

Partial volume segmentation in 3D of lesions and tissues in magnetic resonance images

Brian Johnston, M. Stella Atkins[†], Kellogg S. Booth

Department of Computer Science, University of British Columbia
Vancouver BC, V6T 1Z4, Canada

[†]School of Computing Science, Simon Fraser University
Burnaby BC, V5A 1S6, Canada

ABSTRACT

An important first step in diagnosis and treatment planning using tomographic imaging is differentiating and quantifying diseased as well as healthy tissue. One of the difficulties encountered in solving this problem to date has been distinguishing the partial volume constituents of each voxel in the image volume. Most proposed solutions to this problem involve analysis of planar images, in sequence, in two dimensions only. We have extended a model-based method of image segmentation which applies the technique of iterated conditional modes in three dimensions. A minimum of user intervention is required to train the algorithm. Partial volume estimates for each voxel in the image are obtained yielding fractional composition of multiple tissue types for individual voxels. A multispectral approach is applied, where spatially registered data sets are available. The algorithm is simple and has been parallelized using a dataflow programming environment to reduce the computational burden. The algorithm has been used to segment dual echo MRI data sets of Multiple Sclerosis patients using lesions, gray matter, white matter, and cerebrospinal fluid as the partial volume constituents. The results of the application of the algorithm to these datasets is presented and compared to the manual lesion segmentations of the same data.

1. ITERATED CONDITIONAL MODES

A modeling method which has previously been applied to the restoration of corrupted images^{2,5} has recently been adapted, with varying degrees of success, to the segmentation problem.^{3,4,8-10} Using this model, an image is assumed to represent a locally dependent Markov random field (MRF), exploiting the assumption that voxels in a given local neighbourhood of the image have similar intensity.

Consider the 3D cubic lattice of the image space as a set, S , of n voxels indexed in some manner from $i = 1, 2, \dots, n$. Let the observed data y represent one set of data values on S . The value y_i denotes the observed record at position i . A segmentation of y will be represented by a set x . The value x_i denotes the segmentation value at pixel i . Let x^* represent the true segmentation of y .

The true segmentation x^* is a realization of a locally dependent MRF with distribution $p(x)$. A probability distribution $p(x)$ is a MRF if the following condition holds:

$$p(x_i | \{x_{k \setminus i}\}) = p(x_i | x_{N_i})$$

where $x_{k \setminus i} = \{x_k | k \neq i\}$, N_i indexes a neighbourhood system around voxel i , but not including index i , and $x_{N_i} =$

$\{x_j | j \in N_i\}$. Of all the possible interactions between these neighbourhood systems and voxel i we shall be concerned with pairwise interactions only. The problem then is to determine the *a posteriori* probability of a given segmentation, x , given the observed data, y and this neighbour interaction model. The state, x' , that maximizes this probability is the maximum *a posteriori* (MAP) estimate of the true segmentation x^* and is the mode of the posterior distribution of x^* . Given the MRF assumption, finding the MAP estimate is equivalent to maximizing

$$p(y_i | x_i) p(x_i | x'_{N_i}) \quad (1)$$

at every voxel site, i , in the image.

Application of Equation 1 to each voxel in an image constitutes a single iteration of what Besag has coined *Iterated Conditional Modes* (ICM).²

2. 3D MULTISPECTRAL STOCHASTIC VOLUME SEGMENTATION

We have adapted the Iterated Conditional Modes approach to develop a quasi-automatic solution to the segmentation problem in 3D.⁶ Extending the theory of MRF's to 3D, we determine partial membership of voxels in several tissue types.

The partial volume effect is one of the major difficulties encountered in medical image segmentation. We have endeavoured to provide a solution which yields a probability classification for each voxel, indicating the percentage composition of different tissues in that voxel. Each x_i will now be considered a vector of values indexed from $k = 1, 2, \dots, t$, where t is the number of tissue types present in the data volume. The k th value of element x_i , which we shall call its labeling, represents the probability that voxel i is of tissue type k . These probabilities can be viewed as the partial volume contribution from each tissue type to voxel x_i . As such, $\sum_{k=1}^t p(x_{ik}) = 1.0$ must hold at each iteration of the algorithm.

It has been shown^{2,6,8} that given only pairwise neighbour interactions and two estimates of the segmentation that differ only at index i , the probability that voxel x_i belongs to a particular tissue type, k , is given by

$$p(x_{ik}) = p(y_i | x_{ik}) * p(x'_{ik}) e^{-Z} \quad (2)$$

where

$$Z = \sum_{l=1}^t \beta_{kl} u_i(l)$$

and $p(y_i | x_{ik})$ is the conditional probability of the record, given its labeling (obtained from a set of initial histograms as described below), $p(x'_{ik})$ is the provisional estimate of the probability of pixel i belonging to tissue type k and β_{kl} is a measure of the interaction between tissues labeled k and l obtained using partial memberships of voxels in a particular tissue class. The term $u_i(l)$ is the *number* of neighbours of voxel x_i having label l . It is likewise determined from partial volume memberships weighted by inter-voxel distances.

A minimum of user intervention is required to train the ICM algorithm by requiring the manual outlining of regions of interest in a single image selected from the volume. Each of these possibly disjoint regions represents the user's belief that the voxels contained within the circumscribed regions are *pure voxels*. Normalized, low-pass-filtered histograms are obtained from each of these regions and used as input into the relaxation algorithm. These histograms provide the initial estimate of the segmentation, which in turn provides the initial neighbour interaction parameters and energy term of Equation 2.

Application of Equation 2 to each voxel in 3D constitutes one iteration of the algorithm. After each iteration or a predetermined number of iterations, the histograms and neighbourhood interaction parameters may be recalculated. Relaxation of the segmentation occurs after approximately six to ten iterations. Following convergence, a partial volume estimation is produced, indicating the probability that each candidate voxel is of a particular tissue type.

In order to accommodate a partial volume solution some modifications have been adopted to automatically calculate neighbour interactions and neighbour counts in 3D.⁶ In light of the fact that voxels are rarely cubic, the neighbour interactions and neighbour counts are further weighted by the intervoxel distances between pixel i and its 26 neighbours in 3D.

We have defined the neighbour interaction parameters β_{kl} to be the inverse of the probability that tissues k and l are adjacent, as opposed to the notion of the probability that tissue k is adjacent to tissue l .⁶ These adjacencies are weighted by the partial volume consistencies of each voxel and by the relative distance from each voxel to its neighbours in 3-space. The probabilities are determined from the tallied adjacencies for each tissue pair (k,l) in the provisional estimate of the segmentation, x' . These neighbourhood interaction parameters are determined from the initial segmentation estimate and may be updated after a predetermined number of iterations, thereby refining the accuracy of the segmentation.

When available, two or more independent sets of data may be used to drive the segmentation algorithm. We have performed the segmentation on multiple echo MRI sequences, but other combinations of data are just as valid, provided the data sets are independent and spatially registered. For example, patient scans from PET analysis may be combined with MRI or CT. Since the data sets are independently acquired, the resulting probabilities calculated by the ICM algorithm may be combined by multiplying corresponding values. The result of the multiplication is again normalized such that $\sum_{k=1}^t p(x_{ik}) = 1.0$. We have found that the results obtained from using two or more data sets are consistently more accurate than those obtained from using either of the data sets alone.⁷

The MRF assumption allows for high levels of parallelism in the implementation of ICM in 3D. Since, at each iteration, the new segmentation estimates are entirely neighbourhood based, we could conceivably allow updating of each voxel to be carried out by a separate processor. This would require the use of highly specialized hardware, such as a transputer network. The parallelism that we have achieved is provided by the image processing development environment that we adopted as part of the implementation of this algorithm.^{1,6}

3. TESTING AND RESULTS

We have applied our version of the ICM algorithm in 3D to MRI scans of the head obtained from Multiple Sclerosis (MS) patients. Here we discuss the results of using our algorithm to quantify the extent of Multiple Sclerosis lesions.

For our analysis, we have used dual echo MRI sequences (T1 and T2), each consisting of 256x256x27 slices acquired in the axial plane on a General Electric 1.5 Tesla MRI scanner. The slices are contiguous (no interslice gap) with each voxel measuring 0.781mm x 0.781mm x 5.0mm in the x , y , and z planes respectively.

We have attempted to segment four different tissues from the scans: white matter, gray matter, cerebrospinal fluid and MS lesions. For this analysis, we have elected to recalculate histograms and neighbourhood interaction parameters just once, following convergence of the first segmentation which occurred after approximately eight iterations. Convergence of the second run of the segmentation occurred following approximately six iterations of the algorithm.

In order to reduce bias introduced by the experimenter, the initial regions of interest required by the algorithm have been manually outlined on selected images by a radiologist. This was done independent of our research. The ROIs so obtained were applied to the corresponding images in order to obtain characteristic histograms for each of the tissue types under consideration.

3.1. Initial histograms

The initial histograms obtained by ROI analysis of the T1 and T2 images are shown in Figure 1 and Figure 2 respectively. The x axis represents the intensity of voxels in the chosen ROIs while the y axis denotes the probability or frequency of

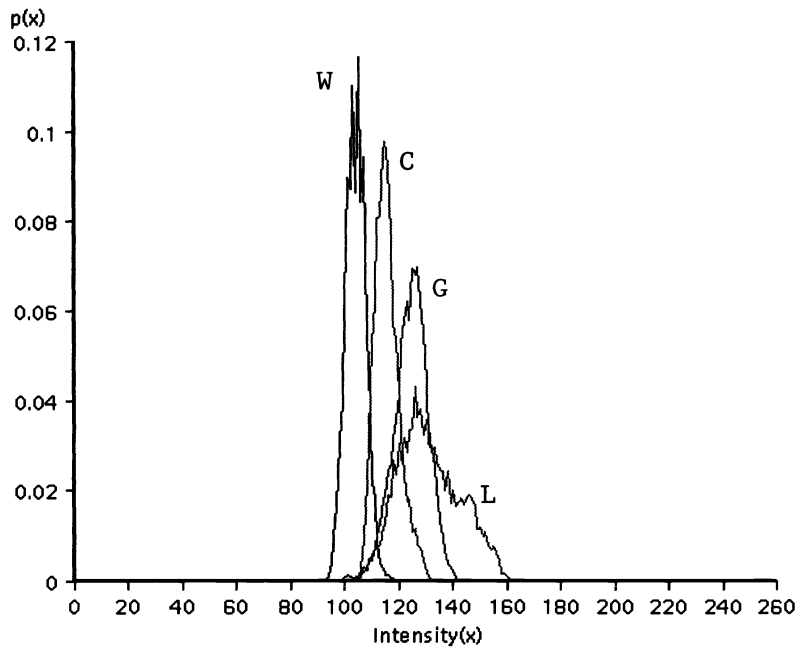


Figure 1: Histograms of T1 weighted MRI series based on manually drawn ROIs over a single image. W = White Matter, G = Gray Matter, C = CSF, L = MS Lesion.

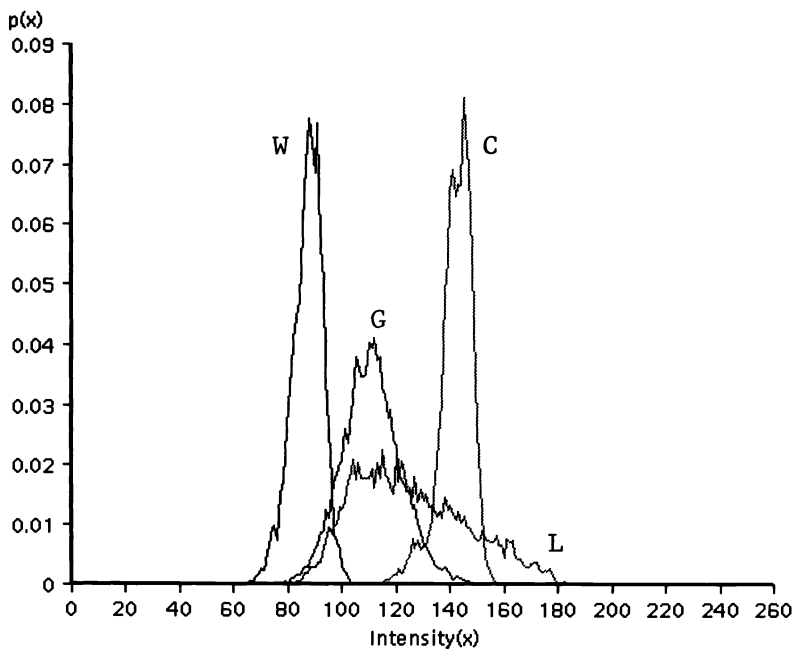


Figure 2: Histograms of T2 weighted MRI series based on same manually drawn ROIs as for Figure 1. W = White Matter, G = Gray Matter, C = CSF, L = MS Lesion.

Table 1: Initial neighbourhood interaction parameters obtained from the T1 image sequence.

<i>Tissue</i>	Gray	White	CSF	Lesion
Gray	12.77	14.26	7.76	7.66
White	14.26	4.47	4.51	35.64
CSF	7.76	4.51	9.59	10.09
Lesion	7.66	35.64	10.09	18.28

Table 2: Initial neighbourhood interaction parameters obtained from the T2 image sequence.

<i>Tissue</i>	Gray	White	CSF	Lesion
Gray	9.02	5.05	35.74	5.90
White	5.05	4.03	8866.84	9.18
CSF	35.74	8866.84	12.84	23.95
Lesion	5.90	9.18	23.95	10.57

voxels in the ROI at a particular intensity. These histograms have been smoothed to reduce noise in the distributions and have been normalized in order to give probabilities as opposed to actual numbers of voxels at particular intensities. As such, the area under each histogram is equal to one. These histograms show the distributions of each of the four tissue types to be segmented in each echo sequence. The ROIs were applied over corresponding images in both MRI series in order to obtain histograms for each spectrum.

There are several observations that can be made by visual inspection of these histograms. First, note that the histogram for white matter is somewhat isolated from that of the other three tissues. This suggests that white matter is most easily segmented from brain MRI images. This is in fact the case, as will be demonstrated in later sections. The second factor to note is that the histograms for the MS lesions are broad and relatively low, indicating the non-homogeneity of these lesions. As well, the lesion histograms illustrate a tendency to be noisier than the *pure* tissues, indicated by the many small sharp spikes present in the lesion histograms. This non-homogeneity is characteristic of many pathological conditions of the brain which result in lesions, including tumours and other neurological disorders. This non-homogeneity is also what magnifies the degree of difficulty in automatically segmenting brain lesions by computer.

Another complicating factor is the amount of overlap exhibited by the lesion histograms. Any segmentation algorithm will, as a result of this extensive overlap, have significant problems differentiating lesions from the remaining tissue types in the volume. As well as the lesion histograms being broad and flat it can be seen from the graphs that the other tissues also overlap, to a significant degree, in intensity profiles. The *shoulders* of each histogram overlap with its neighbours in the graph, sharing intensity distributions.

Lastly, the dynamic range of the data is rather small. For example, the T1 data shows a minimum intensity value of 93 and a maximum of 162. The dynamic range is actually significantly larger for both image spectra since not all tissue types are represented in the histograms (e.g. skin, bone, fat, vessels, etc.). However, the narrow range of the data serves to further complicate the segmentation process.

3.2. Initial neighbourhood interaction parameters

The neighbourhood interaction parameters calculated from the initial histograms of the T1 and T2 image volumes are shown in Tables 1 and 2 respectively. Relatively *low* numbers indicate high probabilities of tissues occurring adjacent

to each other, while *high* numbers illustrate tissues which rarely occur adjacent to each other. This is a direct result of how these parameters are defined and incorporated into the ICM equation. For the most part, the entries in the T1 and T2 neighbourhood interaction tables are consistent with known neuro-anatomy, with some exceptions. These numbers are subject to the accuracy of the initial segmentation, which in turn relies on the initial histograms. Because of this, some numbers in the tables appear to be somewhat inaccurate, in terms of known neuro-anatomical features.

For example, MS lesions occur almost exclusively in the white matter of the brain. However, the (white matter:lesion) adjacency measure given in Table 1 would suggest otherwise. If we keep in mind that the relative numbers of white matter (and other non-lesion) voxels far exceeds the numbers of lesion voxels, then this number seems more realistic. Another reason for the apparent contradiction in the (white matter:lesion) interaction probability is the nature of the MS lesion itself. Although occurring in white matter almost exclusively, the lesions do not have sharp transitions from lesion to white matter. It can be seen on analysis that the intensity of MS lesions declines when measured from the centres of the lesions to the edges. Because the decline is gradual and because white matter is dark and lesion is brightest on both T1 and T2 images, the intensity profile of the MS lesions passes through an area shared with gray matter. This phenomenon is borne out by the histograms of Figures 1 and 2. This results in a higher than expected value in the interaction table for the (white matter:lesion) adjacency measure, and a lower than expected value for (gray:lesion).

Consider the diagonal elements of Table 1. These numbers can be considered *self*-adjacencies, yielding measures of the relative quantities of different tissues and a measure of the convolutedness of the tissue structure. White matter, for example, which constitutes most of the brain matter, yields a very low number, while lesion, which occurs infrequently, and thus must share a larger percentage of its total voxels with other tissues, exhibits a relatively large *self*-adjacency number. This pattern is repeated in Table 2.

When Table 1 is compared with Table 2, the corresponding entries are relatively consistent, but for a few notable exceptions. These differences are due to differences in the characteristics of the T1 and T2 imaging sequences of MRI, as well as the complex nature of tissue interaction as measured by our algorithm. In particular the (CSF:white matter) interaction parameter is highly contrasted between the two tables. This results from the poor differentiability of CSF in the T1 image volume, due to the nature of the nature of the imaging sequence. The high value given in Table 2 is more plausible than the low value of Table 1, as CSF and white matter do not occur adjacently in healthy brain tissue.

These neighbourhood interaction parameters are affected by a number of complex factors, including (i) the response of tissues to the applied echo sequence, (ii) the relative volumes of tissues in the organ being imaged, (iii) the relative amounts of adjoining surface area in different tissues and (iv) the accuracy of the segmentation which has been used to calculate the parameters. The complexity of these factors interacting together makes it extremely difficult to offer accurate analysis of the neighbourhood interaction tables without extensive application of the ICM algorithm on many data sets.

Following initial convergence of the ICM algorithm, the histograms and neighbourhood interaction parameters are recalculated and used as input to the algorithm a second time. This recalculation of parameters is performed in order to refine the segmentation, assuming the new parameters are more accurate than those used initially. Since we are using intensity and segmentation values from the entire volume in the second run of the algorithm, as opposed to just a single slice, this assumption is realistic.

3.3. Recalculated histograms

The new histograms are calculated using partial volume contributions from each tissue to each voxel. Thus, the manner in which the histograms are obtained is fundamentally different than that used to obtain the initial histograms, i.e. manual segmentation. We would expect, theoretically at least, that the new histograms would be quite similar to the initial versions. This expectation is fulfilled despite the dramatically different manner in which the two sets of histograms are obtained. Figures 3 and 4 show the results of reassessment of tissue profiles in the intermediate segmentation of the T1 and T2 image sequences respectively. Note the similarities between these and the original histograms of Figures 1 and 2. Again we notice the overlap in tissue distributions. Most notably, it is shown that the lesion intensity profiles are shared by several other

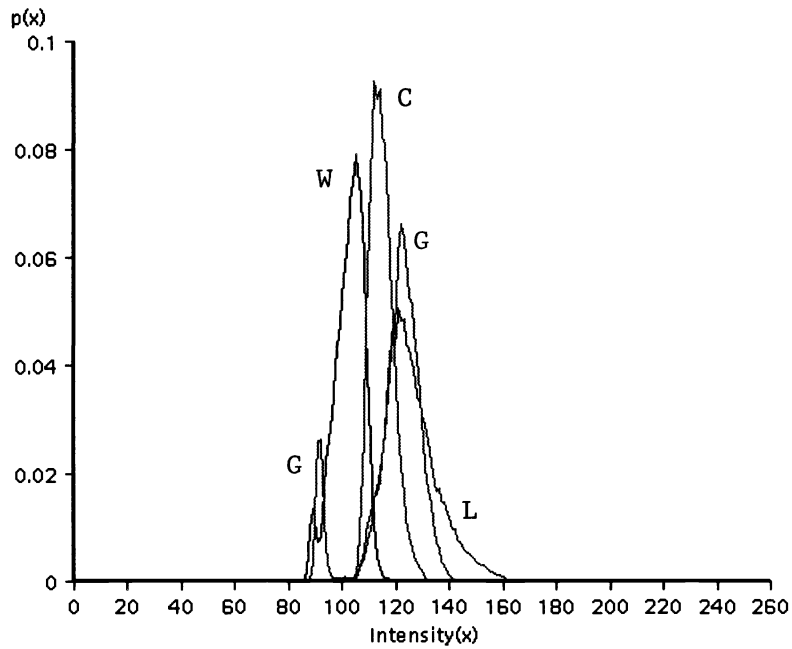


Figure 3: Histograms of T1 weighted MRI series based on recalculation from intermediate segmentation. W = White Matter, G = Gray Matter, C = CSF, L = MS Lesion.

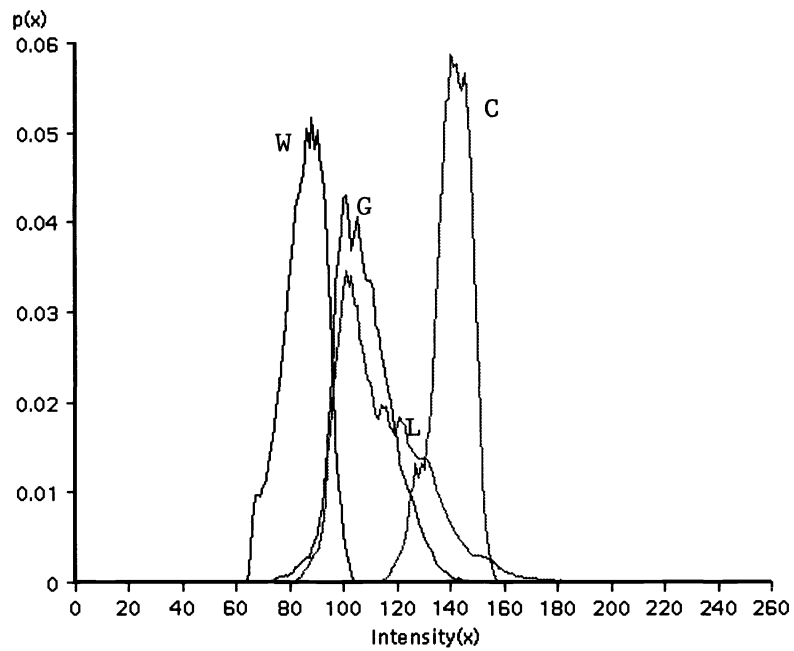


Figure 4: Histograms of T2 weighted MRI series based on recalculation from intermediate segmentation. W = White Matter, G = Gray Matter, C = CSF, L = MS Lesion.

tissues, especially gray matter and CSF. The white matter histogram is, again, less affected by other tissue profiles, though not to the same extent as with the original distributions.

What is different about these histograms is that the lesion profiles have become less distributed. They are taller and thinner, perhaps indicating that the first run of the segmentation isolates the lesions from the tissues. As well as the difference in the shape of the lesion profiles, another interesting observation is that these sets of histograms, especially the lesion histogram, are less noisy than the originals. Both these factors should improve the second run of the segmentation.

One characteristic of the new histograms which is of particular note is the spike in the low intensity end of the T1 gray matter profile of Figure 3. This spike occurs a significant distance away from the main gray matter histogram and may be a result of other tissues in the volume which were not specified in the original histograms, but which were partially classified in the intermediate segmentation.

3.4. Segmentation results

We have applied the ICM algorithm using the aforementioned histograms and neighbourhood interaction parameters to each of the T1 and T2 weighted MRI series and then combined the results by multiplying and normalizing the respective probabilities for each tissue type. We are able to do this because the two series have been acquired independently. We have selected from the segmentation results a single slice from the volume to illustrate the performance of the ICM algorithm. Figure 5 shows the result of the combined segmentation on slice number 18 (of 27) from the volume. The top four images show the segmentation of CSF, gray matter, white matter and lesion, as labeled in the figure. The original slice data (from the T2 sequence), with and without the manually segmented lesions overlaid on the slice, are given in the bottom two images. The gray scale intensity of each of the four segmentation images is in direct proportion to the probability obtained for that tissue in the image.

As anticipated, the results for white matter are by far the most impressive. Some lesions have been incorrectly classified as gray matter and some gray matter incorrectly classified as lesion. However, for the most part we have successfully isolated most lesions in the image. They are clearly differentiable in the lesion segmentation image as the most intense regions of the image.

3.4.1. Post-processing of the segmentation

Following the combined segmentation it was decided to apply some further post-processing to the lesion segmentations in order to more fully extract them from the volume. A morphological opening was applied to the combined result to reduce the number of isolated voxels in the result. The result of this opening is given in Figure 6 (top right). The top left image in the figure is the lesion segmentation resulting from combination of the results from the T1 and T2 segmentations, as given in Figure 5. The opening has clearly removed a number of island voxels without significantly affecting the lesion segmentation.

Since the intensities given represent probabilities that a given voxel is a lesion voxel, we decided to threshold the probability and only consider probabilities which were over a particular measure, in this case 0.4. This thresholding operation was applied to the morphologically opened result obtained previously. Visually, this threshold gave the best result when compared with the manual segmentation. The thresholded final segmentation is given in the bottom right of Figure 6.

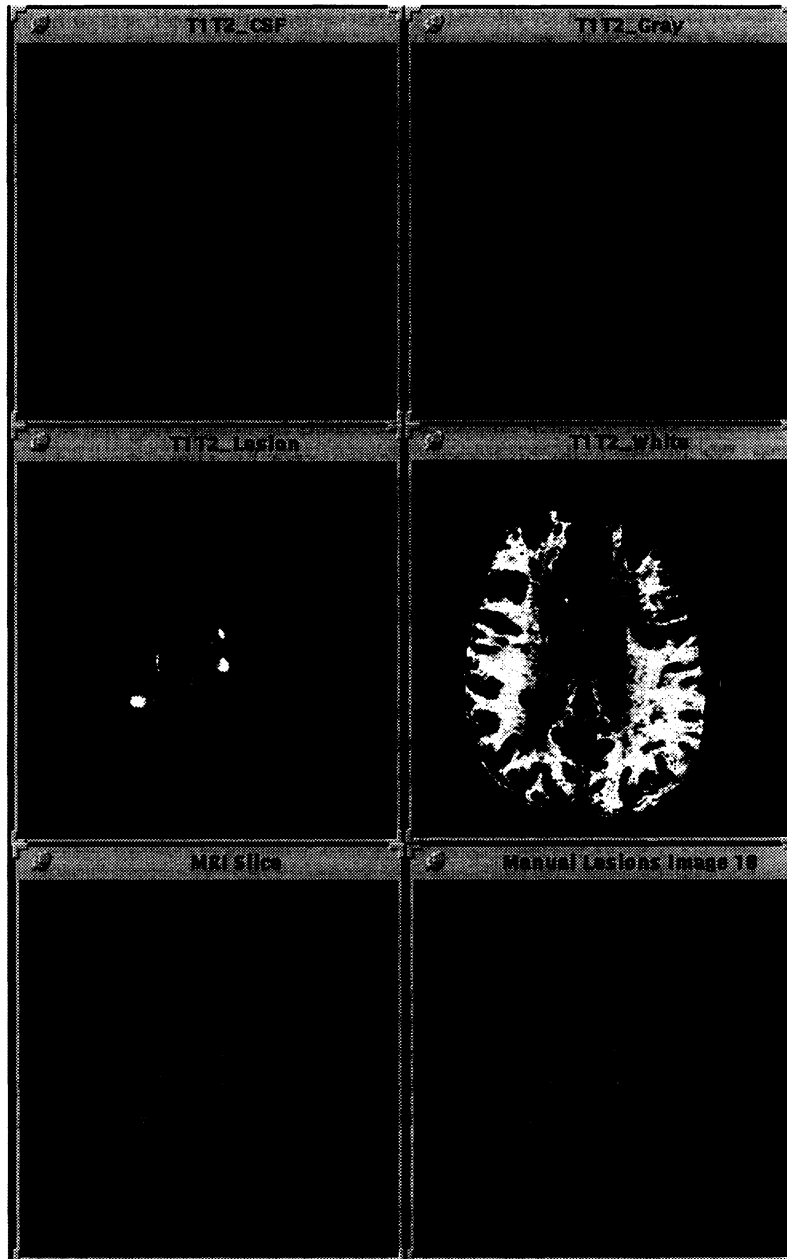


Figure 5: Segmentation of a single slice obtained by combining the T1/T2 Segmentations The bottom images show the original T2 data with (right) and without (left) the overlaid hand drawn lesion segmentations.

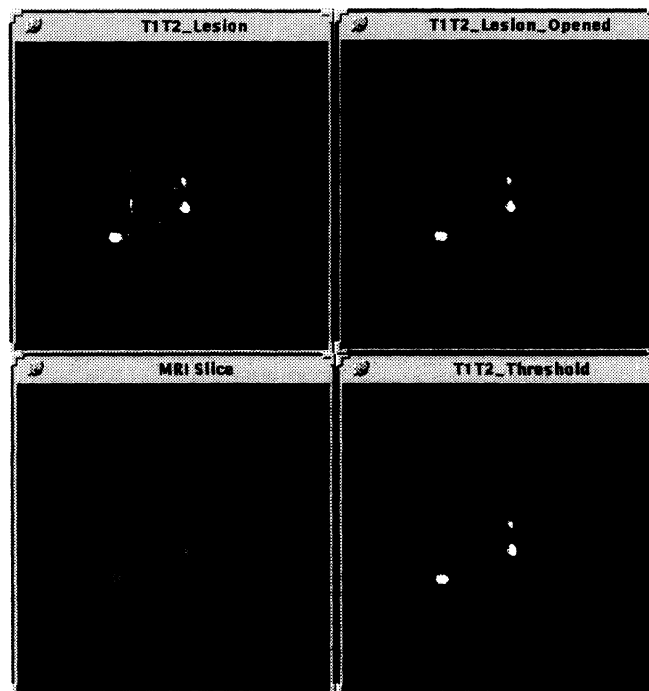


Figure 6: Lesion segmentation of a single slice obtained by applying a morphological opening (upper right) and subsequent thresholding (lower right) to the results of Figure 5. The bottom left image is the original data. (The background intensity has been lightened to enhance visualization.)

3.5. Lesion volume analysis

We have further examined our extensions to ICM by measuring the lesion volumes given by our segmentations. Lesions were manually outlined on every slice in the volume by a radiologist. We analysed the ROIs outlining the lesions in order to measure the lesion volume in each slice. Any voxel which lay within the borders of the manual lesions was considered a complete lesion voxel for our tests. The volumes so obtained can be considered a gold-standard by which we can measure the efficacy of our segmentation algorithm. However, lesion identification and delineation, as admitted by our experts, is an extremely complex and subjective task, requiring years of training and practise. It is with this in mind that we are able to make preliminary judgements regarding the accuracy of our algorithm.

Following our segmentation and the analysis of the manual lesion ROIs, we determined the volumes produced by our own results. We used the partial volume quantity for each voxel in summing the volumes obtained by our segmentations. The volumes of each slice were measured for (i) the T1 segmentation, (ii) the T2 segmentation, (iii) the combined T1/T2 segmentation, (iv) the combined segmentation following morphological opening, and (v) the final results obtained from thresholding the opened segmentation. The results have been plotted and are given in Figure 7. Note how, as the results are first combined and then post-processed, the ensuing volumes tend to more closely match the actual volumes determined from manual segmentation, also presented in Figure 7.

The graphs for the final segmentation and the hand-drawn lesions have been expanded and are shown in Figure 8. For the most part, the volumes obtained by the ICM segmentation are consistently and proportionately less than those measured from the manual segmentation. This is especially true of the middle slices in the volume. We believe this is due to several factors: (i) the manual segmentation uses full values for each voxel included in the lesion ROIs, while our segmentation measures partial volumes, (ii) errors by the radiologist in lesion detection, and (iii) errors in the segmentation. The first factor is, in itself, not to be considered an error but perhaps an adjustment made by our program that is not possible (or at

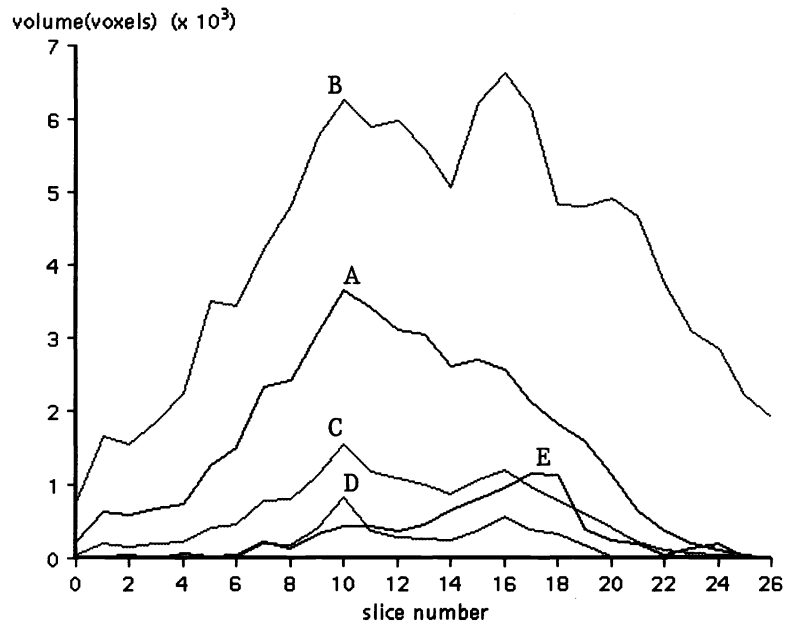


Figure 7: Graphs of lesion volumes determined at all stages of the segmentation, including manual segmentation. (A) = T1, (B) = T2, (C) = T1/T2 combined, morphologically opened, (D) = T1/T2 combined, opened and thresholded, (E) = manual segmentation.

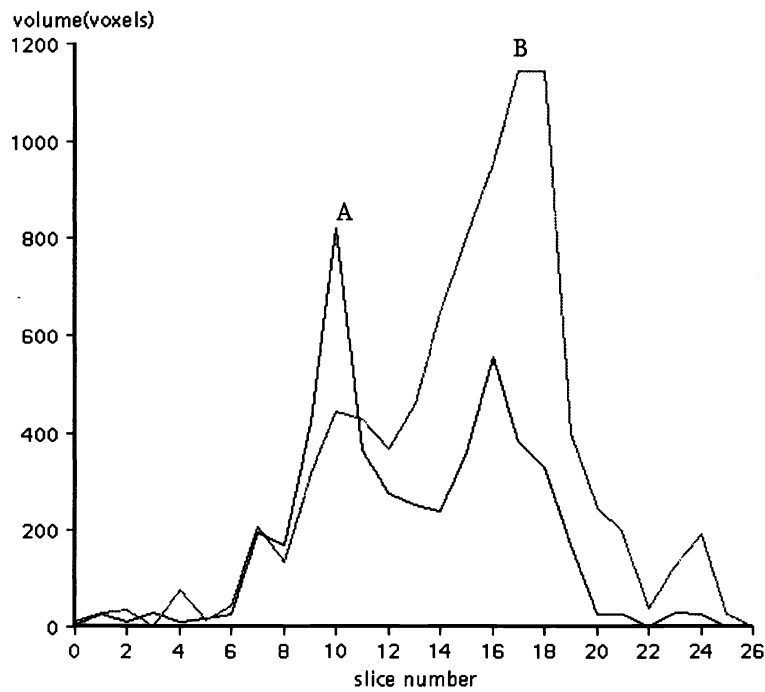


Figure 8: Graphs of lesion volumes of final segmentation and manual segmentation, expanded from Figure 7. (A) = T1/T2 combined and post-processed, (B) = manual segmentation.

least not easily accomplished) by manual outlining of regions. The second factor is usually due to a tendency by clinicians to consistently classify an area as lesion when there is doubt as to the region's composition. The third factor, errors in the segmentation, result from inconsistencies in the histograms and neighbourhood interaction parameters as discussed previously. This will be addressed in ongoing work to be completed in this area.

4. CONCLUSION

The problem of accurate segmentation of MRI and other medical images is a complex, as yet unsolved problem in medical image analysis. We have presented a reasonable, model-based solution using neighbourhood and histogram analysis. In so doing, an attempt was made to build a model of the image on a case-by-case basis, and use this model as a foundation on which to refine the segmentation. We have developed this algorithm in 3D using a dataflow-based visual programming environment. We have maintained a local neighbourhood analysis scheme throughout the algorithm in order to preserve the potential for parallel computation.

5. ACKNOWLEDGEMENTS

This research has been partially funded by the Natural Sciences and Engineering Research Council of Canada. The authors wish to thank the University of British Columbia MS/MRI Group for supplying the MRI data on which the algorithm was tested.

6. REFERENCES

1. T. Arden and J. Poon. *WIT Users Guide*. Logical Vision Ltd., Burnaby, B.C., Canada, 1993.
2. J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 48(3):259–302, 1986.
3. J.A. Broekhuijsen, S.C. Becker, and W. A. Barrett. Probabilistic segmentation of myocardial tissue by deterministic relaxation. In *Computers in Cardiology*, pages 99–102, Long Beach, California, 1989. IEEE Computer Society.
4. H.S. Choi, D.R. Haynor, and Y. Kim. Partial volume tissue classification of multichannel magnetic resonance images—a mixel model. *IEEE Transactions on Medical Imaging*, 10(3):395–407, September 1991.
5. S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, November 1984.
6. B. Johnston, M.S. Atkins, and K.S. Booth. Three-dimensional partial volume segmentation of multispectral magnetic resonance images using stochastic relaxation. *Proceedings of the IS&T/SPIE Symposium on Electronic Imaging Science and Technology*, volume 2180, February 1994.
7. B. Johnston. Three Dimensional Multispectral Stochastic Image Segmentation. Master's thesis, The University of British Columbia, Computer Science Department, Vancouver, B.C., January 1994.
8. N. Karssemeijer. Three dimensional stochastic organ models for segmentation in CT scans. In *Proceedings of the SPIE: Biostereometrics '88—Fifth international meeting*, volume 1030, pages 177–184, Chapel Hill, N.J., 1988.
9. K.R. Smith and L.A. Kendrick. Bayesian computer vision methods for improved tumor localization and delineation. In *Proceedings of the IEEE Medical Imaging Conference*, pages 2140–2144, Santa Fe, 1991.
10. R.R. Stringham, W.A. Barrett, and D.C. Taylor. Probabilistic segmentation using edge detection and region growing. In *Proceedings of the SPIE: Visualization in Biomedical Computing 1992*, volume 1808, pages 40–51, Chapel Hill, N.J., 1992.