

Intensity-Based Object Extraction from 3D Medical Images Including a Correction for Partial Volume Errors

S.A. Röll, A.C.F. Colchester, P.E. Summers, L.D. Griffin
Dept. of Neurology, UMDS, Guy's Hospital, London SE1 9RT, U.K.
s.roell@umds.ac.uk

Abstract

Feature extraction by applying a threshold window to image intensity values is a simple and common image processing technique. We consider the case of 3D images where intensity based feature extraction is used to determine the volume of objects of interest. We show that the accuracy of volume determination is limited by partial volume (PV) effects. We outline a new method for correcting for PV effects based on object geometry and object intensity. Although this PV correction has been developed with respect to a specific application in magnetic resonance imaging, it is applicable to volume extraction by thresholding in any image and can possibly be extended to other intensity-based extraction techniques.

1 Introduction

Intensity thresholding techniques, such as histogram analysis or intensity-based region growing, are frequently used for segmentation of arbitrarily shaped objects from a data volume such as a 3D magnetic resonance (MR) image. At the boundary of an object, part of the volume of each voxel is occupied by the object and part by a background. This creates a layer of voxels with intensities intermediate between the object and background values. We refer to this intensity averaging due to image sampling as the partial volume (PV) effect. Thus, for any sampled image, the estimated *volume* of an object depends on the intensity threshold used for extracting the object. This is the case even if pure object and pure background intensities are well separated.

We reason that it is possible to apply a systematic correction to control the pollution of the volume calculation induced by PV effects. Our correction to the volume is based on the *surface area* of the thresholded object.

This work emerges from the problem of segmenting multiple sclerosis (MS) lesions in MR brain images as accurately as possible. Lesion volume is a widely used measure of disease burden, and its determination with maximum precision is desirable. For this application, intensity-based segmentation techniques have potential advantages over edge-based techniques, as they are less dependent on

operator interaction and are much faster when multiple small lesions need to be identified. However, they have been shown to underestimate lesion volume by 10%-20% [1],[2],[3]. The fact that lesions are typically small relative to voxel size (10-500 voxels), and thus have a high percentage of voxels at their boundary, suggests that PV effects account for much of this error.

With this application in mind, we will use the term 'lesion' equivalent to 'an object of interest in an 3D MR image'.

2 Related Work

Soltanian-Zadeh et al. [4] presented a feature space analysis method for MR images which includes a correction for PV effects. Their approach is aimed at obtaining maximum image information from two or more of registered MR images showing different contrast. Accordingly, two or more intensity values are available for each image voxel. A statistical analysis of partial volume tissue classification, relying again on vector valued images, has been suggested by Choi et al. [5]; in contrast, we seek to obtain maximum information from a scalar valued image.

More recently Santiago and Gage [6] have shown that a model of PV effects can also be incorporated in the statistical analysis of scalar valued images. However, in contrast to the correction presented here, their method dispenses completely with the spatial information of the image.

Koenderink and coworkers have developed a rigorous mathematical treatment of object shape at multiple resolutions (see e.g. [7]). They show how image resolution is responsible for changes of object shape ('blurring'). Although our method extracts sub-voxel information from the image, it does not deal with reversing the effect of blurring. Hence, we appreciate Koenderink's concept of scale space as it shows that we have to expect limitations on the accuracy of our correction.

3 Theory and Computer Simulation

Our aim is to measure the volume of an arbitrarily shaped lesion, which can be identified as an object because at least its interior voxels have intensity values different from the background. We use the term 'interior' voxel for a voxel completely filled by object intensity. A minimum size of an object to be considered is given by the requirement that there is at least one interior voxel. For the sake of simplicity we assume that the object background is either of higher or of lower intensity than the object (homogeneous background). We further assume that the data values of the volumetric MR image represent intensity values which are averaged over a voxel (of any shape). Finally, we only encounter PV effects as introduced above if lesions possess well-defined boundaries, i.e. if transition regions between lesion and background are small compared to voxel size. We introduce the term *geometric* PV first and consider then how to correct PV with respect to the local *intensity* distribution in the image.

3.1 Geometric Partial Volume

We define geometric partial volume V_{gPV} as the true volume V_{true} of an object minus the volume of its interior voxels. We expect V_{gPV} to be proportional to the number of boundary voxels, or, in the same way, to the surface area of the object. In initial experiments we found the latter measure to be more accurate.

In order to test the hypothesized dependence of V_{gPV} on surface area, we consider various synthetic objects created by a thresholded distributions of Coulomb charges (fig. 1a). The true volumes of these objects are accessible with arbitrary accuracy. Image sampling is simulated by placing the objects in a grid made up by cuboidal voxels. V_{gPV} can be determined for each object for the given grid resolution.

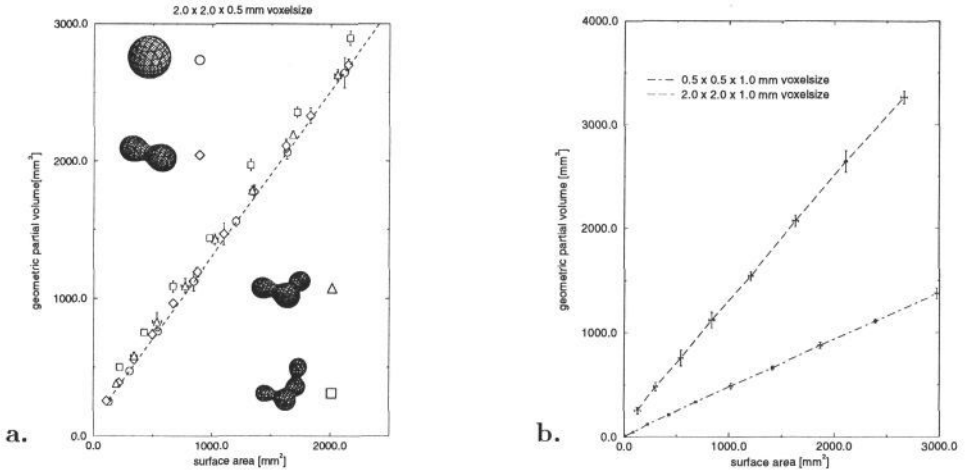


Figure 1: Geometric PV versus surface area (obtained by surface reconstruction).

a. Four different objects of various sizes, obtained by thresholding from one up to four Coulomb charges, at a certain voxel size. The dashed lines indicates the result of linear regression for spheres.

b. Spheres with diameters ranging from 3 to 33 mm at different voxel sizes.

Data points in a. and b. are averaged over translations of the object origins by ten increments along the voxel diagonal line. Error bars represent maximum deviation of the mean due to this “phase” effect.

Surface areas of objects are computed by a modified ‘Marching Cubes’ algorithm [8],[9]. This local method of surface reconstruction is applicable to our synthetic objects placed in grids as well as to real objects in MR datasets, yielding triangular surfaces of the same quality in both cases. Connected groups of internal voxels in the MR image (or grid points of the charge distributions) are covered by the same surface. In contrast to surfaces constructed by the original Marching Cubes algorithm [8], surfaces constructed according to [9] are guaranteed to be leakless.

As figure 1a shows, V_{gPV} increases linearly with surface area and is essentially independent of object shape. Therefore, knowledge of the surface area of an arbitrary object sampled at a known voxel size allows us to find its geometric partial

volume V_{gPV} , by looking up the geometric partial volume of a sphere of the same surface area sampled at the same voxel size. The result of linear regression on V_{gPV} of spheres is given by the dashed line in figure 1a.

In figure 1b, geometric partial volume is plotted for spheres sampled at two different voxel sizes. As expected, V_{gPV} increases with voxel size.

3.2 Intensity-based Partial Volume Correction

Geometric PV V_{gPV} represents the upper boundary of PV losses an object can suffer. For a volume V_{thresh} obtained by intensity thresholding, volume underestimation due to neglecting partially filled voxels is always less severe than it is predicted by geometric PV ($V_{true} - V_{thresh} < V_{gPV}$), since some almost completely filled voxels are taken into account for V_{thresh} . Moreover, if the number of partially filled voxels taken into account for V_{thresh} is large enough, V_{thresh} may even overestimate true lesion volume V_{true} , since the partially filled voxels contribute to V_{thresh} with their complete volume. To eliminate the effect of thresholding, we desire a volume correction V_{corr} so that ideally $V_{true} = V_{thresh} + V_{corr}$, or that at least the corrected V_{thresh} lies closer to V_{true} .

We assume that interior lesion voxels may possess a range of intensities within the interval $[I_{min}, I_{max}]$. By considering only these interior voxels, an intensity probability distribution $p(I_L)$ of lesion intensities I_L not affected by PV averaging is computed for each lesion. In the case of a bright lesion surrounded by a lower mean background intensity $\bar{I}_{bg} < I_{min}$, the lower intensity threshold I_{min} is relevant with respect to the boundary definition of the lesion. For a lesion intensity I_L , $I_{min} < I_L < I_{max}$, at a threshold I_{min} , boundary voxels which are partially filled to a threshold fraction of

$$t = \frac{(I_{min} - \bar{I}_{bg})}{(I_L - \bar{I}_{bg})} \quad (1)$$

and greater are regarded as lesion voxels. Their complete volume is contained by V_{thresh} .

The further treatment of the problem would be straightforward if we assume all degrees of partially filling to occur with the same probability. This is the case if cuboidal objects are randomly sampled by aligned cuboidal voxels. However, we found it necessary to consider a more realistic probability distribution $p_{fill}(x)$ of degrees of fillings x , as e.g. p_{fill} for a sphere sampled by cuboidal voxels (see fig. 2).

To understand qualitatively how such non-constant distributions can arise, we may consider a sphere sampled by spherical voxels. This case is simpler than the case of cuboidal voxels, since the distance d between sampling and sampled sphere is the only parameter determining the degree of partial filling x . Due to the curvature of the sampled sphere the relationship between d and x is not linear. An interval of degree of filling at one distance $|x(d_1) - x(d_1 + \Delta)|$ is generally not equal to the interval $|x(d_2) - x(d_2 + \Delta)|$ at another distance. Since due to random sampling all distances d are equally probable, not all degrees of filling x are equally probably. We found it sufficiently accurate to determine the probability of degrees of filling by placing spheres randomly in a cuboidal grid and counting degrees

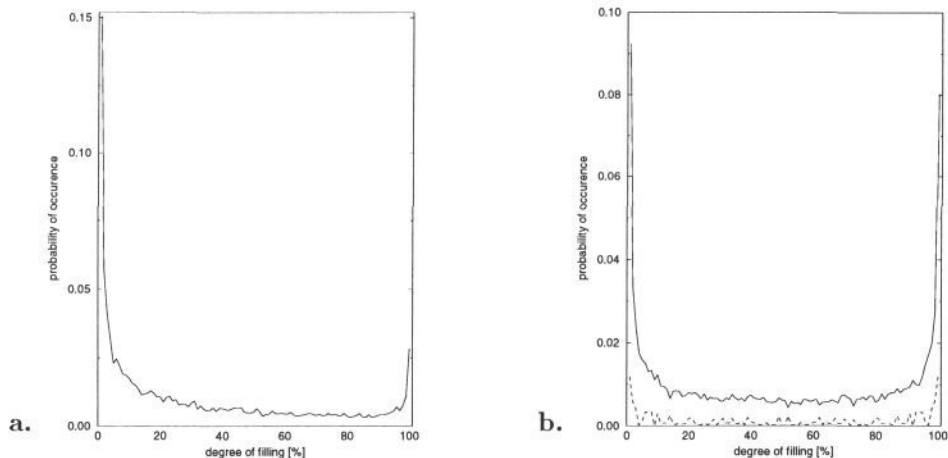


Figure 2: Probability distribution of degrees of filling of cuboidal voxels of size $1.0 \times 1.0 \times 1.0$ mm by spheres. In a., a sphere of 3 mm diameter yields an asymmetric distribution, whereas in b., a sphere 33 mm diameter yields an almost symmetric distribution. The amount of the deviation from symmetry is indicated by the dashed line. The wiggles in a. and b. are due to the fact that the graph was obtained by a Monte-Carlo simulation. For each of the graphs 10^5 partially filled voxels was examined.

of filling. This Monte-Carlo method avoids problems connected with deriving these distributions analytically. We emphasize that only at this point does the additional assumption of cuboidal voxel shape enter our model. This assumption is appropriate for the case of MR images.

Figure 2a shows that for spheres which are small compared to voxel size, i.e., spheres with a relatively large curvature, less filled voxels are more likely than almost filled ones. For large spheres (2b), the distribution becomes almost symmetrical, with maxima for the lowest and highest degrees of filling. Comparing distributions for spheres to those for the other synthetic objects shown in figure 1 of the same surface areas did not yield a significant difference. Hence, we approximate the statistic of partially filled voxels p_{fill} of an arbitrary object by the statistic of a sphere of the same surface area.

Geometric PV V_{gPV} is assumed to add up according to p_{fill} . The contribution of each degree of filling x to V_{gPV} is given by $f(x) = c_1 x p_{fill}(x)$, with the constant c_1 chosen so that $\int_{0+}^{1-} f(x) dx = V_{gPV}$. In order to find the partial volume overestimated by counting almost filled voxels as full we also need the complementary function $\tilde{f}(x) = c_2 (1-x) p_{fill}(x)$, expressing how x contributes to a V_{gPV} made up by the *not* filled voxel volumes. Again, c_2 is determined by the requirement $\int_{0+}^{1-} \tilde{f}(x) dx = V_{gPV}$. PV correction for a lesion intensity I_L is obtained by summing up PV according to $f(x)$ for $x = 0+$ up to the threshold $x = t$ (equation 1), and subtracting overestimated PV according to $\tilde{f}(x)$ for $x = t$ up to $x = 1-$.

The following equation summarizes partial volume correction:

$$V_{corr} = \sum_{I_L=I_{min}}^{I_L} p_I(I_L) \cdot \left(c_1 \sum_{i=0}^{t_{bin}(I_L)} x(i)p_{fill}(i) - c_2 \sum_{i=t_{bin}(I_L)}^{100} (1-x(i))p_{fill}(i) \right). \quad (2)$$

Lesion intensities are treated as integer values and degrees of partial filling and the statistic of filling are partitioned into 100 bins denoted by $x(i)$ and $p_{fill}(i)$, respectively. The threshold degree of partial filling is denoted by the threshold bin t_{bin} .

At an optimal threshold t_{opt} , the second term correcting for overestimated PV cancels the first term for underestimated PV. We use the term ‘optimal’ threshold since at this point the correction crosses zero and the volume V_{thresh} obtained by thresholding can be expected to lie close to V_{true} . The value of t_{opt} depends on the shape of the p_{fill} . However, for a symmetric distribution of partial filling $p_{fill}(x) = p_{fill}(1-x)$, the optimal threshold is always $t_{opt} = 0.5$, i.e. midway between lesion and background intensity, which matches ones intuitive expectation. For the distributions we have encountered so far (fig. 2) we found t_{opt} in the range of $0.45 < t_{opt} < 0.5$.

For each lesion the ‘scaled relevant threshold’

$$I^* = \frac{(I_{min} - \bar{I}_{bg})}{(\bar{I}_L - \bar{I}_{bg})}, \quad (3)$$

$0 < I^* < 1$, obtained by replacing the lesion intensity I_L in equation (1) by its mean value \bar{I}_L , indicates the average cut-off point of degree of filling, above which intensity is considered as lesion intensity and below which it is not. I^* depends essentially on the selection of the with respect to the lesion background relevant intensity threshold, in our case I_{min} . Accordingly, for $I^* > t_{opt}$, V_{thresh} is expected to underestimate V_{true} , for $I^* = t_{opt}$, V_{thresh} will be close to V_{true} , and for $I^* < t_{opt}$, V_{thresh} is expected to overestimate V_{true} .

An unavoidable circularity of our method turns out to be not severe. When a correction is applied, we do not know the surface area of a lesion at the optimal threshold $I^* = t_{opt}$, but only the surface area of the thresholded lesion. This uncorrected value of the surface area is used to estimate V_{gPV} (which in turn determines c_1 and c_2 in equation 2). However, generally surface area increases less than linearly with increasing lesion volume. Applying our method to MR images, we found lesion surface area to be much less sensitive against threshold variations than lesion volume.

4 Verification by Phantom Studies

Phantom lesions of various shapes and volumes ranging from 0.02 - 3.0 cm³ made from paraffin wax were embedded into an agarose gel. By measuring the density of the wax, their true volumes V_{true} could be derived from their weight within

a confidence interval of 2%. Several 3D- and multislice imaging experiments, performed on a Philip ACS/II 1.5 T scanner, yielded data sets of voxel sizes between $0.5 \times 0.5 \times 1.0$ and $2.0 \times 2.0 \times 2.0$ mm. Fig. 3a shows a slice of one of these data volumes. The fact that the magnetic resonant signal of paraffin is much weaker than the signal of the surrounding gel, leads to images of inverse contrast. This inverse case can be treated as completely analogous to the case discussed so far. E.g., in this case I^* is defined by the upper lesion defining threshold I_{max} ,

$$I^* = \frac{(\bar{I}_{bg} - I_{max})}{(\bar{I}_L - \bar{I}_{bg})}.$$

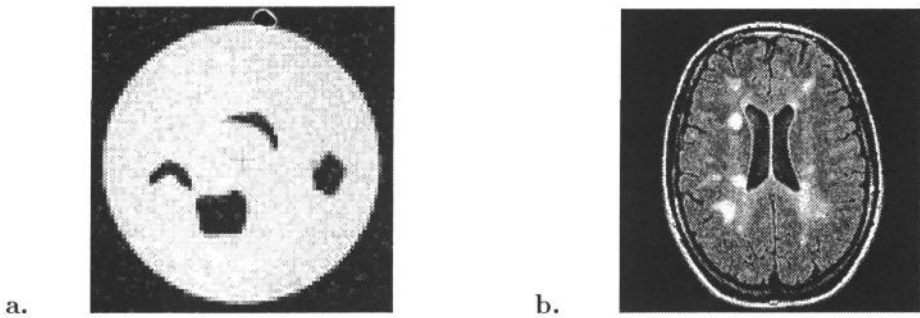


Figure 3: a. Paraffin phantom lesions appear as dark spots in a Spin Echo MR image. The voxel size is $1.0 \times 1.0 \times 1.0$ mm. b. A slice of a FLuid Attenuated Inversion Recovery (FLAIR) MR image of a brain shows MS lesions as bright spots. The voxel size is $0.9 \times 0.9 \times 2.0$ mm.

In figure 4, the relative deviation ϵ of the uncorrected volume V_{thresh} and the corrected volume $V_{thresh} + V_{corr}$ from the true volume V_{true} are plotted versus the scaled relevant threshold I^* .

For a lesion of 9 mm linear size at a voxel size of $2 \times 2 \times 1$ mm (fig. 4a), a scaled threshold in the range of $0.74 > I^* > 0.38$ leads to an error range of uncorrected volumes of $-36.2 < \epsilon_{uncorr} < 19.1$ percent. For a lesion of this size at the given voxel size, our correction crosses zero at scaled threshold $I^* = 0.46$ (note that this root of the correction is given in each of the graphs of fig. 4 by the intersection point of the corrected and uncorrected error). The correction keeps the volume at an almost constant value. However, it overestimates the ground truth V_{true} by about 7%. This error is likely to be caused to a large part by an underestimation of the true volume of the phantom. Therefore, for range $0.55 > I^* > 0.45$, the uncorrected volume lies closer to V_{true} than the corrected volume. We have included this result as an example of less than average performance of our method.

In figure 4b, a result above average is shown. Over the full realized range of I^* apart from the zero crossing of the correction at $I^* = 0.48$, the corrected volume approximates the known volume better than the uncorrected volume V_{thresh} . The slope of V_{thresh} is effectively suppressed.

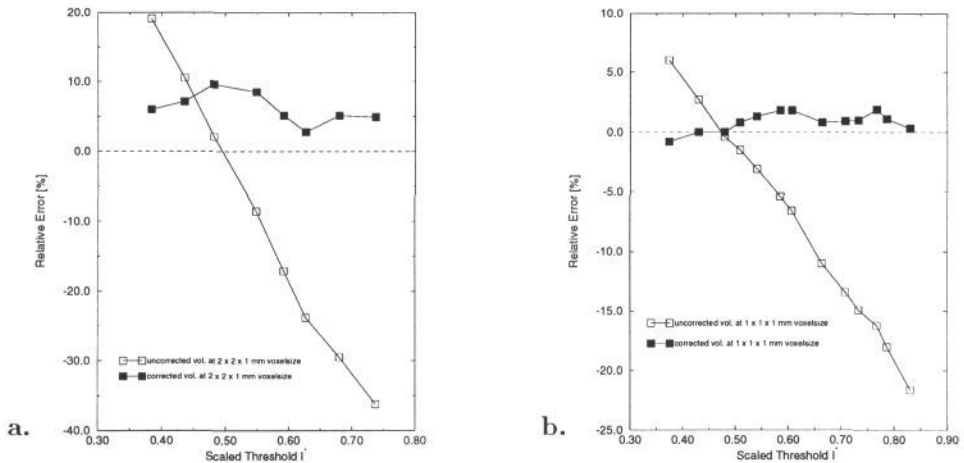


Figure 4: Relative errors of volume measurements without PV corrections $\epsilon_{uncorr} = |V_{thresh} - V_{true}|/V_{true}$ (\square), and with PV corrections $\epsilon_{corr} = |V_{thresh} + V_{corr} - V_{true}|/V_{true}$ (\blacksquare).

a. A phantom lesion of $V_{true} = 420 \text{ mm}^3$ ($\approx 9 \text{ mm}$ linear size) at a voxel size of $2 \times 2 \times 1 \text{ mm}$.

b. A phantom lesion of $V_{true} = 968 \text{ mm}^3$ ($\approx 12 \text{ mm}$ linear size) at a voxel size of $1 \times 1 \times 1 \text{ mm}$.

The following statements summarize our results:

a) At $I^* = t_{opt}$, our correction crosses zero and volumes obtained by thresholding agree within their error bounds with the known volumes V_{true} .

b) As expected, PV effects become significant for $I^* < 0.4$ or $I^* > 0.6$ and their relative importance increases with increasing voxel size and decreasing lesion size.

c) Except for the interval $0.4 < I^* < 0.6$, relative errors of corrected volumes were smaller by at least a factor of two compared to relative errors of uncorrected volumes obtained by thresholding.

d) Within the interval $0.4 < I^* < 0.55$, uncorrected volumes are occasionally closer to the true volume than corrected volumes, if V_{true} deviates from the measured volume at $I^* = t_{opt}$. However, these differences are rarely more than 5.0% and our correction can safely be applied over the full range of I^* .

e) Compared to uncorrected volumes, corrected volumes are essentially constant over the entire range of I^* . This fact shows that the circularity of using uncorrected values for surface area does not have a major impact on volume correction. For our phantoms, we found surface area to increase slightly with decreasing I^* , and to vary in some cases unsystematically with I^* . These variations of surface area are responsible for the remaining variations of the corrected volumes.

5 Application to Multiple Sclerosis Lesions

Whereas for our phantom images it is possible to obtain results of volume measurements for large range of intensity thresholds ($0.3 < I^* < 0.9$), this is not the case for MR images of the brain with enhanced contrast for MS lesions. Due to a higher fraction of noise in the image and the adjacency of other objects to MS lesions in image space as well as in intensity space, only a small band of 'possible' threshold intensities, varying for each lesion in the image, is applicable in order to obtain good intra-lesion connectivity and a sensible boundary definition. For the scan of a brain shown in figure 3b, we found these possible thresholds $I_{possible}^*$ to deviate for various lesions substantially from the optimal thresholds $I^* = t_{opt}$ ($0.35 < I_{possible}^* < 0.7$, whereas $0.45 < t_{opt} < 0.5$). In these cases, volume correction is necessary. Relative differences between corrected and uncorrected volumes ranged up to 20%.

6 Discussion and Conclusion

We propose a systematic correction method for PV errors which arise when using intensity-based methods for estimating volumes in MR images. No adjustable parameters are required. We have shown that the correction improves accuracy and consistency of volume measurements in MR phantom and that it is applicable to MR brain data.

An alternative method of dealing with volume quantitation in sampled images would be the following. For a volume V containing the lesion as well as some background, the background intensity I_{bg} divided by the average voxel intensity I_{ave} measures lesion volume: $V_{lesion} = ((I_{ave} - I_{bg})/I_{ave})V$. This method circumvents geometrical considerations elegantly by taking intensities averaged by PV effects into account as such. As a potential disadvantage, its accuracy depends directly on the estimation of the background intensity. This is not the case for the method presented here, where the magnitude of PV averaging is additionally determined by the surface area of the defined object. Furthermore, the purely intensity-based method does not define objects and hence does not allow visual control of what parts of the volume are counted as lesion volume. In future work we will compare the two approaches in detail and evaluate the accuracies achieved.

So far our correction is implemented only for simple thresholding in a one-dimensional histogram. The concept can be extended to other intensity-based segmentation methods; its implementation for intensity-based region growing segmentation is under way.

References

- [1] E.F. Jackson, P.A. Narayana, J.S. Wolinsky, and T.J. Doyle, Accuracy and Reproducibility in Volumetric Analysis of Multiple Sclerosis Lesions, *J. Computer Assisted Tomography*, 17(2), 200-5, 1993

- [2] M. Kamber, D.L. Collins, R. Shinghal, G.S. Francis, and A.C. Evans, Model-based 3D Segmentation of Multiple Sclerosis Lesions in Dual-Echo MRI Data, its Visualization in Biomedical Computing, SPIE 1808, 590-99, 1992
- [3] D.A.G. Wicks, P.S. Tofts, D.H. Miller, G.H. du Boulay, A. Feinstein, R.P. Sacares, I. Harvey, and R. Brenner, Volume Measurement of Multiple Sclerosis Lesions with Magnetic Resonance Images, *Neuroradiology*, 34, 475-79, 1992
- [4] H. Soltanian-Zadeh, J.P. Windham, and A.E. Yagle, Optimal Transformation for Correcting Partial Volume Averaging Effects in Magnetic Resonance Imaging, *IEEE Trans. Nuclear Science*, 40(4), 1204-12, 1993
- [5] H.S. Choi, D.R. Haynor, and Y. Kim, Partial Volume Tissue Classification of Multichannel Magnetic Resonance Images - A Mixel Model, *IEEE Trans. Medical Imaging*, 10(3), 395-407, 1991
- [6] P. Santago and H.D. Gage, Quantification of MR Brain Images by Mixture Density and Partial Volume Modeling, *IEEE Trans. Medical Imaging*, 12(3), 566-74, 1993
- [7] J.J. Koenderink and A.J. van Doorn, Dynamic Shape, *Biological Cybernetics*, 53, 383-96, 1986
- [8] W.E. Lorensen and H.E. Cline, Marching Cubes: A High Resolution 3D Surface Construction Algorithm, *Computer Graphics* 21(4), 163-70, 1987
- [9] S. Röhl, Segmentieren und Visualisieren dreidimensionaler MR-Bilder, *Diplomarbeit*, Physikalisches Institut der Universität Würzburg, 1993.