Segmentation of Volume Images using a Multiscale Transform*

Tod Courtney and Narendra Ahuja
Beckman Institute
University of Illinois at Urbana-Champaign
405 N. Matthews Ave., Urbana IL 61801 USA
E-mail: tod, ahuja@stereo.ai.uiuc.edu

Abstract

This paper presents a new method for multiscale segmentation of volume images. The segmentation is achieved using a recent nonlinear transform which leads to well-characterized regions at different spatial and intensity scales. The detected three-dimensional regions are closed and are homogeneous relative to their surround. A pyramid is generated containing the region information extracted across a range of homogeneity scales. The pyramid represents the multiscale volumetric structure. Experimental results are given for magnetic resonance data as well as video sequences.

1. Introduction

Recent advances in computer imaging, specifically in medical imaging and high speed CCD cameras, have produce reliable systems for capturing true volume data sets. Volume images, like two-dimensional images, contain structures at many levels of detail. As an example of this, consider an MRI of a human brain. At a fine scale, many different regions in the image can be identified corresponding to fine photometric and structural variations in the brain structure. At a coarser scale, the main physiological structures of the brain, such as the cerebrum and the brain parenchyma, are visible. At the coarsest scale, the entire brain can be seen as one region. The different regions in the image have different sizes and photometric characteristics, i.e., different spatial and intensity scales. It is imperative for the analysis of volume data that good automatic volumetric segmentation methods are developed.

During the last decade, many different approaches to volume segmentation have been presented, including some multiscale methods. Many of these techniques are extensions of common two dimensional segmentation methods. One approach is to perform a two dimensional segmentation on the first slice of the volume, and then extend this segmentation through the other slices, using deformable contours called snakes[2] [5]. Another is to perform a full three-dimensional segmentation, as in [7], where three-dimensional gradient and Laplacian operators are used. A three-dimensional segmentation algorithm which produces a multi-resolutional pyramid based on the Marr-Hildreth operator is presented in [3]. A multiscale method using hyperstacks consisting of blurred versions of the original image was developed in [10].

Other methods rely on specific knowledge of the image domain along with standard segmentation techniques, as in the rule-based system developed by Raya [8] to identify the three main components of the brain: the cerebrum, the brain parenchyma and the cerebrospinal fluid. Chen et al. [6] used morphology with knowledge-based filters specialized to the particular segmentation domain. A more ad hoc method[4] is based on threshold based masks where the thresholds are chosen such that the desired features of the brain are captured.

The algorithm presented in this paper is a three-dimensional extension of a general nonlinear transform developed in [1]. The transform differs from previous multiscale methods in that it incorporates two components of scale which specify the spatial and the homogeneity variations in the data. The spatial scale quantifies the size of the regions, while the homogeneity scale quantifies the amount of intensity variation allowed within a region. The transform encodes the region border information into a three-dimensional vector field. Region bounds are then extracted to form an initial segmentation. This segmentation forms the base of a multiscale pyramid whose higher levels contain increasingly coarser structure. The entire pyramid cap-

*This research was supported by grant N00014-93-1-1167 from the Advanced Research Projects Agency.

1015-4651/96 $5.00 © 1996 IEEE
Proceedings of ICPR '96

tures the hierarchical structure present in the volume data.

The remainder of this paper is organized as follows. Section 2 reviews the nonlinear transform used in this work. Section 3 describes the method of extracting the finest scaled regions from the output of the transform. Section 4 describes the multiscale segmentation pyramid. Experimental results are presented in section 5, with concluding remarks in section 6.

2. Transform

The foundation of this segmentation algorithm is the multiscale transform developed in [1] and [9]. The transform maps the given three-dimensional volume image to a volume of three-dimensional force vectors. The vector at each point represents the force of attraction from the region in which that point belongs. The vectors are directed away from the borders between homogeneous regions and toward the region's medial axis. This allows the borders to be identified by locating divergence of neighboring vectors.

The discrete representation for the transform producing force vector F(x,y,z) at the point P(x,y,z) of a volume V is :

$$F(x,y,z) = \sum_{(u,\ v,\ w)\in R} D_s(x,y,z) D_g(x,y,z) \frac{\vec{r}}{\| \vec{r} \|} \tag{1}$$

where

$$D_s(x,y,z) = d_s(\| \vec{r} \|, \sigma_s(x,y,z)) \tag{2}$$

$$D_g(x,y,z) = d_g(\Delta V(x,y,z), \sigma_g(x,y,z)) \tag{3}$$

$$d_x(x,y) = \begin{cases} 1 & \text{if}(x \le y), \\ 0 & \text{if}(x > y). \end{cases} \tag{4}$$

$$\Delta V = |V(x,y,z) - V(x+u, y+v, z+w)| \tag{5}$$

$$\vec{r} = u\hat{i} + v\hat{j} + w\hat{k} \tag{6}$$

The force vector F(x,y,z) at point P(x,y,z) is computed on a cubic block of voxels centered at P, called the of region of application and denoted as R. The size of R is variable and is determined automatically. Each point, N(x,y,z), within R is tested to see if it should contribute a subvector. The functions d_s and d_g impose spatial and homogeneity scale constraints which will be defined below. If the two points satisfy these scale constraints, the subvector is added to the total force at P.

The functions d_s and d_g are based on two scale parameters which represent represent the spatial and homogeneity variations in the data. The value of the spatial scale, denoted as σ_s, specifies the maximum distance between any point and the region boundary. By varying the value of σ_s, the transform can detect many types of edges efficiently. Sharp edges in the image require processing on only a small block of the volume. Blurred edges require a larger block in order to correctly identify the edge. Through automatic determination of the optimal spatial scale at each point, the transform is able to correctly identify both sharp and gradual edges.

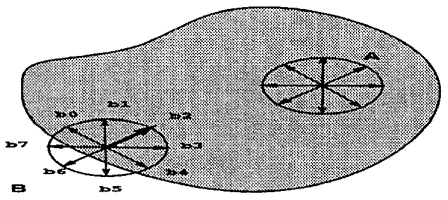


Figure 1. At Point A, the subvectors cancel for σ_s inside a region. For Point B, the subvectors b0-b4 contribute force, while b5-b7 do not. The final force is b2.

The value of the homogeneity scale, referred to as σ_g, determines the amount of intensity variation allowed among voxels within a homogeneous region. At a given value of σ_g, voxels within a identified region have intensity variation less than or equal to σ_g. The variation between voxels inside and outside of a region is greater than σ_g. Specifying the homogeneity scale determines the minimum contrast that must exist between neighboring regions.

Each subvector contributing to the force at P is directed toward a neighbor, N, which is of similar spatial and homogeneity characteristics as P. Since the total force at P is a summation of these subvectors, the resulting force is directed toward other voxels within the homogeneous region containing point P. The force vector at P is guaranteed to accurately indicate the direction of the interior of a homogeneous region.

In order to compute the vector field, σ_g is defined to be spatially invariant. At a given σ_g, homogeneous regions in the image will be of varying sizes and σ_s must be determined for each three-dimensional point in the volume image. The value of σ_s is based on the intensity variation of the neighbors around that point. If a large neighborhood of voxels all have intensity within σ_g of the intensity at point P, σ_s for that point will be large. Thus, σ_s can be thought of as a function of σ_g and the voxel neighborhood of P, although it is never directly expressed as such.

The value of σ_s for a given point is determined by evaluating the transform for an initial value of σ_s ($\sigma_s = 1$) and then repeatedly evaluating the transform at increasingly larger values of σ_s until a stable force vector is produced. A stable force vector is one which is non-zero, and constant in direction over a range of σ_s. Requiring stable force vectors allows the transform to ignore effects of noise and insures that the transform identifies the significant region borders in the data. Once the vector is stable, it is recorded for that point. This process is repeated for every point in the volume.

3. Initial region extraction

The initial regions form the base of the hierarchical segmentation pyramid. These initial regions should consist of the smallest perceptually significant structures in the data. It was experimentally found by [9] that an initial σ_g=3 was sufficient to capture the finest scaled regions. At lower values of σ_g, insignificant regions are found as a result of noise in the imaging system.

To extract the initial regions, the transform is applied to the entire data set to produce the vector field. From the vector field, two signatures of the initial regions are identified. These signatures are edges along the region's borders and skeletons along the region's medial axis.

The signatures are identified using a fundamental property of the transform which guarantees the vectors are directed toward the interior of the regions. As a result, vectors along the borders between regions diverge from one another, and vectors along the region's medial axes converge to form the skeletons. To identify these two signatures, the vectors of neighboring voxels are compared. If the vectors diverge, the two voxels are along an edge and are labeled as edge voxels. If the vectors converge, the voxels are along a skeleton and are labeled as skeleton voxels.

The edge and skeleton voxels are used by a hybrid region growing algorithm to extract the regions from the vector field at the initial σ_g. The algorithm uses the skeletons as initial region seeds. The seed regions are then expanded to include all surrounding unlabeled non-edge voxels. Region growth is halted by the presence of edge voxels. The identified regions represent the finest structures present in the volume data and form the base of the multiscale pyramid.

4. Multiscale pyramid

Multiple segmentations at increasing values of σ_g form a pyramid, encapsulating all the structural information in the data in a range of fine to coarse scales. The lowest level of the pyramid contains the finest scaled regions identified at the initial σ_g. These were identified from the vector field using the method discussed in the previous section. Higher levels in the pyramid have increased values of σ_g and contain coarser segmentations.

The natural formulation for building this pyramid would be to apply the transform directly to the original data for all values of σ_g. Each application of the transform would find all the stable homogeneous regions for that scale. However, applying the transform independently for each value of σ_g is inefficient, since the structural knowledge from finer levels of the pyramid is ignored. Savings in the formation of the new σ_g level can be achieved by using some of the properties of the pyramid.

An important property of this pyramid is causality of scale. Causality of scale states that once a voxel is identified as a member of a given region it will remain in that region for all larger σ_g. The property is derived from the definition of a homogeneous region. Causality of scale allows us to model the regions at a given level of σ_g by regions of the same shape with constant intensity equal to the mean intensity of the original region. The volume image containing these constant intensity regions will be referred to as the mean-intensity region volume.

When the transform is applied to a data volume, the regions are extracted by detecting the divergence of the force vectors along the region boundary. The boundary vectors are computed by summing subvectors from the current point to each of its neighbors within a given region, as described in Section 2. Each subvector has a component which is based on the intensity of the current point, I[x] and the intensity of the neighbor, I[Nbr(x)]. If the relationship $|I[x] - I[Nbr(x)]| < \sigma_g$ is satisfied, the subvector is added to the total force.

In the mean-intensity region image, all voxels within two neighboring regions have the same intensity. The computation of each subvector results in a repeated testing of $|I[x] - I[Nbr(x)]|$. This is a source of inefficiency which is compounded considering the computation must be performed at each voxel along both sides of the border between the two regions. It should be sufficient to test only one subvector at two points on opposite sides of the border between the two regions. If the subvectors diverge, the whole border still exists. If the subvectors do not diverge, the two regions are

homogeneous at the new scale.

When applying the transform to the mean-intensity region volume, the intensity used in the test, $|I[x] - I[Nbr(x)]| < \sigma_g$, is the average intensity of the region at x. Substituting this mean for the actual intensity at x transforms the subvector calculation to the spatially invariant equation: |Mean(R)-Mean(R(Nbr(x)))| $< \sigma_g$ where R is the region containing point x. If this test holds the regions should merge at the new scale, σ_g.

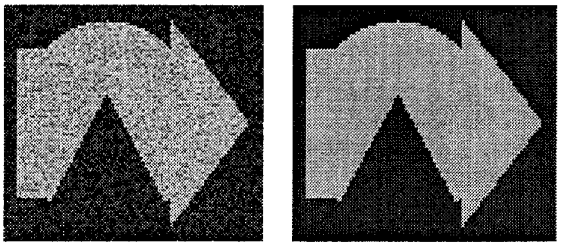
This new algorithm does not need to be applied to the data volume for each level of the pyramid. Instead a data structure containing the region intensity and topological information is used. The data structure records the size, average intensity and list of neighboring regions for all regions at a given σ_g scale. To construct the next level of the pyramid, the algorithm traverses the data structure and compares the average intensity of each region with the average intensity of each of its neighbors. Whenever a region is determined to merge with one of its neighbors, the data structure is updated accordingly for both regions.

5. Experimental results

The segmentation method presented here has been successfully applied to many different sets of volume data, including synthetic objects with noise, single images, motion sequences, and medical volume images. In all cases, the segmentation performed well. For the synthetic volumes, performance was measured quantitatively, by comparing the results with the known structures present. Since the correct regions are not a priori known in real data, quantitative comparisons are not possible. The performance for real data was measured perceptually. Future validation of the real data will compare the results with those of medical experts. The results presented here are cross section images taken from the volume data, with segmented regions colored by their average value.

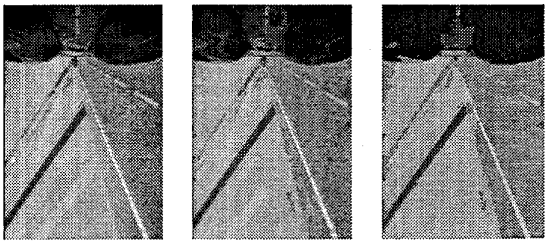
Figure 2 shows cross-sections of a synthetic three-dimensional object and the resulting segmentation. The synthetic object contains independent, identically distributed noise with a uniform distribution. Performance was measured by calculating the number of voxels which were incorrectly labeled. Results in all tests found that all voxels were correctly labeled, verifying the segmentation is robust in the presence of noise.

An example of a segmented video sequence is shown in Figure 3. The spatio-temporal volume is formed from the images in a video sequence. The sequence is taken from a camera mounted on the underside of an aircraft during takeoff. The transform found the regions present in the sequence and tracked their motion through time.



(a) (b)

Figure 2. (a) Cross-section of a synthetic solid object overlaid with random noise and (b) the resulting segmentation.



(a) (b) (c)

Figure 3. Cross-section of a airplane motion sequence volume. (a) Original Image. (b) Fine (σ_g=5) and (c) coarse(σ_g=30) segmentations.

Finally, cross sections are presented from two different MRI volumes (Figure 4). The top row contains the original data. The second row shows a segmentation for the initial value of σ_g, which maintains much of the detail found in the original image. The bottom row shows the structure at a higher level of the pyramid. At this coarse scale, the main structures of the brain such as the gray and white matter are identified.

6. Conclusion

We present a new method of volume segmentation based on a three-dimensional nonlinear transform. The transform incorporates a special concept of scale by independently handling the spatial and intensity properties of the regions in the volume. The transform produces a three dimensional field of force vectors, representing the force applied to each voxel by the region in which it belongs. Because of the nonlinear transform and the properties of the vector field, the transform

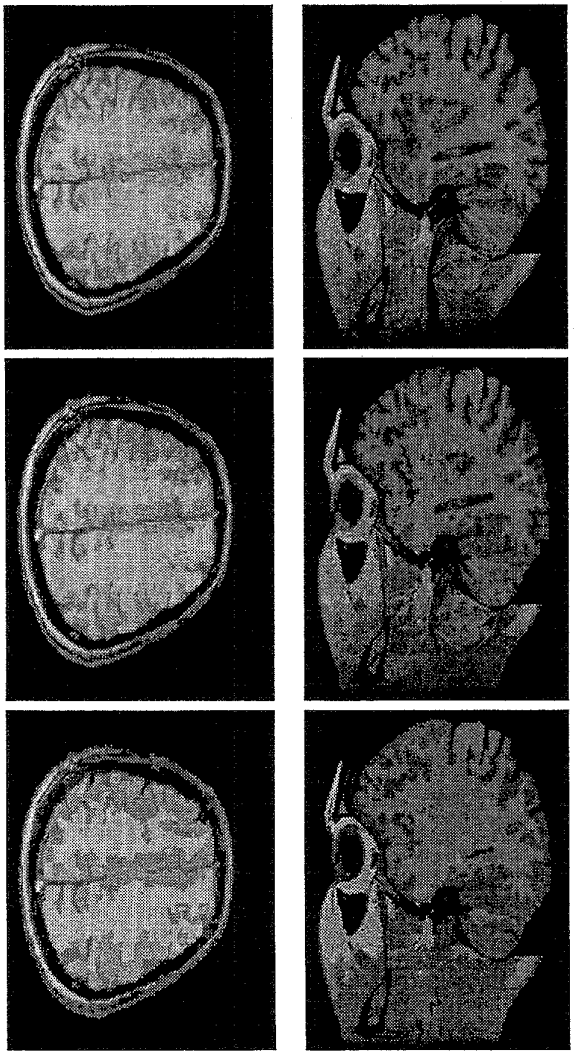


Figure 4. Two different MRI volumes. Original data in first row, fine and coarse segmentations in middle and bottom rows.

produces segmentations which are robust and immune to noise variations.

Regions are extracted from the vector field through a hybrid region growing algorithm. A multiscale pyramid is formed, representing the hierarchical structure present in the data. Experimental results are provided on both real and synthetic data. A detailed viewing of the results reveals the quality of the segmentations.

Future work could focus on higher dimensional algorithms to segment four dimensional data sets, such as three-dimensional volumes changing with time.

References

[1] N. Ahuja. A transform for detection of multiscale image structure. In *Computer Vision Pattern Recognition*, pages 780–781, New York, New York, June 1993.

[2] E. Ashton, M. Berg, K. Parker, C. W. Chen, J. Weisberg, and L. Ketonen. Segmentation and feature extraction techniques, with applications to biomedical images. In *Proceedings of the International Conference on Image Processing 1994*, pages 670–677, Austin, Texas, 1994.

[3] M. Bomans, M.-H. Höhne, U. Tiede, and M. Riemer. 3-D segmentation of MR images of the head for 3-D display. *IEEE Transactions on Medical Imaging*, 9:177–183, June 1990.

[4] M. E. Brummer, R. M. Mersereau, R. L. Eisner, and R. R. J. Lewine. Automatic detection of brain contours in MRI data sets. *IEEE Transactions on Medical Imaging*, 12:153–166, June 1993.

[5] I. Carlborn and K. M. Harris. Computer-assisted registration, segmentation, and 3D reconstruction from images of neuronal tissue sections. *IEEE Transactions on Medical Imaging*, 13(2):351–362, June 1994.

[6] C. W. Chen, J. Luo, K. J. Parker, and T. S. Huang. A knowledge-based approach to volumetric medical image segmentation. In *Proceedings of the International Conference on Image Processing 1994*, pages 493–497, Austin, Texas, 1994.

[7] O. Monga, R. Deriche, G. Maladain, and J. Cocquerez. Recursive filtering and edge closing. In *European Conference on Computer Vision*, pages 56–62, Antibes, France, Apr. 1990.

[8] S. P. Raya. Low-level segmentation of 3-D magnetic resonance brain images - a rule based system. *IEEE Transactions on Medical Imaging*, 9:327–337, Sept. 1990.

[9] M. Tabb and N. Ahuja. Multiscale image segmentation using a recent transform. In *Image Understanding Workshop*, pages 1523–1530, Monterey, California, 1994.

[10] K. L. Vincken, A. S. E. Koster, and M. A. Viergever. Probabilistic segmentation of partial volume voxels. *Pattern Recognition Letters*, 15:477–484, May 1994.