Robust retrieval of three-dimensional structures from image stacks

Maria Garza-Jinich^{1*}, Peter Meer² and Veronica Medina³

¹Instituto de Investigaciones en Matematicas Aplicadas y en Sistemas, Universidad Nacional Autonoma de Mexico, Apdo. 20-726, Mexico DF 01000

²Electrical and Computer Engineering Department, Rutgers University, Piscataway, NJ, 08855-0909, USA

³Departamento de Ingenieria Electrica, Universidad Autonoma Metropolitana, Av. Michoacan y La Purisima, Iztapalapa, Mexico DF 09340

Abstract

Robust high-breakdown-point location estimators are employed to analyze image stacks under the piecewise constant image structure model. To reduce the effect of bias along the Z-axis, the class parameters are extracted using three consecutive slices. The segmentation algorithm first determines the most reliable seed regions, which are then used in a region-growing procedure supported by local evidence. The robustness and stability of the proposed technique is shown with both synthetic and real data, the latter consisting of one MRI and one confocal microscopy set. The performance of the algorithm is consistent with the ground truth obtained with manual segmentation by physicians.

Keywords: confocal microscope images, image segmentation, magnetic resonance images, robust statistics

Received November 25, 1996; revised June 10, 1997; accepted July 14, 1998

1. INTRODUCTION

The sequences of two-dimensional (2-D) cross-sectional slices produced by X-ray computed tomography and magnetic resonance (MR) provide information about anatomically significant features which can be used to generate twoand three-dimensional (3-D) renditions of internal structures (Herman and Liu, 1979; Robb and Barillot, 1989; Tiede *et al.*, 1990). Through the years, the possible use of medical imaging as a powerful quantitative tool in surgery and diagnosis has become evident (Cho *et al.*, 1993). Many surgical procedures require highly precise 3-D location of a target tissue, and reliable volumetric analysis techniques are needed to evaluate the extent of diseases and monitor treatment. A full description of the potential of an end-to-end system for 3-D image-guided surgery can be found in Grimson (1995).

The segmentation of medical images into 3-D anatomically significant structures is essential for further registration, atlas mapping, volume quantization and 3-D spatial location, and has always been a challenging task. The bottleneck is the lack of highly accurate and reproducible segmentation methods for isolating 3-D structures from the stack (Udupa, 1991), and manual slice editing is still performed in most clinical applications (McInerney and Terzopoulos, 1996). MR images are well suited for automatic analysis because they are produced with unique contrast and high spatial resolution. However, processing MR images is not straightforward due to several factors: the large amount of data contained in a set; the unknown number of tissues at different levels of the stack: and the noise produced by the scanner (Smith and Ranallo, 1989), which may vary from set to set, from image to image, and even within an image (Wells et al., 1996). The partial volume effect, anatomic structures smaller in size than the voxels, can also produce artifacts. All the above-mentioned problems should be taken into account when designing an automatic segmentation algorithm.

Edge/surface detection methods are often proposed in the literature (Kennedy et al., 1989; Bomans et al., 1990; Deklerck et al., 1993), although they all encounter the problem of how to obtain closed and connected boundaries (Udupa, 1994). Graph theory (Raman et al., 1993; Wu and Leahy, 1993; Robinson et al., 1994) and differential geometry (Sander and Zucker, 1990) were employed to correctly select and delineate the significant boundaries. Active contours and surfaces have been adapted to the extraction of anatomical structures (Davatzikos and Prince, 1995; Snell et al., 1995; Székely et al., 1996) yielding promising results, although they are time consuming and their output is highly dependent on the accuracy of the initial contour or surface model; see McInerney and Terzopoulos (1996), for a complete description of the development and application of deformable models in medical image analysis,

A second approach is region/volume delineation, where direct segmentation of 3-D objects is achieved through the extension of conventional region-growing techniques to volume growing, supported by statistical modeling of the image content (Liou and Jain, 1993). These methods often assume that the noise in the images is ergodic and uncorrelated with the signal (Soltanian-Zadeh *et al.*, 1992). To obtain accurate results with such techniques it is necessary that three basic procedures: estimation of the number of classes, estimation of their parameters and the labeling (classification) of the pixels are reliably performed. Although a general solution requiring the total automation of each step is desirable, it is not an easy task and most authors propose algorithms that rely either on user interaction or on *a priori* knowledge.

Automatic estimation of the number of classes is rarely done (Liang *et al.*, 1994), and usually this number is supplied by the user. In medical image analysis, the number of classes depends on the application, and most systems use a predefined number of classes (Cline *et al.*, 1990; Li *et al.*, 1993; Arata *et al.*, 1995; Tsai *et al.*, 1995; Johnston *et al.*, 1996). While in some cases only soft-tissue and non-softtissue discrimination is needed (Brummer *et al.*, 1993), in others discrimination between different soft tissues is also required (Joliot and Mazoyer, 1993; Taxt and Lundervold, 1994; Cheng *et al.*, 1996).

Automatic estimation of class parameters is done more frequently, although intervention of the user may still be required. Estimation of the parameters from the graylevel histogram of a region of interest (ROI) (Cline *et al.*, 1990; Joliot and Mazoyer, 1993; Johnston *et al.*, 1996), or by supervised clustering (Taxt and Lundervold, 1994), will inevitably depend on the accuracy of the user-supplied data. Employing the mixture of Gaussians model for gray-level histogram analysis (Brummer *et al.*, 1993; Tsai *et al.*, 1995) relies on the existence of well-defined peaks and valleys in the histograms, which cannot be guaranteed in the general case. Clustering algorithms such as K-means (Arata *et al.*, 1995) or iterative pixel classification such as relaxation or using Markov random fields (MRF) (Cheng *et al.*, 1996), require good initialization in order to converge to a satisfactory solution.

The labeling procedure should take into account local spatial information to remove isolated pixels or small regions. This is usually achieved by using mathematical morphology operations (Brummer *et al.*, 1993; Tsai *et al.*, 1995; Kapur *et al.*, 1996), filtering (Cline *et al.*, 1990) or heuristic local analysis (Joliot and Mazoyer, 1993).

Sequential processing, i.e. executing all three procedures sequentially for one class at a time (Brummer *et al.*, 1993; Arata *et al.*, 1995; Tsai *et al.*, 1995) can introduce artifacts since the hard decision thresholds do not take into account the overlap between the gray-level value distributions of the neighboring classes. Very few proposed methods (Li *et al.*, 1993; Taxt and Lundervold, 1994) consider and solve this issue. Labeling produced by MRF implementations (Liang *et al.*, 1994; Johnston *et al.*, 1996) gives better results, although the algorithms are time consuming.

Regardless of the technique chosen, satisfactory image segmentation, like most low-level computer vision tasks, cannot be achieved by an exclusively bottom-up (datadriven) approach. In a practical segmentation method several trade-offs must be satisfied. The most important one is between over- and under-segmentation. In oversegmentation, the amount of erroneously delineated regions is kept low at the expense of the image being broken into too many parts. Over-segmentation is recommended whenever sufficient a priori information is available, and the parts can be assembled together with a top-down process. This is often the case with MR images, since the anatomical constraints are well understood. The dual of over-segmentation is under-segmentation, where the number of delineated regions is kept low at the expense of erroneous fusions. It is a frequent artifact of techniques which estimate the decision thresholds exclusively from global statistics derived from the data.

The segmentation of a sequence of 2-D images in the search for 3-D structures raises another trade-off. In the sliceby-slice approach, the 2-D slices are analyzed individually and the 3-D structure recovered from these results. Discontinuities of the volumes between slices, however, are highly probable and should be accounted for. On the other hand, the volume-oriented approach is based on sets of consecutive slices, from where the 3-D elements are delineated. The trade-off is between the amount of computation and the use at the earliest stages of the processing of all the available information. A final trade-off is between choosing a local or a global solution. A purely local approach may lead to the delineation of insignificant features from a global point of view. On the other hand, a purely global solution may lead to loss of detail. This, local-global duality, appears under different guises in almost every low-level image understanding task.

The 3-D structure extraction method proposed in this paper was designed to optimize all the above-mentioned issues.

- The bottom-up stage uses robust image analysis techniques (Meer *et al.*, 1991), and the number of classes, the parameters of the classes and the labels of the pixels are produced automatically.
- The quality of segmentation (under/over) is controlled only by context-independent thresholds, resulting in a weak dependence of the performance on user-set parameters.
- The volume-oriented approach employed is restricted to local 3-D neighborhoods defined by three consecutive slices, and it is able to deal with a possible bias in the *Z*-axis due to the equipment.
- The parallel region-growing algorithm takes into account that the gray-level distribution of different tissues can overlap, starts with the highest confidence pixels and solves ambiguous cases with the aid of local information.
- The method is highly modular, having three main processing steps: (i) robust estimation of the parameters of the significant 3-D regions; (ii) computation of significant classes; (iii) volume segmentation by 3-D region growing, which are completely separated and allow integration of *a priori* information (if available) or performance improvement through consensus (if increased computational burden is acceptable).

The paper is organized as follows. In Section 2 the fundamental computational module, the robust location and scale estimation is analyzed. In Section 3 the segmentation algorithm is presented using a synthetic set of images as an example, and validated with numerical data. Experimental results with one set of MR images, an application to confocal microscopy, as well as comparison with manual segmentations are shown in Section 4. Conclusions are included in Section 5, where the method is discussed, and its potential and limitations are presented.

2. ROBUST LOCATION AND SCALE ESTIMATION

Two estimates are of interest when segmenting a gray-level image into constant patches. For convenience assume the data, i.e. the pixel values, are accessed by a single index. The data are modeled as additively corrupted with zero-mean identically and independently distributed (IID) noise ϵ_i , $x_i = \theta + \epsilon_i$; i = 1, ..., n; where x_i are integers between 0 and (say) 255. Using terminology from statistics, the value of the constant characterizing the patch is obtained with a (one-dimensional, 1-D) location estimator $\hat{\theta}$, while the spread of the values from this constant by a (1-D) scale estimator \hat{s} . The residuals are defined as $r_i = x_i - \hat{\theta}$; i = 1, ..., n; and the value of the estimate is obtained by minimizing a function of them. The traditional, least-squares method, uses $\min_{\hat{\theta}_{LS}} r_i^2$ and the location estimator is the mean

$$\hat{\theta}_{\rm LS} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

while the scale estimator is the standard deviation

$$\hat{s}_{\text{LS}} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \hat{\theta}_{\text{LS}})^2}{n-1}}.$$
 (2)

Both these estimators are non-robust. Assume that a single data point has an arbitrarily large value disregarding for the moment that the range of x_i is limited. This single outlier corrupts both estimates (1) and (2). The least-squares estimators of location and scale have zero asymptotic breakdown point [see Rousseeuw and Leroy (1987), Subsection 1.2, for a rigorous definition of the breakdown point of an estimator].

A robust location estimator is the mode. For a continuous probability density function (PDF) the mode is the most probable value of the variable, i.e. the maximum of the PDF. In practice the PDF is not known and the mode must be estimated from discrete data. Maximum of the PDF means that the values around the mode have a higher probability of occurence and the mode can be detected by locating the highest-density region in the ascendingly ordered sequence of the data points. Often the window used in measuring the local density contains half data points. The mode-detection algorithm can be summarized as follows.

Mode detection algorithm

- 1. Order the data in ascending order, $x_{i:n}$.
- 2. Define a sliding window containing half the data points. Let half the size of the window be

$$\delta_i = (x_{i+h-1} - x_i)/2$$
 $h = [n/2] + 1.$

3. Find i_m such that

$$i_m = \arg\min \delta_i$$
.

This is the location of the densest region, and the mode is computed as

$$\theta_M = (x_{i_m+h-1} + x_{i_m})/2.$$



Figure 1. An example of the failure of the mode detection algorithm: (a) The 1-D test data; (b) histogram of the samples.

Thus, the mode is detected with a shortest-half algorithm [Rousseeuw and Leroy (1987) Subsection 4.2]. If i_m is non-unique, the final value is obtained by averaging the $\hat{\theta}_M$. When x_i is a histogram, with i = 0, ..., 255, the histogram has first to be reshaped into a sequence of length $\sum_{i=0}^{255} x_i$.

It can be shown [Rousseeuw and Leroy (1987) Subsection 4.2] that the mode, computed as above, is the least-median-of-squares location estimate, $\min_{\hat{\theta}_M} \text{ med}_i r_i^2$. The use of squares instead of absolute values assures the uniqueness of the solution if n is an even number. The presence of the median makes the mode a robust high-breakdown estimator. Indeed, if up to half the data points do not belong to the distribution of inliers, the mode is still detected correctly. The least-median-of-squares scale estimator is obtained from the mode detection algorithm as $\hat{s}_M = c\delta_{i_m}$, where the constant c is chosen to obtain consistent results for the assumed PDF of the inliers in the data. For example, if the inliers are normally distributed, c = 1.4826. The data can now be dichotomized into inliers and outliers by comparing $|x_i - \hat{\theta}_M|$ with (say) $2 \times \hat{s}_M$.

The mode detection algorithm, however, may fail when the data is almost bimodal. For example, the noisy piecewise constant data in Figure 1a was generated using two normal distributions: 100 samples from (100, 25^2) and 80 samples from (175, 25^2). The number of inliers thus only slightly exceeds the number of outliers. The mode estimator should return a value ~100. While the separation of the two constant regions can be seen clearly, the histogram (Figure 1b) shows that the input into the mode detection algorithm is not helpful. The highest density region is not well defined and the algorithm returns $\hat{\theta}_M = 118.5$ with a scale estimate

 $\hat{s}_M = 53.37$. A detailed analysis of the causes for the estimator's failure can be found in Mintz *et al.* (1992).

The above example is typical of noisy step discontinuities in images. Thus, in spite of using a robust high-breakdownpoint location estimator, such noisy data cannot be segmented reliably into two homogeneous parts. The performance of the mode detection algorithm can be improved significantly if the spatial structure of the outliers is exploited. Under the piecewise constant model, both the inliers and the outliers are constants corrupted by zero-mean noise. In the data window (which in this example is all the data) without loss of generality, the noise can be assumed IID; that is, the scale estimate (standard deviation) derived from pixels in a homogeneous neighborhood should be similar. The standard deviation computed in small analysis windows, however, increases significantly if the neighborhood contains the discontinuity between the two homogeneous regions. This property, which is valid only for the piecewise constant model (Mintz et al., 1992), can be used to design a twostep mode detection procedure. It should be noted that the reliability of the technique is obtained at the expense of being restricted to noisy piecewise constant data.

2.1. Enhanced mode detection (EMD) procedure

- 1. Tessellate the data by overlapping analysis windows of length w (w an odd number) with centers separated by m locations.
- 2. In every window, compute the mean $\hat{\theta}_{LS,i}$ Equation (1) and standard deviation $\hat{s}_{LS,i}$ Equation (2). These values are associated with the *i*th pixel, the center of the window.
- 3. Form the distribution of all the $\hat{s}_{LS,i}$ values associated to all windows. Find the mode of this distribution. It is the robust scale estimate \hat{s}_{EM} derived from the data. Mark the outliers of the $\hat{s}_{LS,i}$ distribution.
- 4. Remove from the $\hat{\theta}_{LS,i}$ distribution those points which are associated with an outlier in the $\hat{s}_{LS,i}$ distribution. With high probability the discontinuity is in the neighborhood of these points.
- 5. Find the mode of the 'censored' $\hat{\theta}_{LS,i}$ distribution. It is the robust location estimate $\hat{\theta}_{EM}$ derived from the data.

The enhanced mode detection procedure performs significantly better for noisy piecewise constant data than the original mode detection algorithm. The technique can be extended to piecewise polynomial data. If the data is close to unimodal the EMD defaults into the original mode detection algorithm. See Mintz *et al.* (1992) for details. The example



Figure 2. 3-D rendering of the uncorrupted synthetic objects.

in Figure 1a processed with EMD (the 1-D analysis windows contained w = 15 samples, located at every m = 3 samples), yields $\hat{\theta}_{\rm EM} = 92.13$ and $\hat{s}_{\rm EM} = 28.84$, much closer to the correct values.

The application of this technique in MRI is justified by the fact that the assumption of constant intensity plus additive noise within a given anatomical structure is reasonable in regions with low texture (Ayache, 1995). Furthermore, the assumption of slowly varying intensities is a good estimate to model MR data, where a specific intensity level can be assumed to be approximately constant over the image plane if the MR scanner is optimally tuned and if non-linearity artifacts are precorrected (Gerig *et al.*, 1992). In order to take into account discontinuities in the z-direction, the EMD procedure can be extended to consider 3-D data windows. This is suitable for MRI stacks, where hundreds of slices must be segmented.

3. ROBUST ANALYSIS OF IMAGE STACKS

The proposed 3-D segmentation method will be described and validated using a stack of synthetic data. Four geometrical 3-D objects, one solid sphere and three hollow half-spheres in a dark background, were defined with a set containing 100 slices (Figure 2). The 3-D volumes were rendered with the system 3DViewnix developed by the Medical Image Processing Group at the University of Pennsylvania. Both objects and background were corrupted by white, zero-mean, Gaussian noise with the same standard deviation (8.0), and the mean gray value of the distribution of each object was slightly changed across the stack (see Figure 5a). This example was designed to simulate the intensity inhomogeneities between slices produced by an MR scanner. Besides, not all objects appear in all slices, as happens in medical images:

significant anatomical features may extend only over a subset of slices. [We refer the reader to Garza-Jinich *et al.* (1996) for another validation example, where the data contains no inhomogeneities between slices.]

One slice of the noisy stack is shown in Figure 3a. Note that in the gray-level histogram of the slice (Figure 3b) the five classes cannot be distinguished.

3.1. Estimation of robust parameters $\hat{\theta}_j$ and \hat{s}_j

The *j*th delineated region in the *l*th slice of the stack is characterized by an estimated gray-level value $\hat{\theta}_j$ and standard deviation \hat{s}_j , j = 1, ..., J. (The second index *l* is not used for ease of exposition.) To segment the slice, the $\hat{\theta}_j$ and \hat{s}_j corresponding to all *J* significant regions must be determined. Note that *J* is not known *a priori*. An optimal segmentation procedure should balance the use of global information (e.g. histograms derived from the image) and of local information (e.g. spatial constraints).

To exploit the 3-D information contained in the image stack the analysis of the *l*th slice is based on a $3 \times 9 \times 9$ data window defined over three consecutive slices l - 1, l, l + 1. The data windows tessellate the *l*th slice by being shifted 3 pixels along the two cardinal directions. Each data window is analyzed independently with the EMD procedure, employing a small, $3 \times 3 \times 3$ analysis window. The analysis windows are offset with m = 2 along either cardinal direction, and can extend over the boundaries of the data window. Both windows are centered, i.e. the output of any processing is allocated to the pixel in the center.

The size of these windows defines the minimum area a structure should have in order to make the recovery of its parameters possible. In other words, the algorithm will recover the parameters of a region if its area is big enough to contain >50% of the analysis windows in one data window. If this condition does not hold, it will not be possible to detect a region. Thus, to achieve maximum sensitivity to small 3-D structures only three consecutive slices are chosen for a data window. The experimental results, with synthetic and real data (to be described in the following section) support the adequacy of this approach. The size of both windows (data/analysis) in any of the three axis may be changed by the user by choosing from pre-set options.

For the *k*th data window, $k = 1, \ldots, K$; the EMD procedure provides a robust estimate of the gray-level value of the majority of the pixels, $\hat{\theta}_k$ (the mode $\hat{\theta}_{\rm EM}$), and the robust estimate of the standard deviation of those pixels, \hat{s}_k (the scale estimate $\hat{s}_{\rm EM}$). The former characterizes the local structure according to the piecewise constant model, the latter the deviations from this model. The histogram of the local mode values, $\hat{\theta}_k$ (rounded to the closest integer) is the global information employed. Three consecutive slices from the



Figure 3. Synthetic data. (a) An example of a noisy slice from the stack shown in Figure 2. (b) Gray-level histogram of the slice.



Figure 4. Robust estimates of the local structure derived from three slices. (a) Histogram of the local modes. (b) The most probable standard deviation of the local mode values.

stack, with the one in Figure 3a (center), yield the histogram of local modes in Figures 4a; cf Figure 3b. Note that the peaks corresponding to the objects and the background are now clearly distinguishable.

In order to guarantee that the local modes indeed represent only homogeneous windows, and avoid the case when a non-homogeneous window produces accidentally the same gray-level value as a homogeneous one, the histogram of all standard deviations is produced. The value in the 70percentile is used to censor the $\hat{\theta}_k$ values and the histogram of local modes is updated.

The mode of the standard deviations mapped into a bin (i.e. associated with similar $\hat{\theta}_k$ values) is the robust, most probable

standard deviation for the local structures mapped into that bin. It is denoted \hat{r}_i , i = 0, ..., 255; and in Figure 4b these values are shown.

3.2. Computation of significant classes

The significant classes in the lth slice are extracted from the histogram of local modes by a robust, multilevel thresholding algorithm which is a 1-D application of a robust clustering technique (Jolion *et al.*, 1991).

- 3.2.1. Analysis of the histogram of local modes
 - 1. Stretch the histogram into an ordered list of length n. Set j = 1.

Table 1. True parameters versus estimated parameters for slice l = 62 (Figure 3a).

Background	$\mu_1 = 37$	$\sigma_1 = 8.0$	$\hat{\theta}_1 = 37$	$\hat{s}_1 = 6.9$
Object 1	$\mu_2 = 71$	$\sigma_2 = 8.0$	$\hat{\theta}_2 = 70$	$\hat{s}_2 = 7.2$
Object 2	$\mu_3 = 93$	$\sigma_3 = 8.0$	$\hat{\theta}_3 = 93$	$\hat{s}_3 = 6.4$
Object 3	$\mu_4 = 135$	$\sigma_4 = 8.0$	$\hat{\theta}_4 = 135$	$\hat{s}_4 = 7.5$
Object 4	$\mu_5 = 153$	$\sigma_5 = 8.0$	$\hat{\theta}_5 = 153$	$\hat{s}_5 = 7.9$

- 2. Set b = 0.5. Set c = 0.2.
- 3. Apply the mode detection algorithm using a window containing $h = [b \times n] + 1$ points. The mode corresponds to index $i_{m,h}$. The class candidate is defined by $\hat{\theta}_{M,h}$ and $\hat{r}_{m,h}$, where $m = i_{m,h} + [0.5 \times b \times n]$. Note that the most probable standard deviation is associated with the class.
- 4. Compute the mass ω_h contained in the bins which contributed to the $i_{m,h}$ th window. Define the compactness measure

$$\eta_h = \frac{\omega_h}{2\delta_{i_{m,h}}}$$

- 5. $b \leftarrow b 0.1$. If $b \ge c$ return to 3.
- 6. The class extracted at this iteration corresponds to $h_c = \arg \min_h \eta_h$, i.e. $\hat{\theta}_j = \hat{\theta}_{M,h_c}$, and $\hat{s}_j = \hat{r}_{m,h_c}$.
- 7. Remove from the histogram, all the bins falling within $2\hat{s}_j$ from $\hat{\theta}_j$.
- If the number of remaining points in the histogram is larger than c × n, j ← j + 1, return to 2; else STOP.

The algorithm stops if the number of points in the histogram becomes <100c% of the original number. The parameter c affects the number of classes identified, and by modifying it the user can increase or decrease the number of classes, i.e. the resolution of the segmentation.

In Table 1 the values used in generating the synthetic data, μ_j and σ_j (for slice l = 62), are compared with their estimates from the three slices used in Figure 4, $\hat{\theta}_j$ and \hat{s}_j . Note the accuracy of the estimation.

In order to find the changing parameters of the classes through all the stack, the estimation of significant classes was done iteratively as described above, for all possible sets of three consecutive slices in the stack. Figure 5a represents the true mean gray value of the background and objects, as they change slightly through the stack of images. The x-axis corresponds to the gray-level value, the y-axis corresponds to



Figure 5. Variation of the mean gray-level value across the stack. (a) Original; (b) recovered.

the number of slices in the stack. The leftmost continuous line represents the background. From left to right the next lines represent respectively the large hollow half-sphere, the small solid sphere, and the two small hollow half-spheres. As can be seen from Figure 5b, the number of existing objects and the local changes in the mean gray-level values are recovered correctly for each slice across the stack.

3.3. Segmentation of 3-D regions

The 3-D structures are extracted by classification, i.e. each pixel is allocated to a class j = 1, ..., J. As the classification becomes more ambiguous, the role of local spatial information is increased by iteratively relaxing the decision parameter a. Each iteration comprises two traversals of the image stack of L slices. First, the seed regions containing those pixels that have the highest probability of belonging to a class are determined (they are the pixels far from class boundaries). These seed regions are then used in a region-growing algorithm. Both algorithms are based on the 3-D neighborhood of a pixel.

- 3.3.1. Segmentation algorithm
 - 1. Set the value of a, set the value of a^* , N = 15 and M = 1.
 - 2. Run 'seed-planting algorithm' with parameter a.
 - 3. Run 'region-growing algorithm' with parameter a.
 - 4. If there remained unlabeled pixels in the stack, or $a < a^*$,

 $a \leftarrow a + 0.5$, return to 2; else, STOP.



Figure 6. The role of the decision threshold a. (a) a = 1.0; (b) a = 1.5 (see text).

3.3.2. Seed-planting algorithm

1. Set slice index l = 1.

- 2. A pixel in slice *l* is labeled as class $j = 1, ..., J_l$, if its gray level lies in the interval $(\hat{\theta}_j - a\hat{s}_j, \hat{\theta}_j + a\hat{s}_j)$; and at least *N* of its 3-D neighbors have gray levels in the same interval or were already labeled as class *j*.
- 3. $l \leftarrow l + 1$. If $l \le L$ return to 2; else STOP.

3.3.3. Region-growing algorithm

1. Set slice index l = 1.

2. An unlabeled pixel in slice l is labeled as class $j = 1, \ldots, J_l$,

if its gray level lies in the interval $(\hat{\theta}_j - a\hat{s}_j, \hat{\theta}_j + a\hat{s}_j)$; and at least M of its 3-D-neighbors was already labeled as class j.

- 3. If no more labeling is possible, proceed to 4; else return to 2.
- 4. $l \leftarrow l + 1$. If $l \le L$ return to 2; else STOP.

The parameter *a* controls the flow of global information (class definition's). This parameter, combined with parameter *N*, is of basic importance in MR images, where the existence of a possible bias in the intensities may produce overlapping of the distributions of different tissues. Choosing one global threshold to separate two distributions (Brummer *et al.*, 1993) leads to substantial misclassification. When the value of the parameter *a* is small the pixels are allocated based on narrow intervals around $\hat{\theta}_j$, and only the pixels in that interval are candidates (Figure 6a). As the algorithm proceeds, the value of *a* is increased, a pixel may become a candidate in more than one class (Figure 6b). Such a pixel will be

allocated to the class to which the majority of its classified neighbors belong. Thus the problem of the appearance of new isolated seed regions, a problem simple thresholding methods invariably produce, is avoided. In practice, the proposed algorithm results in an adaptive thresholding, where the value that separates two distributions depends on the values of the pixels in the neighborhood. The parameter $N \le 26$ influences the amount of under- or over-segmentation. For example, if N is close to the upper limit, the classification is based on very strict conditions and the slice will be broken into many small pieces. Note that N depends implicitly on a. The implementation of the algorithm used N = 15. More complicated rules, like 'N increases linearly with a' did not seem to have a significant influence for real (MRI) data.

The number of already classified neighbors (M) required in the region-growing algorithm controls the speed of aggregation and affects the shape of the delineated regions. A small M (close to 1) allows the delineation of long elongated features, while a large M (close to N) makes the algorithm strict, and the regions grow slowly. If there is a tie in the local majority rule, the pixel is allocated to the class which yields the smallest Mahalanobis distance (computed with $\hat{\theta}_i$ and \hat{s}_i).

Note that N = M = 0 corresponds to a simple thresholding operation, while non-zero values for these parameters impose local spatial constraints in the segmentation process. Such constraints have a regularization effect and similar rules are employed in any practical segmentation algorithm, e.g. Beveridge *et al.* (1989).

The increase of the parameter a can be stopped at $a^* = 2.5$ or 3, in which case a few pixels will not be classified. These pixels can then be allocated based on the already delineated local structure or used to define new features. Experiments with the MRI data have shown that for $a^* = 3.0$ the delineation is very close to final. Note that parameters a^* , N and M are context independent, i.e. they do not depend on the processed image stack.

The classes determined from all possible three consecutive slices (Figure 5b) were used to analyze the stack of noisy synthetic data derived from Figure 2. The results are shown in Figure 7. The volume recovery algorithm extracts volumes-of-interest (VOI) from all labeled slices. The user specifies the VOI by pointing to any pixel inside of it in any labeled slice where part of the volume is visible. The system marks the whole volume in all slices through a 3-D-filling algorithm (Foley *et al.*, 1990). It is possible to recover several VOIs simultaneously. Given the open design of the algorithm, additional modules to delete false links using mathematical morphology (Brummer *et al.*, 1993; Kapur *et al.*, 1996), or to integrate *a priori* knowledge concerning the location and shape of the VOIs already at the level of decision making, may be incorporated (Raya, 1990; Li *et al.*, 1993;



Figure 7. 3-D rendering of segmented objects from the synthetic image stack.

Sonka *et al.*, 1996). The four objects were segmented and displayed separately, keeping the same orientation they have in Figure 2.

4. EXPERIMENTAL RESULTS

4.1. Magnetic resonance images

We have performed experiments with three brain sets: axial, sagittal and coronal. Here only the sagittal brain set is described. There is no information available about the acquisition protocol parameters used. Furthermore, the set was neither filtered nor preprocessed before being segmented.

The sagittal set contains 109 slices, where each slice is a 256 \times 256 image with 4096 gray levels, rescaled (using the same function) to 256 gray levels. To investigate the sensitivity of the algorithm to the size of the data windows the set was processed using different data window sizes (the analysis window size was always $3 \times 3 \times 3$) and the class parameters were obtained for each triplet. The results are shown in Figure 8. The classes plotted, from left to right, are: class 1 corresponds to background, class 2 corresponds to cerebrospinal fluid (CSF), class 3 corresponds to gray matter and class 4 corresponds to white matter.

In Figure 9 the segmentation of the brain from the sagittal set is shown. The operational parameters of the system were data window size $3 \times 9 \times 9$, analysis window size $3 \times 3 \times 3$, $c = 0.1, a = 1.0, a^* = 3.0, N = 15$ and M = 1. Figure 9a, c



Figure 8. Variation of the mean gray-level value of the estimated classes across the sagittal set. (a) Results with data window size $3 \times 7 \times 7$. (b) Results with data window size $3 \times 9 \times 9$. (c) Results with data window size $3 \times 11 \times 11$.

and e show examples of the original slices. Figure 9b, d and f show the labeling of those slices. All pixels not belonging to any of the four classes were assigned to an 'unknown' class (white) which corresponds to fat tissue.

Two 3-D views of the brain are shown in Figure 10. Figure 10a shows the largest component with pixels labeled as classes 3 and 4. Figure 10b shows only the largest connected component with pixels labeled as class 4. Figure 10c shows the volume occupied by CSF (pixels labeled as class 2) in the lateral ventricles displayed at a larger scale. The procedure of segmenting 109 images size 256×256 in five classes took 13 min on a Sun Spare 20.

Finer structures contained in a subset of slices can also be identified by tuning the parameters of the system to a predefined subset of slices. In this case, however, a semiautomatic procedure is recommended. A triplet of slices containing the structure of interest is chosen and its class definition is extracted. The subset of slices in which the substructure appears is then defined. While this step can be performed automatically, we did it manually in the example. The extracted class parameters are used across the whole subset of slices to delineate the structure of interest.

The result shown in Figure 11 was produced by setting the data window size to $3 \times 7 \times 7$, analysis window size $3 \times 3 \times 3$, c = 0.02, a = 1.0, $a^* = 5.0$, N = 15 and M = 1, and centering the estimation on slice l = 38, where this substructure of interest is visible. Only those images where the substructures can be identified (l = 35, ..., 74) were labeled. Six classes were obtained in this case: in addition to the five classes of the previous example, a sixth class in the



Figure 9. Segmentation of the brain from the *sagittal* set. (a) Slice l = 48. (b) Labeling of slice l = 48. (c) Slice l = 38. (d) Labeling of slice l = 38. (e) Slice l = 25. (f) Labeling of slice l = 25.

white matter area was obtained due to the fact that 98% of the histogram of local modes was now analyzed. The sixth class corresponds to the putamen, which although it is very faintly delineated in the original image, the algorithm is able to detect it as a different homogeneous region. Figure 11a shows the contour of the sixth class in the central area of the brain in slice l = 38, Figure 11b shows the contours of the sixth class in the central area of the brain in slice l =45. These borders correspond to the putamen, the pulvinar nuclear complex and the caudate nucleus. Figure 11c is the volume rendering of the connected component corresponding to the substructures.

The above example shows the advantage of using the proposed method as a preprocessing technique for deformable models algorithms (Davatzikos and Prince, 1995; Snell *et al.*, 1995; Staib and Duncan, 1996; Székely *et al.*, 1996). Given





Figure 10. Segmentation of the brain in the sagittal set. (a) Volume rendering of the gray and white matter. (b) Volume rendering of the white matter only. (c) Volume rendering of CSF in the lateral ventricles.

с



Figure 11. Segmentation of the sagittal set into substructures. (a) Central slice (l = 38) used for parameter estimation. (b) Labeling of the central slice. (c) Volume rendering of connected component corresponding to pulvinar nuclear complex, caudate nucleus and putamen.



Figure 12. Segmentation of the cell from the neuron set. (a) Central slice (l = 100) used for parameter estimation. (b) Labeling of slice l = 100. (c) Slice l = 80. (d) Labeling of slice l = 80. (e) 3-D rendering of the segmented neuron viewed from the top of the stack. (f) 3-D rendering of the segmented neuron viewed from the bottom of the stack.

the robustness of the estimation and the quality of the regiongrowing algorithm, it is a suitable method for producing accurate initial contour or surface models.

4.2. Confocal microscope images

Since the method proposed is designed to work on any stack of 2-D slices, we applied the technique to an image stack obtained with a confocal microscope.

Confocal systems offer the chance to image thick biological tissue in three dimensions. They operate in the bright-field and fluorescence modes, allowing the formation of high-resolution images with a depth of focus sufficiently small that all the detail which is imaged appears in focus and the out-of-focus information is rejected (Wilson, 1990). Some of the current applications in biological studies are in neuron research (Cameron *et al.*, 1997). In materials science or semiconductor device technology, this technique allows the measurement of surface topography (Sheppard and Cogswell, 1990). As there is a lack of good equipment for 3-D visualization in this area (Forsgren *et al.*, 1990), we propose our algorithm as a suitable technique for the confocal community.

In Figure 12 the segmentation of a neuron is shown. The set contains 150 images, and as there are no intensity inhomogeneities between slices, the estimation of parameters was made from three consecutive images: l = 99, 100, 101, using the default parameters of the system. Figure 12a shows the central slice (l = 100) of the triplet used to estimate the parameters. Figure 12b shows the labeling of the central slice in eight classes: four corresponding to different intensities of the background and four to different intensities of the cell. Figure 12c shows another slice (l = 80). Figure 12d is the labeling of slice l = 80. The 3-D view of the neuron as seen from the top of the stack is shown in Figure 12f.

4.3. Comparison with manual segmentation

A set of seven randomly chosen images from the sagittal set were segmented manually by two physicians with expertise in the analysis of MR images (see Acknowledgements). One of the chosen images is shown in Figure 13a. Figure 13b and c show the automatic and manual segmentation of gray matter respectively. Figure 13d and e show the automatic and manual segmentation of white matter. In order to validate the results of the proposed method, two different comparisons were made between automatic and manual segmentations. Table 2 shows the average confusion in the seven images. The rows in the confusion matrix indicate each tissue segmented manually. For example, of all the white matter identified by the experts, 92.4% was correctly labeled as white matter by our method, 7.6% was labeled as gray matter and the rest (0.1%) was labeled as CSF. The comparison of manual versus automatic gray matter segmentation shows more confusion, mainly between gray matter and CSF. Further experiments comparing the confusion among several manual segmentations should be performed, in order to conclude whether the differences are within tolerance.

A quantitative comparison of statistics between manual and automatic segmentation was also carried out. In Tables 3 and 4, the second column (left to right) shows the mean and standard deviation for the manual segmentation of white and gray matter respectively, for each of the seven images. The remaining columns show the parameters of the classes obtained





Figure 13. Comparison of automatic versus manual segmentation. (a) Original image. (b) Automatic segmentation of gray matter. (c) Manual segmentation of gray matter. (d) Automatic segmentation of white matter.

 Table 2. Comparison of manual segmentation versus automatic segmentation.

Manual versus automatic	White matter	Gray matter	CSF	
White matter	92.4%	7.6%	0.1%	
Gray matter	5.4%	84.5%	10.1%	

automatically when using different data window sizes: $3 \times 7 \times 7$, $3 \times 9 \times 9$ and $3 \times 11 \times 11$. As was shown previously in Figure 8, the numerical differences between the parameters estimated automatically using different data window sizes is acceptable. The difference with the computed mean and standard deviation of the area delineated by the expert is in

 Table 3. Comparison of manual segmentation versus automatic segmentation for white matter.

Image	Manual	$3 \times 7 \times 7$	$3 \times 9 \times 9$	$3 \times 11 \times 11$
1	$\mu = 123.85$	$\hat{\theta} = 126.5$	$\hat{\theta} = 126.5$	$\hat{\theta} = 128.5$
	$\sigma = 9.16$	$\hat{s} = 8.37$	$\hat{s} = 9.06$	$\hat{s} = 12.93$
2	$\mu = 122.41$	$\hat{\theta} = 125.5$	$\hat{\theta} = 125.5$	$\hat{\theta} = 125.0$
	$\sigma = 12.53$	$\hat{s} = 8.5$	$\hat{s} = 8.7$	$\hat{s} = 6.7$
3	$\mu = 121.72$	$\hat{\theta} = 132.0$	$\hat{\theta} = 132.5$	$\hat{\theta} = 132.5$
	$\sigma = 16.15$	$\hat{s} = 10.6$	$\hat{s} = 13.4$	$\hat{s} = 13.8$
4	$\mu = 119.69$	$\hat{\theta} = 117.5$	$\hat{\theta} = 119.0$	$\hat{\theta} = 114.5$
	$\sigma = 16.76$	$\hat{s} = 12.15$	$\hat{s} = 15.52$	$\hat{s} = 16.3$
5	$\mu = 128.71$	$\hat{\theta} = 131.0$	$\hat{\theta} = 132.0$	$\hat{\theta} = 132.0$
	$\sigma = 14.23$	$\hat{s} = 11.75$	$\hat{s} = 7.58$	$\hat{s} = 13.64$
6	$\mu = 133.54$	$\hat{\theta} = 131.0$	$\hat{\theta} = 130.5$	$\hat{\theta} = 130.0$
	$\sigma = 13.65$	$\hat{s} = 8.27$	$\hat{s} = 12.23$	$\hat{s} = 11.61$
7	$\mu = 132.49$	$\hat{\theta} = 128.0$	$\hat{\theta} = 130.5$	$\hat{\theta} = 129.5$
	$\sigma = 12.85$	$\hat{s} = 6.63$	$\hat{s} = 6.59$	$\hat{s} = 11.85$

Table 4. Comparison of manual segmentation versus automatic segmentation for gray matter.

Image	Manual	$3 \times 7 \times 7$	3 × 9 × 9	$3 \times 11 \times 11$
1	$\mu = 87.83$	$\hat{\theta} = 83.5$	$\hat{\theta} = 83.5$	$\hat{\theta} = 83.0$
	$\sigma = 8.98$	$\hat{s} = 10.39$	$\hat{s} = 11.94$	$\hat{s} = 14.93$
2	$\mu = 86.59$	$\hat{\theta} = 88.0$	$\hat{ heta} = 85.0$	$\hat{ heta} = 84.0$
	$\sigma = 15.56$	$\hat{s} = 10.3$	$\hat{s} = 11.9$	$\hat{s} = 12.8$
3	$\mu = 89.67$	$\hat{ heta} = 87.0$	$\hat{ heta} = 88.5$	$\hat{\theta} = 92.0$
	$\sigma = 16.42$	$\hat{s} = 13.9$	$\hat{s} = 14.16$	$\hat{s} = 16.1$
4	$\mu = 86.27$	$\hat{\theta} = 81.0$	$\hat{\theta} = 84.0$	$\hat{\theta} = 82.5$
	$\sigma = 12.26$	$\hat{s} = 14.06$	$\hat{s} = 15.9$	$\hat{s} = 13.12$
5	$\mu = 92.61$	$\hat{\theta} = 95.0$	$\hat{\theta} = 90.5$	$\hat{ heta} = 90.0$
	$\sigma = 10.05$	$\hat{s} = 11.90$	$\hat{s} = 13.19$	$\hat{s} = 14.67$
6	$\mu = 79.97$	$\hat{ heta} = 81.0$	$\hat{\theta} = 83.0$	$\hat{\theta} = 82.5$
	$\sigma = 11.19$	$\hat{s} = 10.07$	$\hat{s} = 13.35$	$\hat{s} = 13.91$
7	$\mu = 78.04$	$\hat{\theta} = 77.0$	$\hat{\theta} = 77.5$	$\hat{ heta} = 84.0$
	$\sigma = 7.71$	$\hat{s} = 6.12$	$\hat{s} = 10.10$	$\hat{s} = 11.96$

some cases more significant, although for practical purposes, as the confusion matrix shows, it can be considered within tolerance. The fact that the standard deviation of the manual segmentation is in some cases greater than that obtained automatically is due to the method focusing on the most homogeneous regions, and the physicians including voxels in the borders containing partial volumes. The artifact created by partial volumes were discussed recently in Niessen *et al.* (1998).

5. CONCLUSIONS

We have described a new method for extracting 3-D information from MR images. The method makes extensive use of robust statistical techniques, and it is an automatic, bottom– up procedure.

The algorithm uses the assumption of piecewise constancy. By using all the triplets for 'local' estimation of the class parameters, as in the full set MRI example, smooth changes in the class parameters as one moves along the stack are automatically compensated. The intensity bias within an image is a harder problem. Most often it is equipmentdependent and can be characterized and eliminated through adequate calibration procedures. Image processing techniques to eliminate the artifacts introduced by equipment are also available (Gerig et al., 1992; Wells et al., 1996). If a more complex model is to be employed, the EMD procedure described in the paper can be generalized to piecewise polynomial data (Mintz et al., 1992) or methods based on random fields can be used (Marroquin, 1992). However, substituting the piecewise constant image structure model with a (say) piecewise planar model introduces the problem of choosing locally the adequate model order. Even robust methods will yield different results if the underlying assumptions are different. A safer procedure is to use the traditional piecewise constant model and then by analyzing the residuals assess the significance of the bias present.

The desired resolution of analysis is controlled by a few parameters whose effect is straightforward and totally data independent. The main characteristic of the method is a local/global equilibrium: only significant local homogeneous regions become representative global classes. The modular design also allows the recovery of correct class definitions when intensity inhomogeneities between slices exist, as well as the identification of different number of classes at different levels of the stack. Further improvements are also allowed by the modular design, such as increasing the resolution of object delineation by multiple and/or hierarchical estimation of the classes. The technique described can serve as the main module of a complete MR data analysis system (incorporating a priori knowledge about the anatomy), a system in which the constraints on the user-machine interface can be relaxed due to the robustness of the image analysis. The delineated volumes can provide an excellent initial guess for the deformable surface techniques whose performance is strongly dependent on a reliable starting point in the minimization (Davatzikos and Prince, 1995; Snell et al., 1995; Staib and Duncan, 1996; Székely et al., 1996). Such techniques can also be used to further refine the segmentation.

ACKNOWLEDGEMENTS

We thank the anonymous reviewers for their constructive comments which contributed to a significant improvement in the presentation of the paper. The sagittal set was provided courtesy of Siemens Medical Systems, Inc., Iselin, NJ, and is contained in the Chapel Hill Volume Rendering Test Data Site. These images were taken on the Siemens Magnetom. The neuron set was provided courtesy of Dr Martin Wessendorf, Department of Cell Biology and Neuroanatomy, School of Medicine, University of Minnesota. The authors would also like to thank Dr Ramon Gutierrez Alvarado and Dr Bernardo Escutia, Instituto Nacional de Neurologia y Neurocirugia, Mexico City, who segmented the images manually. Peter Meer was partially supported by the the National Science Foundation under the grant IRI-9530546.

REFERENCES

- Arata, L. K., Dhawan, A. P., Broderick, J. P., Gaskil-Shipley, M. F., Levy, A. V. and Volkow, N. D. (1995) Three-dimensional anatomical model-based segmentation of MR brain images through principal axes registration. *IEEE Trans. Biomed. Engng*, 42, 1069–1078.
- Ayache, N. (1995) Medical computer vision, virtual reality, and robotics. *Image Vision Comput.*, 13, 295–313.
- Beveridge, J. R., Griffith, J., Kohler, R. R., Hanson, A. R. and Riseman, E. M. (1989) Segmenting images using localized histograms and region merging. *Int. J. Comp. Vision*, 2, 311– 347.
- Bomans, M., Hohne, K. H., Tiede, U. and Riemer, M. (1990) 3D segmentation of MR images of the head for 3D display. *IEEE Trans. Med. Imag.*, 9, 177–183.
- Brummer, M. E., Mersereau, R. M., Eisner, R. L. and Lewine, R. R. J. (1993) Automatic detection of brain contours in MRI data sets. *IEEE Trans. Med. Imag.*, 12, 153–166.
- Cameron, D. L., Wessendorf, M. W. and Williams, J. T. (1997) A subset of VTA neurons is inhibited by dopamine, 5-HT, and opioids. *Neuroscience*, 77, 155–166.
- Cheng, K. S., Lin, J. S. and Mao, C. W. (1996) The application of competitive Hopfield neural network to medical image segmentation. *IEEE Trans. Med. Imag.*, 15, 560–567.
- Cho, Z. H., Jones, J. P. and Singh, M. (1993) Foundations of Medical Imaging. John Wiley & Sons, Inc., New York.
- Cline, H. E., Lorensen, W. E., Kikinis, R. and Jolesz, F. (1990) Three-dimensional segmentation of MR images of the head using probability and connectivity. J. Comp. Aided Tomogr., 14, 1037–1045.
- Davatzikos, C. A. and Prince, J. L. (1995) An active contour model for mapping the cortex. *IEEE Trans. Med. Imag.*, 14, 65–80.
- Deklerck, R., Cornelis, J. and Bister, M. (1993) Segmentation of medical images. *Image Vision Comput.*, 11, 486–503.

- Foley, J. D., van Dam, A., Feiner, S. K. and Hughes, J. F. (1990) Computer Graphics: Principles and Practice, The Systems Programming Series, pp. 979–986. Addison-Wesley Publishing Company, New York.
- Forsgren, P. O., Franksson, O. and Liljeborg, A. (1990) Software and electronics for a digital three-dimensional microscope. In Wilson, T. (ed.), *Confocal Microscopy*, pp. 65–92. Academic Press, London.
- Garza-Jinich, M., Meer, P. and Medina, V. (1996) Robust retrieval of 3D structures from magnetic resonance images. *Proc. 13th Int. Conf. on PR*, Vienna, Vol. 3, pp. 391–395 IEEE Computer Society Press.
- Gerig, G., Kübler, O., Kikinis, R. and Jolesz, F. A. (1992) Nonlinear anisotropic filtering of MRI data. *IEEE Trans. Med. Imag.*, 11, 221–232.
- Grimson, W. E. L. (1995) Medical applications of image understanding. *IEEE Expert*, 10, 18–28.
- Herman, G. T. and Liu, H. K. (1979) Three dimensional display of human organs from computed tomograms. *Comp. Graphics Image Process.*, 9, 1–21.
- Johnston, B., Atkins, M. S., Mackiewich, B. and Anderson, M. (1996) Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. *IEEE Trans. Med. Imag.*, 15, 154– 169.
- Jolion, J. M., Meer, P. and Bataouche, S. (1991) Robust clustering with applications in computer vision. *IEEE Trans. PAMI*, 13, 791–802.
- Joliot, M. and Mazoyer, B. M. (1993) Three-dimensional segmentation and interpolation of magnetic resonance brain images. *IEEE Trans. Med. Imag.*, 12, 269–277.
- Kapur, T., Grimson, W. E. L., Wells, W. M. and Kikinis, R. (1996) Segmentation of brain tissue from magnetic resonance images. *Med. Imag. Anal.*, 1, 109–127.
- Kennedy, D. N., Filipek, P. A. and Caviness, V. S. (1989) Anatomic segmentation and volumetric calculations in nuclear magnetic resonance imaging. *IEEE Trans. Med. Imag.*, 8, 1–7.
- Li, C., Goldgof, D. B. and Hall, L. O. (1993) Knowledge-based classification and tissue labeling of MR images of human brain. *IEEE Trans. Med. Imag.*, 12, 740–750.
- Liang, Z., MacFall, J. R. and Harrington, D. P. (1994) Parameter estimation and tissue segmentation from multispectral MR images. *IEEE Trans. Med. Imag.*, 13, 441–449.
- Liou, S. P. and Jain, R. C. (1993) An approach to three-dimensional image segmentation. *Comp. Vision Graphics Image Process.*, 53, 237-252.
- McInerney, T. and Terzopoulos, D. (1996) Deformable models in medical image analysis. Proc. Medical Models in Biological Image Analysis 1996, San Francisco, CA, pp. 171–180. IEEE Computer Society Press, Los Alamitos, CA.
- Marroquin, J. L. (1992) Random measure fields and the integration of visual information. *IEEE Trans. SMC*, 22, 705-716.
- Meer, P., Mintz, D., Rosenfeld, A. and Kim, D. Y. (1991) Robust regression methods for computer vision: a review. Int. J. Comp. Vision, 6, 59-70.

- Mintz, D., Meer, P. and Rosenfeld, A. (1992) Consensus by decomposition: a paradigm for fast high breakdown point robust estimation. *Proc. 1991 DARPA IUW*, La Jolla, CA, pp. 345-362.
- Niessen, W. J., Vincken, K. L., Weickert, J. and Viergever, M. A. (1998) Three-dimensional MR brain segmentation. 6th Int. Conf. on Computer Vision, Bombay, India, January 1998, pp. 53-58.
- Raman, S. V., Sarkar, S. and Boyer, K. L. (1993) Hypothesizing structures in edge-focused cerebral magnetic resonance images using graph-theoretic cycle enumeration. *Comp. Vision Graphics Image Process.*, 57, 81–98.
- Raya, S. P. (1990) Low level segmentation of 3D magnetic resonance brain images—a rule-based system. *IEEE Trans. Med. Imag.*, 9, 327–337.
- Robb, R. A. and Barillot, C. (1989) Interactive display and analysis of 3-D medical images. *IEEE Trans. Med. Imag.*, 8, 217–226.
- Robinson, G. P., Colchester, A. C. F. and Griffin, L. D. (1994) Model-based recognition of anatomical objects from medical images. *Image Vision Comput.*, 12, 499–507.
- Rousseeuw, P. J. and Leroy, A. M. (1987) Robust Regression & Outlier Detection. John Wiley & Sons, New York.
- Sander, P. T. and Zucker, S. W. (1990) Inferring surface trace and differential structure from 3D images. *IEEE Trans. PAMI*, 12, 833-854.
- Sheppard, C. J. R. and Cogswell, C. J. (1990) Three-dimensional imaging in confocal microscopy. In Wilson, T. (ed.), *Confocal Microscopy*, pp. 143–169. Academic Press, London.
- Smith, H. J. and Ranallo, F. N. (1989) A Non-Mathematical Approach to Basic MRI. Medical Physics Publishing Corporation, Madison, WI.
- Snell, J. W., Merickel, M. B., Ortega, J. M., Goble, J. C., Brookeman, J. R. and Kassell, N. F. (1995) Model-based boundary estimation of complex objects using hierarchical active surface templates. *Pattern Recogn.*, 28, 1599–1609.
- Soltanian-Zadeh, H., Windham, J. P., Peck, D. J. and Yagle, A. E. (1992) A comparative analysis of several transformations for enhancement and segmentation of magnetic resonance image scene sequences. *IEEE Trans. Med. Imag.*, 11, 302–318.
- Sonka, M., Tadikonda, S. K. and Collins, S. M. (1996) Knowledgebased interpretation of MR brain images. *IEEE Trans. Med. Imag.*, 15, 443–452.
- Staib, L. H. and Duncan, J. S. (1996) Model-based deformable surface finding for medical images. *IEEE Trans. Med. Imag.*, 15, 720–731.
- Székely G., Kelemen, A., Brechbühler, C. and Gerig, G. (1996) Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface model. *Med. Imag. Anal.*, 1, 19–34.
- Taxt, T. and Lundervold, A. (1994) Multispectral analysis of the brain using magnetic resonance imaging. *IEEE Trans. Med. Imag.*, 13, 470–481.
- Tsai, C., Manjunath, B. S. and Jagadeesan, R. (1995) Automated segmentation of brain MR images. *Pattern Recogn.*, 28, 1825– 1837.

- Tiede, U., Hoene, K. H., Bomans, M., Pommert, A., Riemer, M. and Wiebecke, G. (1990) Investigation of medical 3D-rendering algorithms. *IEEE Comput. Graphics Appl.*, 10, 30–41.
- Udupa, J. K. (1991) Computer aspects of 3D imaging in medicine: a tutorial. In Udupa J. K. and Herman G. T. (eds.), 3D Imaging in Medicine, pp. 1–69, CRC Press, Boca Raton, FL.
- Udupa, J. K. (1994) Multidimensional digital boundaries. Comp. Vision Graphics Image Process., 56, 311-323.
- Wells, W. M., Grimson, W. E. L., Kikinis, R. and Jolesz, F. A. (1996) Adaptive segmentation of MRI data. *IEEE Trans. Med. Imag.*, 15, 429–442.
- Wilson, T. (1990) Confocal microscopy. In Wilson, T. (ed.), Confocal Microscopy, pp. 1–64. Academic Press, London.
- Wu, Z. and Leahy, R. (1993) An optimal graph theoretic approach to data clustering: theory and its application to image segmentation. *IEEE Trans. PAMI*, 15, 1101–1113.