Multiscale, Geometric Image Descriptions For Interactive Object Definition

TR89-033

August, 1989

Stephen M. Pizer, John M. Gauch
James M. Coggins, Timothy J. Cullip
Robin E. Fredericksen, Victoria L. Interrante

The University of North Carolina at Chapel Hill
Department of Computer Science
CB#3175, Sitterson Hall
Chapel Hill, NC 27599-3175

To appear in Proc. Die Deutsche Arbeitsgemeinschaft für Mustererkennung, 1989.

MULTISCALE, GEOMETRIC IMAGE DESCRIPTIONS FOR INTERACTIVE OBJECT DEFINITION

Stephen M. Pizer, John M. Gauch*, James M. Coggins,
Timothy J. Cullip, Robin E. Fredericksen, Victoria L. Interrante

Medical Image Display Research Group
University of North Carolina
Chapel Hill, NC 27599-3175, USA

ABSTRACT

A means is described of analyzing two- and three-dimensional images into a directed acyclic graph of visually sensible, coherent regions and of using this DAG as the basis for interactive object definition. The image analysis is in terms of the geometry of the intensity surface via a multiscale approach with a focus on symmetry properties about ridges. The image analysis method, a system for interactive object definition, and results of their use on two-dimensional images are reported.

1. INTRODUCTION

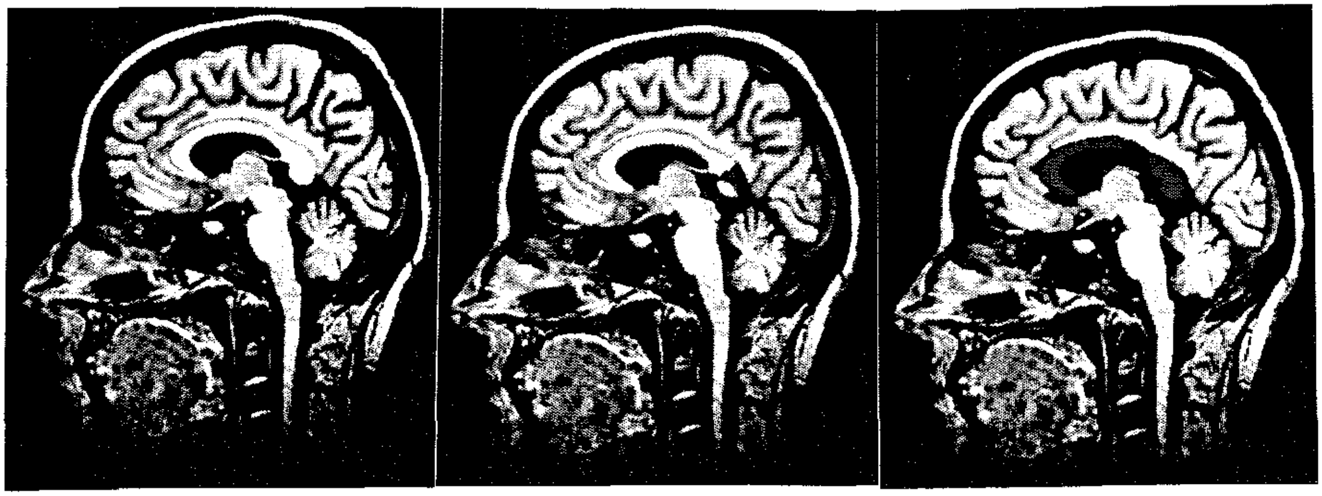
The definition of objects in images can lead to measurements of the objects, display of the objects, or actions based on their properties. Much work in computer vision has taken as its objective the automatic definition of these objects from image data. It seems to us that in an interactive environment a more promising approach is to have the computer analysis provide a framework in which a human who understands the semantics of the scene can quickly define the necessary objects using the image data.

We suggest that without any scene semantics it is possible by automatic computation to describe an image as a related collection of visually sensible, coherent image regions (e.g., see figure 1b) that can provide the basis for quick object definition by the interacting human. More precisely, a method is set forth for automatically computing a directed acyclic graph (DAG) of such image regions with arcs specifying region containment. We demonstrate an interactive system that allows the user, with his or her knowledge of scene semantics, to use this description in the definition of regions containing semantically relevant objects, by interactive selection of DAG regions and possibly some modification of the DAG and *in extremis* the regions defined in its nodes.

Our objective is to demonstrate the power of this approach and the possibility of automatically computing visually sensible, coherent regions without the use of semantics, in particular by a geometric focus on the image as an intensity surface. On the way to this demonstration a number of choices within the analysis approach had to be made. These included the particular axis of symmetry and the means of varying scale. In most cases the alternative chosen was the one which most simplified the initial evaluation. In the following we report encouraging results concerning the power of the overall approach,

* Presently at College of Computer Science, Northeastern University

along with indications of the need to explore some alternatives better matched to the human visual system's function. The paper closes with a discussion of the alternatives needing exploration.



a b c

Figure 1. Visually sensible, coherent image regions: a) the MRI head image used as an example throughout this paper; b) a primitive region; c) a meaningful region.

Our area of application of this work is 3D medical image display. In this area actually defining the part of space making up a meaningful object, or equivalently the surface which encloses that part of space, provides the surface locations and orientations that are the basis for shaded display techniques, whether by surface or volume rendering [Fuchs, 1988; Levoy, 1988]. But once a region has been found that contains the object and parts of its background at contrasting intensities, other more local analysis is available for determining these surface locations and orientations. Therefore, it is also of use simply to define such regions of object containment (*object surrounds* - see figure 1c). In this paper we focus on methods that will lead to such object surrounds while also describing how the same methods can produce the object surfaces themselves.

2. IMAGE DESCRIPTION

We take our goal to be the description of a scalar image in terms of light objects on darker backgrounds. All of the following can be directly transformed for dark objects on lighter backgrounds. The combination of these two analyses is also of interest but is beyond the scope of this paper.

We aim to find image regions lighter than their background that match reasonably well what humans would choose as sensible image regions if they had no knowledge of the semantics of the image scene. We suggest that if the image is viewed as an intensity surface (see figure 2), where height corresponds to image intensity, then shape properties of this surface will determine the region definitions. The need for these regions to be coherent and sensible implies that these shape properties be global properties of the intensities within and around them and not just a collection of local properties such as intensity gradient. That is, the properties will be region-based rather than edge-based. In particular, humans seem to use ridges of intensities as organizing features. Yet, in determining these ridges they are somehow insensitive to rotations of the co-ordinate system and to scaling or monotonic transformations of intensity. (Ideally they should also be insensitive to the context sensitive transformations of intensity effected by the human visual system.) The analysis that we propose must have these characteristics.

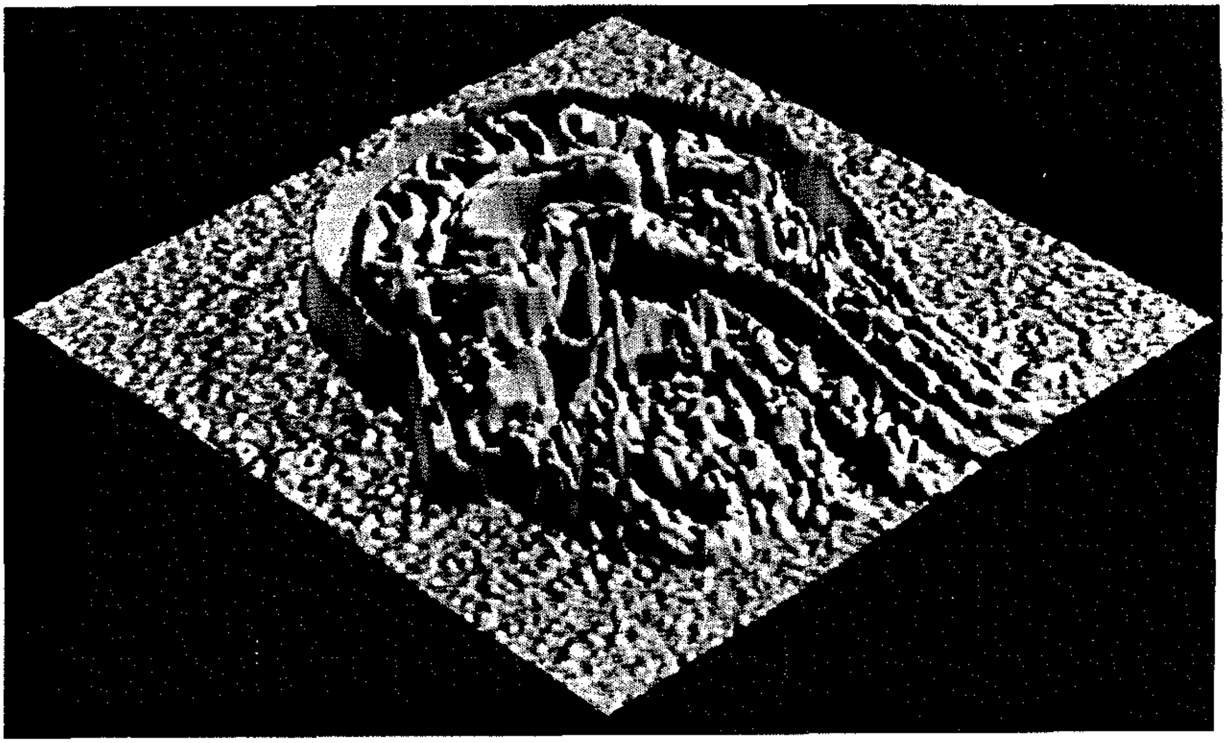


Figure 2. The MRI image of the brain of figure 1, smoothed and viewed as an intensity surface

We suggest that another basic visual property is symmetry -- regions are two sided, so that for each point on one flank of a ridge there is another point on the other flank to which it has some sort of symmetry relation. As a result we define an intensity axis of symmetry (IAS). This axis, made of a forest of branching sheets, fits under the ridges midway between the two symmetric flanks of the ridge (we use the medial axis, though other axes of symmetry are also possible). In order to avoid sensitivity to any monotonic transformation of intensity and also to avoid the incommensurability of intensity and space, the IAS cannot be defined by a symmetry operation in more than the spatial dimensions. Instead, the image (intensity surface) must be considered as a one-parameter family of slices in the intensity dimension. That is, the IAS is the one-parameter family of axes of symmetry of the intersections of the intensity surface with a series of slicing surfaces (see figure 3). Temporarily we have been slicing at isointensity levels, so that each axis of symmetry in the family is the medial axis of a level curve of intensity, even though this slicing focuses too greatly on intensity levels and too little on local image structure. The present IAS has the advantage that there is a 1-1 relation between each branch and a ridge top as defined by the locus of maxima of positive curvature of intensity level curves (these loci are called *vertex curves*) [Gauch, 1988b].

As illustrated in figure 3, associated with each point on the IAS are two points in the image where the maximal disk centered at the IAS point is tangent to a level curve at its intensity. The basic sheets that are the twigs in the IAS forest thus have associated with them a set of image points that are taken as the primitive regions of the image.

The computation of the IAS uses the fact that the medial axis of a boundary is the locus of maximum distance from that boundary. The IAS is computed by collapsing the image's intensity surface toward the locus of maximum distance from that surface. The surface is allowed to move in the spatial direction but held fixed in intensity as it is collapsed toward the IAS.

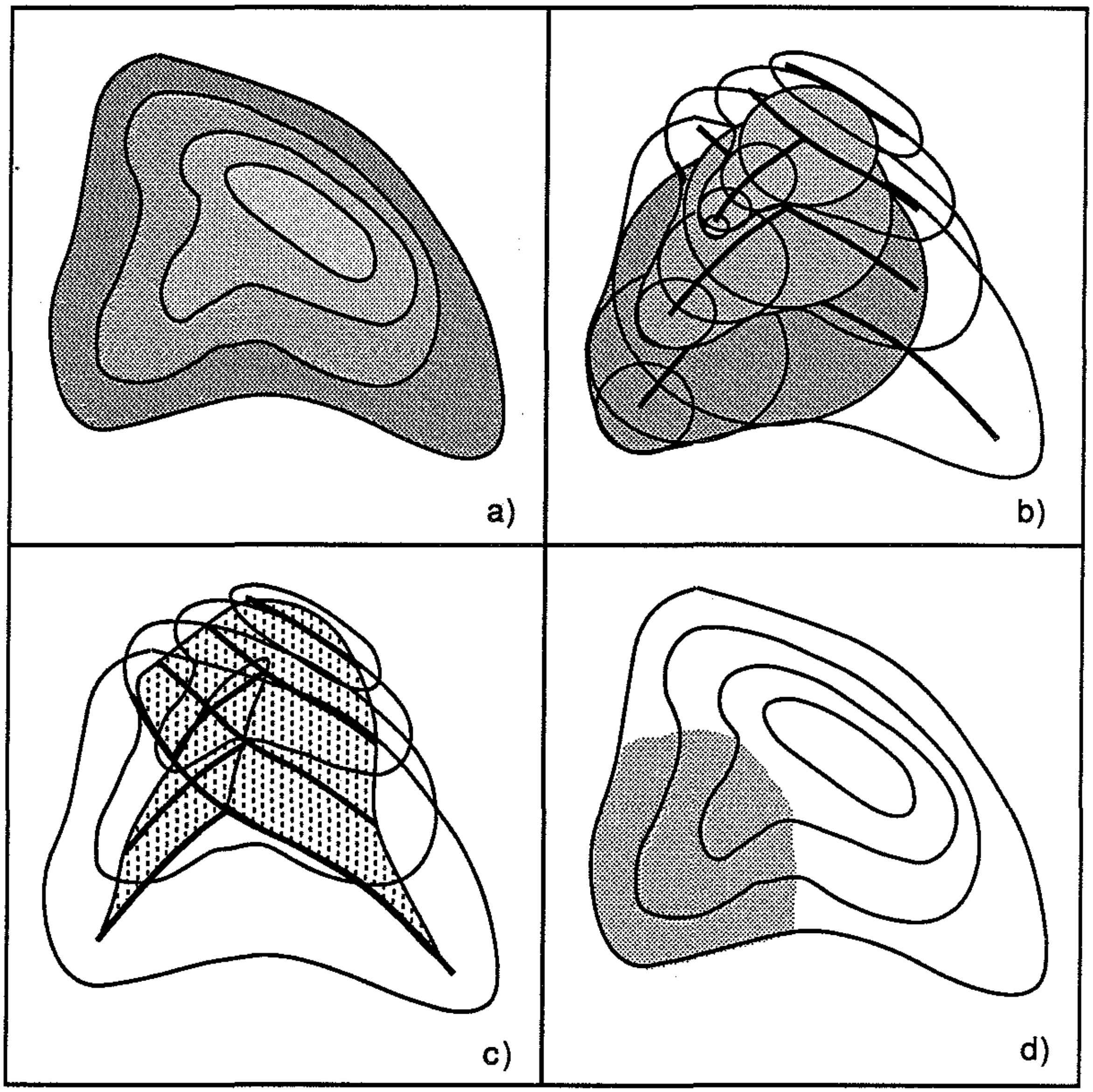


Figure 3. The IAS as a one parameter family of slice axes of symmetry. a) Level curves of a simple image; b) Level curves on intensity surface, their medial axes, and selected maximal disks; c) Level curves on intensity surface, and IAS; d) Image region associated with an IAS branch

In order to maintain surface smoothness while migrating the surface toward the IAS, we extend the concept of active contours [Kass, 1987] to handle surfaces (producing a method of *deformable surfaces*). By relaxation a minimization is obtained of the sum over surface points of a combination of the magnitudes of first and second derivatives of the deforming surface and the negative of the distance the point has moved away from the isointensity contour on which it started. The derivative constraints maintain the integrity of the surface, while the distance function drives the surface toward the IAS.

Once the surface has been collapsed to the IAS, we have a mapping from the original surface to the IAS. We next determine for each point in the image its involute, the point on the other side of the IAS that mapped to the same point on the IAS. These involute pairs are then used in a marching fashion to determine the individual branches of the IAS, and thence the region of the image associated with each branch [Bookstein, 1979]. To compute a branch, the algorithm starts with an arbitrary point on that branch, finds its involute, and then adds neighboring points if their involutes are proper neighbors of the involute of the original point. This operation continues until the current involute is not a neighbor of the previous involute (this occurs at IAS branch points). The result of this is a segmentation of the image that gives one region per branch of the IAS.

For a 256 x 256 image on a 10 MIPS, 5 MFLOPS computer (SUN 4), computing the distance function requires 5 minutes, iterative collapse of the surface requires 20 minutes, finding involutes requires 2 minutes, and identifying branches and regions requires 10 minutes.

Since surround regions for meaningful objects are frequently formed from many of the primitive regions associated with the basic IAS branches, these regions need to be grouped into larger visually sensible, coherent regions. Our strategy is to compute the region containment relations induced by the connectivity of the IAS branches and the annihilation of one branch into another as scale is increased. Since a single IAS branch can attach to more than one other branch (IAS sheets can form a loop around a pit in the intensity surface [Gauch, 1988a]), a single branch can have more than one parent, so a DAG is computed, where the parent-child relation is region containment.

Scale increase is performed by Gaussian blurring, since this form of scale increase most guarantees against the creation of structure [Yuille, 1983]. For simplicity we now use isotropic and stationary Gaussian blurring, but ultimately we expect to vary scale with a blurring kernel that is sensitive to local image structure in both orientation and size. To follow the IAS branches in scale space requires computing the IAS simultaneously at all scales. To accomplish this, we have begun to develop a method of deformable surfaces that applies to surfaces in physical space, intensity, and scale. However, the DAGs that we have computed to date are based on the more computationally efficient approximation of the vertex curves at IAS tops by watershed boundaries and of the IAS branch containment relations by those induced by watershed combination as critical points merge under Gaussian blurring.

All of the above applies not only in two spatial dimensions but also in 3D. To date we have implemented the calculations fully only in 2D. Sample results are given in figure 4.

To this point in the discussion object edges have not been relevant. The bright regions defined by IAS branches are delimited by a locus of a sort of curvature minima of the intensity surface, i.e., by the adjacent bright regions. As such they go past the steepest part of the ridge sides that might be taken to define the object edge -- for regions corresponding to objects they include part of the object surround. The IAS provides the basis for also defining an edge strength that reflects symmetry properties and for computing that edge as a continuous closed contour (in 2D) or surface (in 3D). Instead of the normal edge strength that is based on intensity steepness as the spatial arguments vary, we define the symmetry-dependent edge strength to be the steepness in intensity as the radius of the maximal disk touching the image point varies most strongly, i.e. the component of ∇I in the ∇r direction: $\nabla I \cdot \nabla r / |\nabla r|$. This is a sort of generalization to the IAS of Blum's object angle of the medial axis [Blum, 1978]. The two sides of such a symmetry-dependent edge can be computed together, and the edge is easily followed by following along the associated IAS sheet.

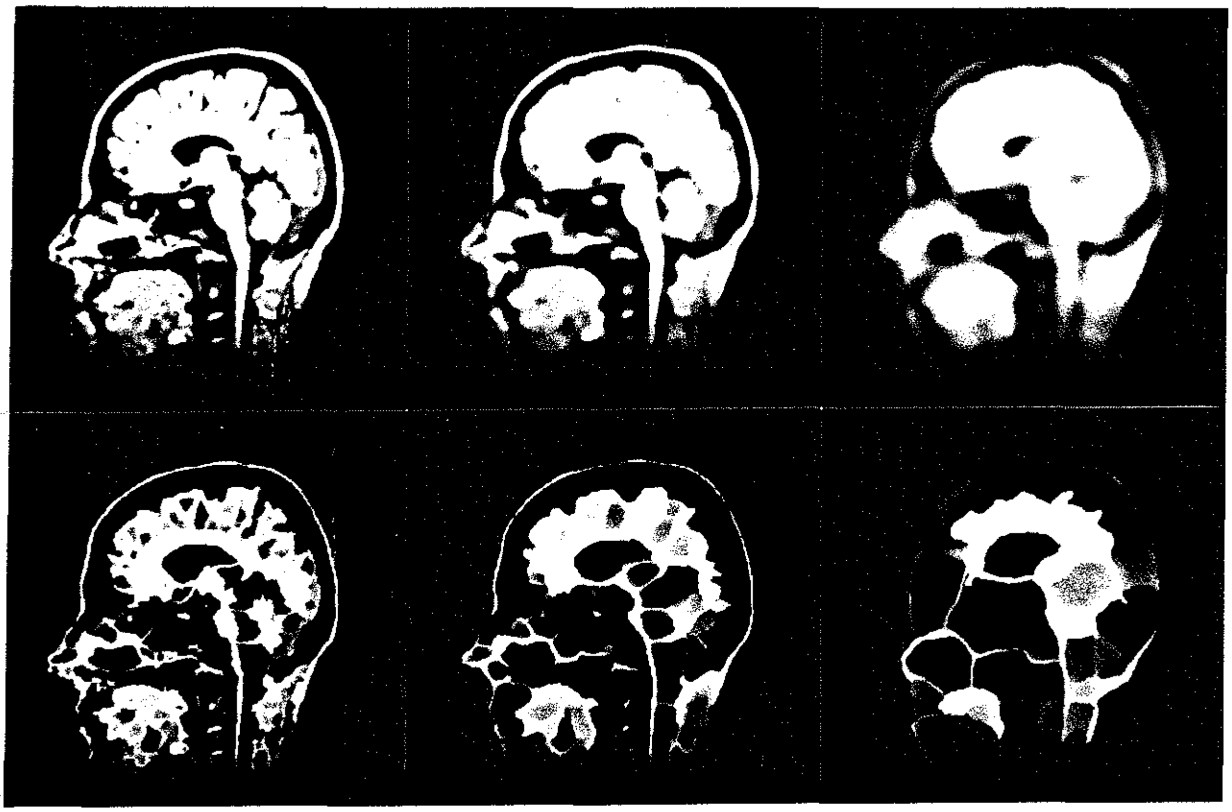


Figure 4. Multiscale sequence of 2D images and their IAS's, based on a 2D MRI head image

3. INTERACTIVE OBJECT-REGION DEFINITION

Automatic image description methods such as the multiscale IAS describe the structure of the image as a DAG of nested regions. Our approach is designed to let the human user label object surrounds of interest by creating regions based on traversals of the region hierarchy. Toward this end, we have developed a tool, the Image Hierarchy Editor (IHE), that we will describe in this section.

To define an object using IHE, the user typically begins by selecting a color for the region to be defined and then selecting a pixel in the desired region. IHE then identifies from the DAG the primitive region containing that pixel and redisplays the pixels in the region using the selected color. Usually, the primitive region is a small component of the object of interest. The user may click on a button to enlarge the labeled region by traversing up the region DAG and labeling the region corresponding to the parent (containing) node. This process may continue until the region labeled matches the desired object. At this point, the user may label this node of the hierarchy as being a "brain stem" or "cerebellum" or any other meaningful object. This region may be retrieved by name at any point in the future.

Alternatively, one can label portions of the object as separate regions and combine them by logical or set operations. While IHE provides all of the logical functions, we have found set union (logical A or B) and set difference (logical A or NOT B) to be most helpful. When a region of interest is created that does not already exist in the hierarchy, the act of labeling the region causes it to be "posted" to the region hierarchy. At present, a posted region is added to the hierarchy as a child of the root node.

The representation of the region DAG in IHE permits traversals from a pixel to its containing region and from a node in the hierarchy to all pixels contained in the node's corresponding region. Manipulation of the image using the DAG is supported by a set of region masks, that allow manual specification of region pixels, color labeling of regions defined in the DAG, and the performance of logical operations between regions.

We have developed a display protocol that permits assignment of a color label at each location while still portraying, at lower contrast, the original gray scale structure of the labeled region. Each region mask is associated with a color. In case a pixel is marked in more than one region mask, a priority ordering is imposed on the region masks; the pixel is displayed in the color corresponding to the highest-priority region mask. An enable byte permits the user to turn off region masks if desired in order to see the lower-priority colors.

We found in testing IHE that the user needed some additional information to determine whether to enlarge the current region by moving to the parent node in the hierarchy or to enlarge it by growing a new region and using logical OR operations to combine the separate regions into one object. We provide this information in a second display in which the region corresponding to the parent node is labeled in a different color. Thus, the user can see the effect of performing the "parent" operation before requesting it. This capability has proven useful since the alternative involves backtracking when the "parent" operation deviates from the desired object - a more frustrating and complex sequence of operations. This presentation is illustrated in figure 5. In the right part of the screen the image, with the selected region colored (shaded, in this grey-scale photograph), appears. In the left part of the screen the selected region and its parent are shaded with two different colors (that cannot be distinguished in grey scale).

The amount of work required to construct a region depends on the quality of the image hierarchy produced by the image description tools. We hope to minimize that work by creating more sophisticated image descriptions. In the meantime the work required to define objects of interest serves as a benchmark for how good the image descriptions are.

An example of regions selectable by these operations are given in figure 6. The tongue region selected in figure 6a required 4 parent operations. The brain cortex region selected in figure 6b required the union of 18 nodes, each requiring 2-4 parent operations. The brain stem region selected in figure 6c required 5 parent operations followed by 4 difference operations.

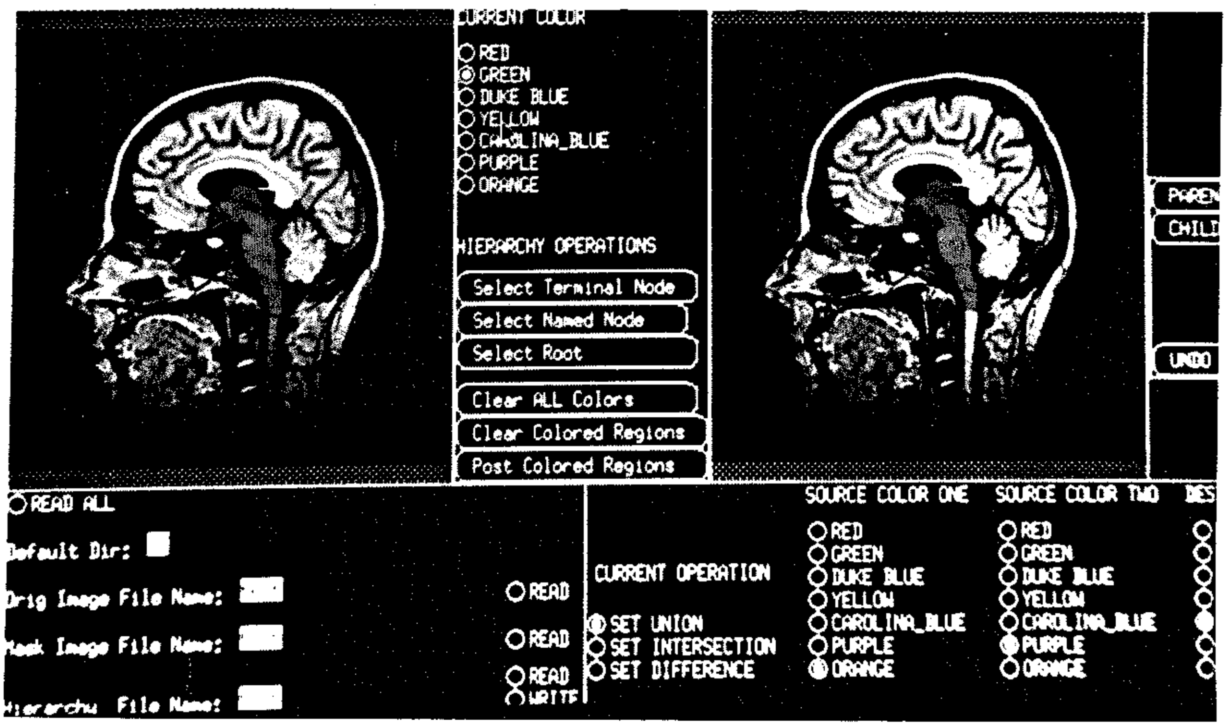


Figure 5. IHE interactive display: region selected by parent relations (darkened, top portion of brain stem in image on right) from 2D MRI head image and same including the parent of the selected region (bottom portion of brain stem in image on left)

a b c

Figure 6. Regions selected by IHE on 2D MRI head image: a) by parent relations alone; b) using unions of regions selected by parent relations; c) using differences among a region selected by parent relations and primitive region

The current implementation of IHE works only with the primitive regions defined by the image description method. Interactive pixel editing is possible within the framework of IHE but is not implemented. Pixel editing would complicate the system since it could modify leaf node structures.

4. EVALUATION

Our results to date suggest that the primitive regions, in both 2D and 3D, based on IAS branches are sensible to humans and can provide the basis for the interactive region definitions needed for display and analysis. The containing super-regions computed to date, while frequently satisfactory, fail too frequently to accord with human choices.

Despite some weaknesses in the image description DAG, our early experience with the interactive approach to defining meaningful object regions is very positive. The image hierarchy editor and its user interface provide a quick way to define object regions in medical images with only a reasonably good image description. This experience is limited to a few tens of images.

Our experience with symmetry-dependent object edges is even more limited. An example of the edge strengths computed using this measure is shown in figure 7. While this result is encouraging, it has weaknesses related to inaccuracies in the IAS on which it is based.

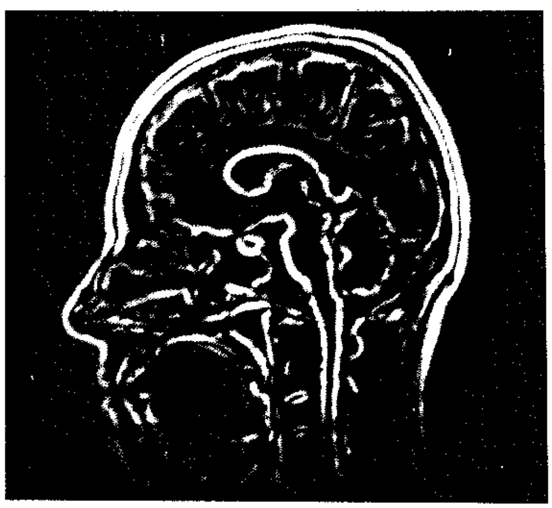


Figure 7. Symmetry-dependent edge strength for 2D MRI head image. The edge strengths shown correspond only to intensity ridges

5. SUMMARY AND FUTURE DIRECTIONS

We have discussed an approach in which human users can easily define objects and object surrounds using an interactive tool with an image description based on ridge symmetries and connectivities in the intensity surface of the image. Our initial results suggest that the idea of interactive object definition is a powerful one and that it is possible, using these geometric analyses, automatically to compute primitive image regions on which such object definitions can be based. Even the DAG based on the watershed approximation to ridge connectivities and annihilations under isotropic, stationary Gaussian blurring seems frequently to produce containment relationships that are useful in combination with the interactive object definition tools. Thus the early results seem to confirm the basic thrust of our approach. However, the methods that we have developed have a number of weaknesses, and these will require some modification in some of the structures and computing methods within our approach.

The Image Hierarchy Editor needs to be improved so that new regions defined by the user can be sensibly placed in the DAG of regions.

Other alternatives to the structure of the IAS must be considered. First, we expect that the calculation of the region parent relations using the actual IAS connectivity and annihilations, rather than the approximation based on watersheds, will give better axes. Second, some redefinition of the IAS should be found to produce a connected DAG of regions, since in a greyscale image the human viewer perceives luminance-based regions with a DAG structure. In contrast, the presently defined IAS forms a forest of disconnected components, sometimes broken at ridge ends. In fact, small scale bumps in the image may be lost in our deformable surfaces approach to computing the IAS precisely because the connectivity enforced by the computing method is not a property of the true IAS. Part of the problem will be solved with a combination of the analysis of light-on-dark with dark-on-light patches, but the solution seems also to be related to the idea that the slicing in the intensity dimension on which our method is based should be determined by the local intensity context, perhaps by the local symmetry information. We expect to investigate transformations of intensity according to local symmetries.

A related issue is the inadequacy of a form of scale change that is independent of local image structure. Some research [e.g., Grossberg, 1985; Koenderink, 1984, 1989] suggests that human vision involves a diffusion of intensity attenuated by edges or local objects. It is known that nonstationary, nonisotropic diffusions retain the guarantees against structure formation available with stationary diffusion [Yuille, 1983; Lifshitz, 1987]. Investigations of ways of generating sensible IAS annihilations by scale-increasing intensity diffusions that are affected by local edge strengths are underway. Means of calculating edge strengths that are related to ridge symmetry will therefore be especially important, providing impetus for the needed development of a multiscale theory of edges based on ridge symmetry. Among the tasks awaiting us are detailed comparison of our symmetry-dependent edge strengths to one-sided edge strengths, the development of a means of computation of a symmetry-dependent closed edge (perhaps by finding the locus of maxima of our symmetry-dependent edge strength along integral curves of IAS radius gradient, as in [Canny, 1983]), and the inclusion of the scale under attention as a parameter of the edge.

Both the need to compute nonstationary intensity slicing and nonstationary scale increase, as well as the need to compute IAS's simultaneously at all scales and intensities leads to a major concern for computing time. Even though the image analysis step is completely automatic and only the interactive object region definition based on this analysis need be done in real time, the analysis computations today in three spatial dimensions plus intensity require tens of hours on a 10 MIPS computer. Adding the dimension of scale and making the two above-mentioned operations nonstationary suggest that only a connectionist scheme will ultimately allow the required computations to be done in a reasonable time. Nevertheless, the attractiveness of the underlying theory leads us to develop the approach and ultimately to find forms of the computations that can be done in a few minutes, even for three-dimensional images.

Finally, the results of our methods based on geometric analysis and annihilations under scale increase must be compared to other multiscale methods with attractive possibilities, including the pyramid-based methods [Burt, 1981; Meer, 1988; Rosenfeld, 1984] or methods based on the sequence of primitive template match strengths across scale, including those based on successive Gaussian templates [Stiehl, 1989; Blom, 1988], those based on Gaussian and orientation templates [Coggins, 1986], and those based on Difference of Gaussian templates [Crowley, 1984].

ACKNOWLEDGEMENTS

We gratefully acknowledge the collaboration of Drs. Fred Bookstein, Robert Gardner, Jan Koenderink, and Bart ter Haar Romeny on the theory of multiscale IAS representations. We thank Dr. Bookstein for his comments on this paper and the UNC Med3D 3D display team, including its leader, Dr. Henry Fuchs, and the designer of our volume rendering methods, Dr. Marc Levoy, for their support and stimulation. We thank Carolyn Din, for help in manuscript preparation and Bo Strain for photographic assistance. The research reported herein was carried out with the partial support of NIH grant number P01 CA47982 and the NASA Center of Excellence in Space Data and Information Science.

REFERENCES

Blom, J, ter Haar Romeny, BM (coaches), Bookelmann, MS, "Edge Behaviour in Scale Space." Internal Report, Dept. of Radiology, State Univ. of Utrecht, 1988.

Blum, H, and Nagel, RN, "Shape Description Using Weighted Symmetric Axis Features." *Pattern Recognition* **10**:167-180, 1978.

Bookstein, FL, "The Line-Skeleton." *Comp. Graphics and Im. Proc.* **11**(2):123-137, 1979.

Burt, P, Hong, TH, Rosenfeld, A, "Segmentation and Estimation of Image Region Properties Through Cooperative Hierarchical Computation."*IEEE Trans SMC* **11**:802-809, 1981.

Canny, JF, "Finding Edges and Lines in Images." Tech Report #720, Artificial Intell. Lab., MIT, Cambridge, MA, 1983.

Coggins, J, Fay, F, Fogarty, K, "Development and Application of a Three-dimensional Artificial Visual System." *Comp. Meth and Prog. in Biomed.* : 69-77, 1986.

Crowley, J & Parker, A, "A Representation for Shape Based on Peaks & Ridges in the Difference of Low-Pass Transform." *IEEE Trans. PAMI* **2**(2):156-169, 1984.

Fuchs, H, Pizer, SM, Creasy, JL, Renner, JB, Rosenman, JG, "Interactive, Richly Cued Shaded Display of Multiple 3-D Objects in Medical Images." *Medical Imaging II*, SPIE **914**, Part B: 842-849, 1988.

Gauch, J, Oliver, W, Pizer, S, "Multiresolution Shape Descriptions and Their Applications in Medical Imaging." *Information Processing in Medical Imaging* (IPMI X, June 1987): 131-150, Plenum, New York,1988a.

Gauch, J and Pizer S, "Image Description via the Multiresolution Intensity Axis of Symmetry." *Proc. 2nd Int. Conf. on Comp. Vis.*, IEEE Catalog #88CH2664-1:269-274, 1988b.

Grossberg, S and Mingolla, E, "Neural Dynamics of Perceptual Grouping: Textures, Boundaries, and Emergent Segmentations." *Perception and Psychophysics* **38**(2):141-171, 1985.

Kass, M, Witkin, A, and Terzopoulos, D, "Snakes: Active Contour Models." *Proc 1st Int. Conf. on Comp. Vis.*, IEEE Catalog #87CH2465-3:259-268, 1987.

Koenderink, JJ., *Solid Shape*. MIT Press, Cambridge, 1989.

Koenderink, JJ "The Structure of Images." *Biological Cybernetics* **50**:363-370, 1984.

Levoy, M, "Display of Surfaces from Volume Data." *IEEE Computer Graphics and Applications* **8**(3):29-37, 1988.

Lifshitz, LM, *Image Segmentation Using Global Knowledge and A Priori Information.* Ph.D. Dissertation, Tech. Report 87-012, UNC Chapel Hill, 1987.

Meer, P, Jiang, S-N, Baugher, ES, Rosenfeld, A, "Robustness of Image Pyramids under Structural Perturbations." *Comp. Vis., Graphics, and Image Proc.* **44**(3):307-331, 1988.

Rosenfeld, A, *Multiresolution Image Processing and Analysis*. Springer-Verlag, Berlin, 1984.

Stiehl, HS, Neumann, H, Back, S, "On Scale-Space Edge Detection in Computed Tomograms." In this volume, 1989.

Yuille, AL and Poggio, T, "Scaling Theorems for Zero-Crossings." A.I. Memo 722, MIT, 1983.