \mid \overline{X}_k) = \sum P(E_{k+1} \mid \overline{X}_{k+1}) P(E_{k+2:t} \mid \overline{X}_{k+1}) P(\overline{X}_{k+1} \mid \overline{X}_k) \tag{7}$$

The sensor model used is given below:

$$P(E_t \mid X_t(0)) = A_l(1/\sqrt{2\pi}\sigma_l) \exp(-(x-\mu_l)^2 / 2{\sigma_l}^2) \tag{8}$$

$$P(E_t \mid X_t(1)) = \sum_i A_i \sqrt{2\pi}\sigma_i \exp(-(x-\mu_i)^2 / 2{\sigma_i}^2) \tag{9}$$

where the probability at time *t* of having a pixel value equal to E_t given that this pixel represents lesion (state $X_t(0)$) is the Gaussian of the lesion tissue class *l*. The probability of having the pixel value E_t given that the pixel does not represent lesion (state $X_t(1)$)is the sum of the other Gaussians in the mixture model. There is a coefficient, α, used to sum the probabilities to 1 as seen in the equation below.

$$\alpha \sum_{i=0}^{1} P(E_t \mid X_t(i)) = 1 \tag{10}$$

A transition model was developed that is dependent on the distance from the lesion border and growth of the lesion. Exponential lesion growth is a simple model often employed [12] and is used here. The likelihood of a transition from lesion to non-lesion or vice versa is greater when the voxel is close to the lesion border (margin). This new model is shown below in equation (11).

$$T = P(\overline{X}_t \mid \overline{X}_{t-1}) = \begin{pmatrix} 1-\exp(-dist/C) & \exp(-dist/C) \\ \exp(-dist/C) & 1-\exp(-dist/C) \end{pmatrix} \tag{11}$$

The coefficient, *C* is a constant, and *dist* represents the distance from the border of the lesion. This distance is modified to take into account exponential lesion growth as follows:

$$dist_{t+1} = dist_t + exp(\gamma * \Delta t) \tag{12}$$

for voxels inside the lesion, and

$$dist_{t+1} = | exp(\gamma * \Delta t) - dist_t | \tag{13}$$

for voxels outside of the lesion. The term Δt represents the time between scans and γ represents the rate of exponential growth. This transition model is therefore unique for each voxel in the image and varies with each volume in the time series.

The choice of γ, the rate of exponential lesion growth, is initially estimated based on results of the E-M algorithm on successive image scans in the time series.

Theoretically, there is a separate γ for each point on the lesion border, allowing for modeling of lesion growth that is anisotropic. The transition model described here requires an estimate of the lesion border on the first scan in the series. This estimate was performed by running the E-M algorithm on the first time point and using the manual tracings to limit false positive results. This semi-automated correction is only performed on the first image of the temporal series.

2.4 Report Generation

The output of the 4D lesion segmentation process is a series of probabilistic images where voxel intensity represents the probability of being tumor. For classification purposes, any voxel whose probability was above 0.5 was labeled as tumor. A total volume of voxels assigned as tumor is computed. An html report is displayed showing tumor volume estimates over time.

2.5 Ground Truth Data

The 5 patients in the study each had 3 MRI scans with near isotropic voxel sizes (0.9375mm x 0.9375mm x 1.0mm). To validate the method, the tumors were manually traced at each time point, using MEDx software. The volumes based on these tracings were computed. Binary image masks were generated from these tracings for the validation process. In addition, the tracings of the tumor in the first scan were used in the correction step for the 4D tumor segmentation technique described in section 2.3.

3 Results

Scans from 5 patients were processed as described in the Methods section. Resulting tumor volumes taken from the generated report were compared with tumor volumes recording during the ground truth creation. Figure 2 shows the correlation between the tumor volumes calculated by the automatic 4D method and manual tracing. The correlation is strong with an R^2 value of 0.89.

In order to determine if the automated method classified as tumor the same voxels as the manual tracing, the Dice similarity coefficient (DSC) was used. The mean DSC value for all time point 2 and 3 measures was 0.71. Table 1 shows these values for the second and third time points in all 5 cases. Time point 1 was left out because the tumor volume estimated from the automated technique is based in part on the manual tracings for time point 1 as described in the methods section.

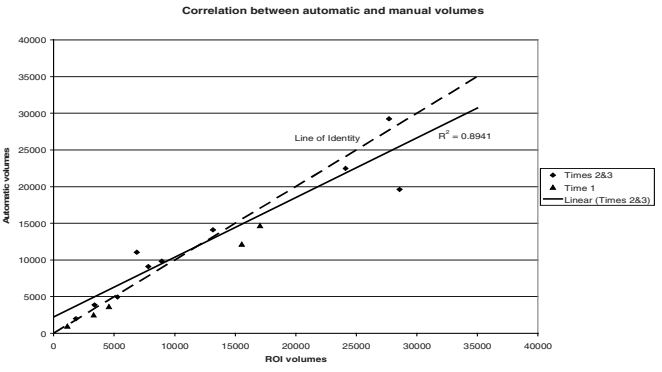

Fig. 2. Comparison of tumor volumes (# voxels) measured by 4D E-M+HMM to that measured by manual tracing. Identical measurements fall on the dashed line. The coefficient of determination, R^2 for the two methods was 0.89 (Correlation was performed for time points two and three as these were obtained without any manual interaction).

Table 1. Dice similarity coefficients for Time points 2 and 3. These measures show the amount of overlap between the automated 4D tumor volume and the manual tracing.

	DSC Time 2	DSC Time 3
Patient 1	0.82	0.84
Patient 2	0.75	0.77
Patient 3	0.70	0.64
Patient 4	0.79	0.75
Patient 5	0.62	0.45

The low DSC for patient 5 is attributed to the manual tracing of the irregularly shaped tumor.

Figure 3 below illustrates one section of the tumor from the MRI of Patient 3. Figure 3A illustrates an enhancing brainstem tumor which is relatively hyperintense and shares signal characteristics with the nearby cavernous sinuses. Figure 3B shows the voxels classified as tumor in white with the black outline representing manual tracing (for validation purposes). Despite the irregularity of the tumor, as well as iso-intensity with the cavernous sinus, the 4D E-M+HMM algorithm reliably classifies the enhancing portion of the tumor and segments it from the cavernous sinuses.

4 Discussion

This paper discusses an automated pipeline process for segmenting and tracking brain tumor lesions from MRI data requiring minimal user interaction. A 4D segmentation algorithm that makes use of probabilistic reasoning over space and time (E-M+HMM) has been tested on 5 patients. Previous work validated this technique on simulated

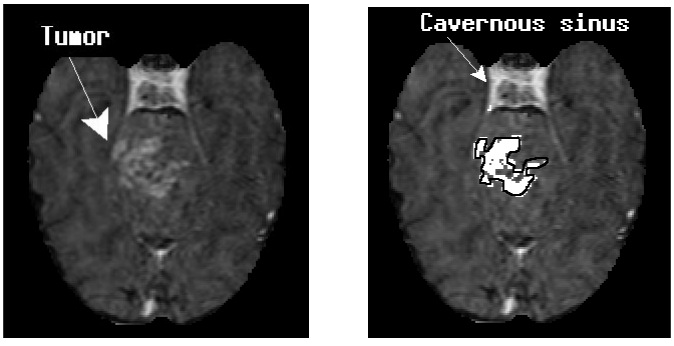

Fig. 3. A) MRI image of a tumor at time point 2 after pre-processing but before tumor segmentation. The tumor shares signal intensity with the vessels of the cavernous sinus. B) Automatic segmentation of the tumor via 4D E-M+HMM (in white) matches the enhancing tumor. The black contour represents manual tracing for validation purposes. Note that the cavernous sinus is not identified as tumor.

images demonstrating that this 4D method yields better segmentation accuracy for the current scan than independently segmenting with the E-M algorithm alone [10]. The proposed segmentation technique also has advantages over change detection techniques. By making use of both spatial and temporal information, our technique will detect tumors even if they are not changing significantly. While our technique is dependent on proper alignment across time points, it is not dependent on intensity normalization since the E-M algorithm is applied for each time point in the series. In addition, changes in tissue distribution (e.g. tumor less or more enhancing) are handled. This automated tumor segmentation process produces a report containing tumor volume change, providing estimates of tumor progression which can be incorporated into treatment planning.

One can certainly ask to what degree temporal information improves segmentation as compared to simply performing E-M at each time point. We could not perform the direct comparison as the E-M alone fails to segment brain tumors in any individual data sets without some form of manual interaction. This is because there are spatially distinct regions in the brain (notably blood vessels) which share signal characteristics with the brain tumor so that E-M classifies both tissue types into the same class. Manually excluding this data on the first data set allows the HMM to propagate this information forward and perform subsequent segmentation in a fully automated way. This is perhaps the most important advantage of applying the HMM in conjunction with the E-M algorithm.

Correlation of brain tumor volumes with manual tracings was strong. However, a degree of mismatch in overlapping areas was caused primarily by difficulty in manual tracing of the tumor boundaries. The tumors are quite irregular, making the manual segmentation task difficult. Future work will include validation on additional patients as well as evaluating how estimates and models of the growth rate of the tumors affect the segmentation results obtained from the HMM model. By estimating the rate of growth for each point on the tumor boundary, it may be possible to predict tumor location at times in the future. This might have further application in treatment planning. Another future goal is to include the Markov Random Field as a prior

probability of the segmentation, modeling spatial correlation among neighborhood voxels in the 4D segmentation algorithm.

Automated tumor segmentation has the potential to play an important role in medical imaging, both in the management of individual patients and in the conduct of clinical trials where changes in tumor volume is a primary endpoint.

References

1. R. Therasse, S. Arbuck, E.A. Eisenhauer et al, "New Guidelines to Evaluate the Response to Treatment in Solid Tumors", *Journal of the National Cancer Institute,* Vol 92, pp. 205-216, 2000
2. A.B. Miller, B. Hoogstraten, M. Staquet et al, "Reporting Results of Cancer Treatment" *Cancer,* 47, pp. 207-214, 1981.
3. G. Sorensen, S. Patel, C. Harmath, et al, "Comparison of Diameter and Perimeter Methods for Tumor Volume Calculation", *Journal of Clinical Oncology*, Vol 19, Issue 2, pp.551-557, 2001.
4. S.M. Haney, P.M. Thompson, T.F. Cloughesy, et al, "Tracking Tumor Growth Rates in Patients with Malignant Gliomas: A Test of Two Algorithms", *Am J Neuroradiology*, 22:73-82, January 2001.
5. K. Van Leemput, F. Maes, D. Vandermeulen, et al, "Automatic segmentation of brain tissues and MR bias field correction using a digital brain atlas" *Proceedings of Medical Image Computing and Computer-Assisted Intervention – MICCAI'98*, volume 1496 of Lecture Notes in Computer Science, pages 1222-1229, Springer, 1998.
6. S. Hajnal, W. Oatridge, B. Young, "Detection of Subtle Brain Changes Using Subvoxel Registration and Subtraction of Serial MR Images", *Journal of Computer Assisted Tomography*, 19(5) pp.677-691, 1995.
7. D. Rey, G. Subsol, H. Delingette, et al, "Automatic Detection and Segmentation of Evolving Processes in 3D Medical Images: Applications to Multiple Sclerosis", *Medical Image Analysis*, 6, pp. 163-179, 2002.
8. G. Gerig, D. Welti, C. Guttmann, et al, "Exploring the discrimination power of the time domain for segmentation and characterization of active lesion in serial MR data", *Medical Image Analysis*, 4, pp.31-42, 2000.
9. D. Welti, G. Gerig, E.W. Radu, et al, "Spatio-temporal Segmentation of Active Multiple Sclerosis Lesions in Serial MRI Data", *IPMI*, *LNCS 2082*, pp. 438-445, 2001.
10. J. Solomon, A. Sood, "*4-D Lesion Detection using Expectation-Maximization and Hidden Markov Model*", Proceedings of the IEEE International Symposium on Biomedical Imaging. (ISBI), pp 125-128. April, 2004.
11. M. Jenkinson, S. Smith, "A global optimisation method for robust affine registration of brain images", *Medical Image Analysis*, 5(2), pp.143-156, 2001.
12. K. Swanson, C. Bridge, J. Murray et al, "Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion", *Journal of the Neurological Sciences*, 216 pp. 1-10, 2003.