Interactive 2D and 3D Object Definition in Medical Images Based on Multiresolution Image Descriptions

February, 1988

Technical Report 88-005

Stephen M. Pizer, John M. Gauch, and Lawrence M. Lifshitz

The University of North Carolina at Chapel Hill
Department of Computer Science
Sitterson Hall, 083A
Chapel Hill, NC 27599-3175

To appear in SPIE, 1988.

UNC is an Equal Opportunity/Affirmative Action Institution.

Interactive 2D and 3D Object Definition in Medical Images Based on Multiresolution Image Descriptions

Stephen M. Pizer[1,2], John M. Gauch[1], Lawrence M. Lifshitz[3]

Departments of Computer Science[1] and Radiology[2]
University of North Carolina
Chapel Hill, North Carolina, USA
and
School of Medicine[3]
University of Massachusetts
Worcester, Massachusetts, USA

ABSTRACT

We present means of interactive definition of anatomic objects in medical images via a description of the image in terms of visually sensible regions. The description is produced by computing structures capturing image geometry and following them through the image simplification of Gaussian blurring. In particular, we suggest that the structure made from intensity "ridge" and "course" curves defined by the locus of intensity level curve vertices, augmented by the pile of internal and external symmetric axes of these level curves, satisfies desirable criteria for a structure on which to base such object definition.

1. INTRODUCTION

Analysis of properties, such as volume and shape, of anatomic objects in medical images require first that the objects be defined. Also, all 3D display methods benefit from, and most such methods require, definition of the objects to be displayed. Clinical usefulness depends on these object definition methods being both accurate and quick. We report some progress in developing such methods.

Object definition in images requires an intelligent understanding of the images. The source of intelligence, be it a human interacting with the display of the image or a computer program exhibiting artificial intelligence, fits its model of the world to the image information in order to define or recognize an object. An image description in terms of a hierarchy of visually sensible regions can provide an important basis for this object definition process. In this paper we first present a means of producing such image descriptions and then lay out methods for their interactive use by a human user to define image objects quickly.

The methods that we discuss are applicable to images of any number of dimensions, though we will discuss only images of two and three spatial dimensions. In dimensions higher than two they operate directly in that space and not slice by slice.

2. MULTIRESOLUTION IMAGE DESCRIPTION VIA ESSENTIAL STRUCTURE ANNIHILATION

A major advance in the study of images has been the realization that they simultaneously represent information at many levels of scale [Burt, 1983; Robson, 1983; Crowley, 1984; Koenderink, 1984; Rosenfeld, 1984]. That is, an understanding of the image requires that global (large scale) properties be combined with more local (smaller scale) properties. An image description in terms of visually sensible image regions should therefore be created by viewing the image at multiple scales.

Our fundamental approach for providing such an image description is to represent the image by a structure that captures essential image information and then define a hierarchy of components of that structure by the order of annihilation of those components as the image is continuously simplified by lowering the scale. The information-capturing "essential structure" is chosen so that image regions are associated with each structure component during the image simplification. Objects can then be defined by taking unions of selected regions.

The idea is that image objects are defined first by regions of large scale, with detail of these objects defined by regions of smaller scale. Regions of large scale are those that are retained as the image is simplified by reducing resolution (blurring), while small-scale regions disappear under less blurring. The small-scale regions define themselves as components of larger-scale regions by blurring to become part of them. Image regions are thus defined by the annihilation of their essential structure components under blurring, or to take a more constructive point of view, by the creation of these components as deblurring is successively applied to the fully blurred image.

To guarantee image simplification [Witkin, 1983; Yuille, 1983; Koenderink, 1984, 1988], successive Gaussian blurring is chosen as the means of scale lowering. The avoidance of local creation of new values of any linear function of derivatives of the image, as the blurring proceeds, is retained even when the Gaussian blurring is non-isotropic or non-stationary [Lifshitz, 1987a]. The variation of the parameters of the blurring Gaussian across the image could be used to reflect *a priori* or tentative knowledge about the scene.

An essential structure should be an image descriptor that has the following five properties:

1. It induces a subdivision of the image into regions.
2. It captures essential region properties, including the way intensity varies across it and the spatial properties of the region, i.e., its shape, and therefore the regions it induces are semantically sensible.
3. The structure relating image components does not change until a component annihilates.
4. It induces a hierarchy of regions by defining for each component the containing component into which it annihilates.
5. It is applicable for images of any spatial dimension.

3. AN EXTREMA-BASED ESSENTIAL STRUCTURE

Based on the idea of Koenderink [1984], we began by choosing as our essential structure the set of intensity extrema, augmented by iso-intensity contour segments [Lifshitz, 1987a,b]. As illustrated in Figure 1, the technique is to follow *extremal paths*, i.e. the tracks of each extremum as image blurring increases and the extremum intensity changes monontonically, until the extremum annihilates with a saddle point. *Iso-intensity paths*, i.e., paths connecting each image location to the closest location at the next higher scale with the same intensity, are also followed until they run into an extremal path.

This essential structure satisfies criteria 1, 4, and 5, and fails criterion 3 only in an unimportant way by allowing extremum creation, but it fails in an important way on criteria 2.

Region Definition. Image regions are defined by associating pixels with extrema according to the extremal paths into which their iso-intensity paths run.

Region Hierarchy. The regions formed by linking iso-intensity paths to extremal paths are associated as subregions of other extremal regions according to the behavior of the iso-intensity paths which begin where an extremum annihilates. The extremal path into which this new iso-intensity path eventually links identifies the parent region for the subregion (for example, see Figure 1).

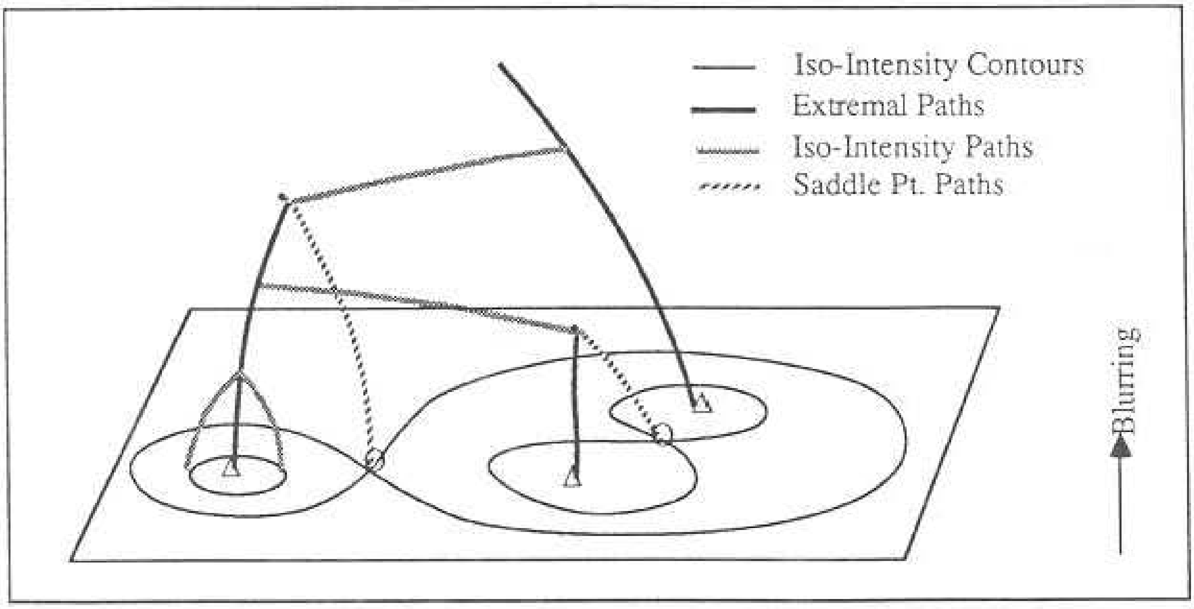

Figure 1: The behavior of extremal paths under resolution reduction. Note that maxima (Δ), and saddle points (O) move together and annihilate. The resulting non-extremal point (•) is then linked via an iso-intensity path to another extremal path.

Generalization to all dimensions. Lifshitz has demonstrated the method for 2D and 3D images and has described how the image representation can be extended to any number of dimensions.

Region sensibleness. Applying these ideas to 2D and 3D medical images, we have shown that the regions in the description thus produced frequently form anatomic objects, or can be easily formed into such anatomic objects using the operations of union and difference (for example, see Figure 2). However, sometimes the resulting regions are not semantically sensible.

First, an extremal region may be made of disconnected components, with neither component itself being an extremal region. If these two components are separated by another region with which each component has a visually apparent edge, we would wish that each edge would form the boundary of an extremal region, but this does not always happen (see Figure 3). This misbehavior results from the fact that the edges of regions may be heavily blurred before either of their iso-intensity sets run into extremal paths, so both sets can run into the same extremal path and thus form parts of the same region.

Second, regions which obviously hang together as a single object do not always combine into single regions in the tree. This problem is especially apparent with long branching regions, such as a blood vessel tree.

We suggest that these difficulties arise from the fact that the essential structure of intensity extrema and iso-intensites inadequately reflects spatial shape and edges. If we view the image as a terrain map, with intensity as height, edges can also be considered as shape features. We thus consider multiresolution description of the shape of such terrain surfaces.

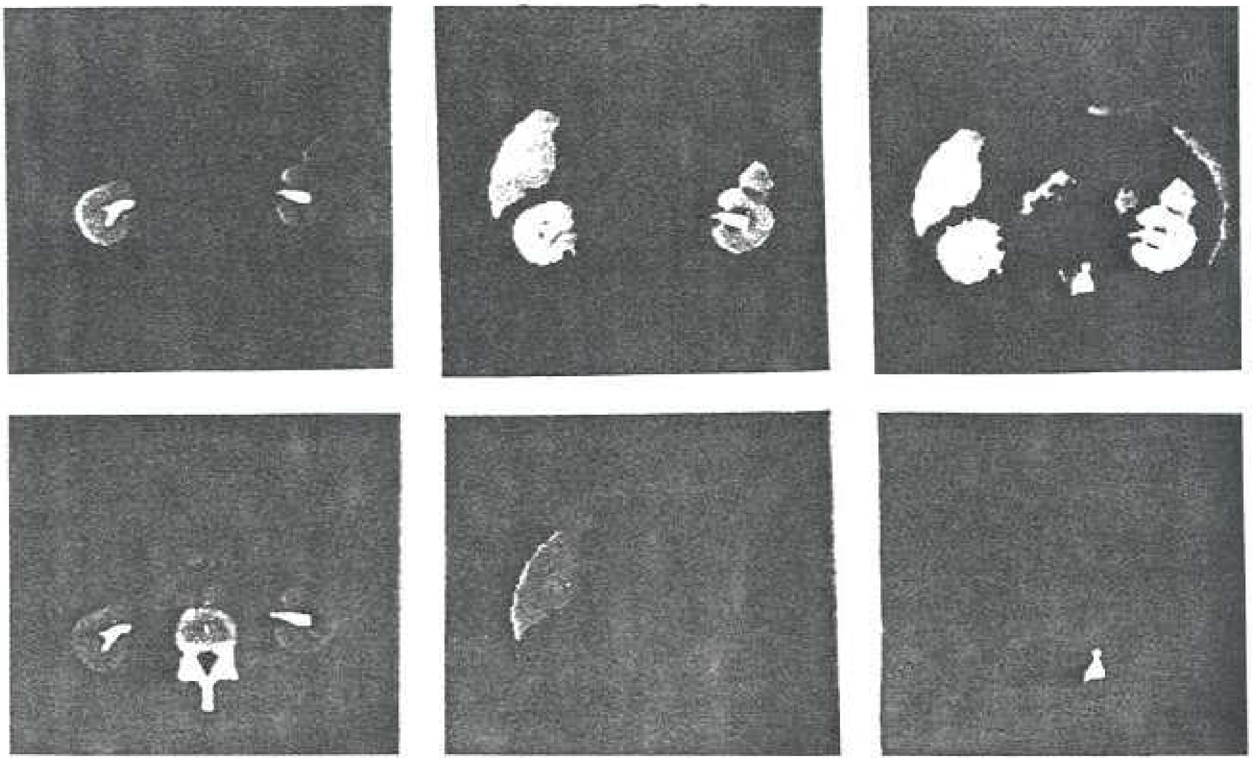

Figure 2: A CT slice through the upper abdomen together with a collection of anatomic regions automatically defined using Lifshitz's program.

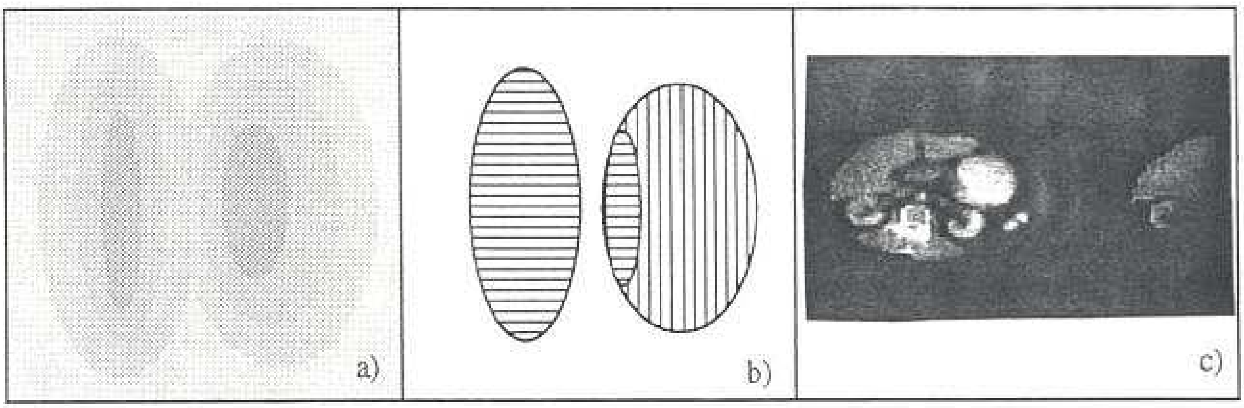

Figure 3: The halftone image in 3a depicts the situation where two dark regions in an image are separtated by a narrow lighter region. The two image segments in 3b illustrate how pixels linking to local extrema can sometimes result in unnatural image segments. In 3c we see an example where Lifshitz's program has identified part of a kidney and the whole liver as a single object in an abdominal CT image.

4. A SHAPE-BASED ESSENTIAL STRUCTURE

Since the intensity dimension is incommensurate with the spatial dimensions, we must treat height specially. At the same time, we would like to make use of structures designed for the spatial shape description of binary objects. We therefore view the terrain map as being made up of a continuous pile of binary images, each corresponding to a (not necessarily planar or level) slice through the terrain, and having value 0 where there is air and value 1 where there is earth. The structures defining the objects on each slice can themselves be thought of as piled on each other in the intensity direction. Such piles ("one parameter families") of slice structures form an attractive shape based essential structure.

The appropriate cut surfaces may be image dependent (see Figure 4); how to choose them is a subject of research, and we avoid that question for now by allowing only the selection of intensity level sets as the cut surfaces after preprocessing of the image (e.g. by a locally adaptive contrast enhancement [Pizer, 1987b]).

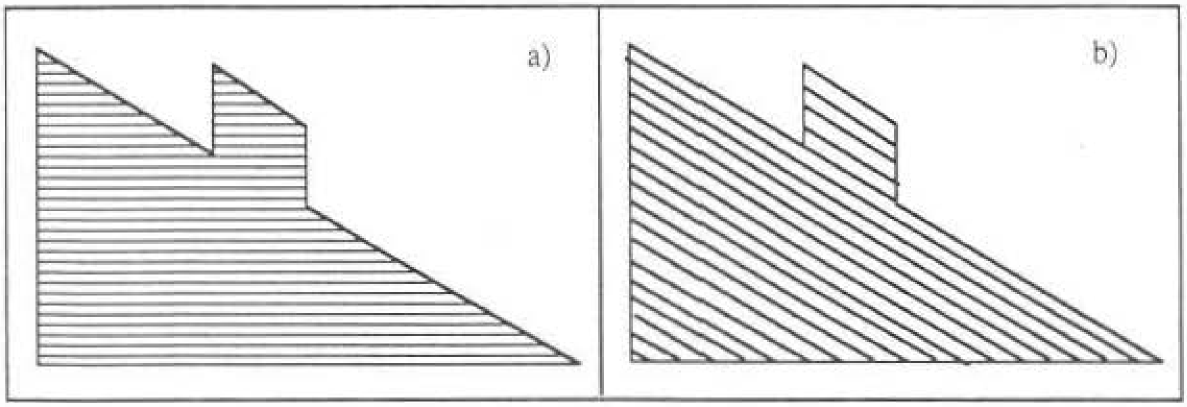

Figure 4: Intensity profile of an image with a raised area on a ramp. The iso-intensity slices in figure 4a may be less natural for describing the image than the inclined slices in 4b.

A shape descriptor whose pile satisifies our criteria for behavior under Gaussian blurring is based on the symmetric, or medial, axis (SA) [Blum, 1978]. The internal SA of a figure is defined as the locus of centers of maximal spheres (disks, in 2D) inside the figure (see Figure 5a). The internal SA of a connected figure is a connected tree that by division at branch points induces a decomposition into regions. For objects without holes, each region has an unbranching axis and two associated boundary sides. When the SA component includes an SA endpoint, the two sides meet at a point of maximal boundary curvature. The axis is the center of a figure and its associated radius function specifies the locations inside the figure. The external SA of a figure is simply the internal SA of the complement of the figure.

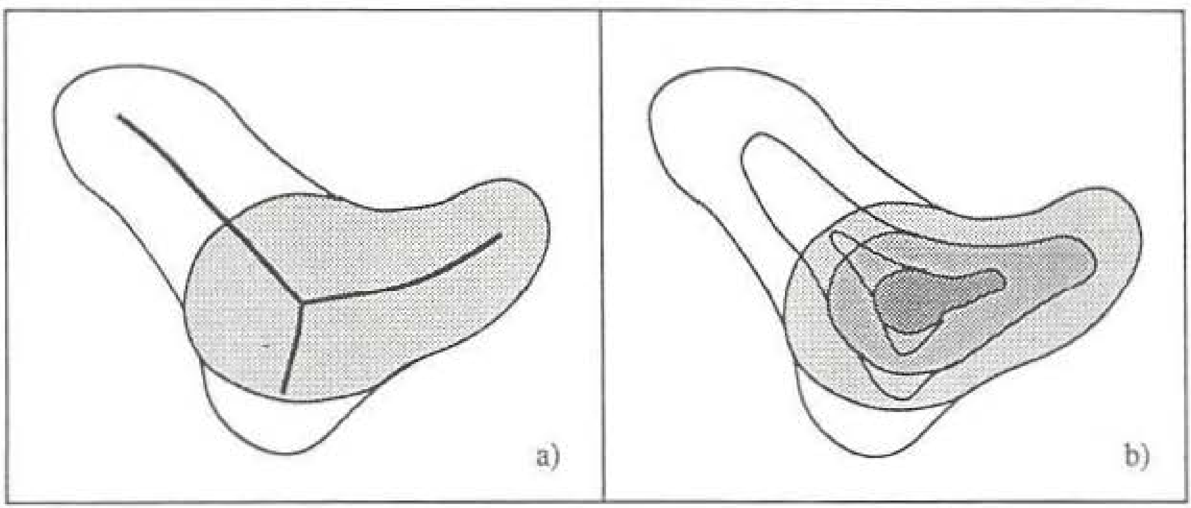

Figure 5: The region associated with an SA branch in 5a is the union of all maximal circles centered on the axis branch. For each SAP sheet in 5b, we combine the SA regions associated with each intensity level to obtain an intensity volume.

The internal SA pile (SAP) captures light objects on dark backgrounds. It is made by piling the internal SAs for earthen regions: the x,y at each terrain map level L such that $I(x,y) \geq L$. The external SAP captures dark objects on light backgrounds. It is made by piling the SAs of the air: x,y such that $I(x,y) \leq L$. Gauch [1987] has shown that each SAP consists of branching sheets, each such branch characterizing shape in both space and intensity of a corresponding part of the image. The internal and external SAPs touch, orthogonally, at intensity saddle points. Furthermore, branches shrink to annihilation under Gaussian blurring of the image I (see Figure 6). A hierarchy of SAP sheets is thus induced. Corresponding to each sheet in the hierarchy is a radius function on the sheet, and a region image R for a sheet is defined by $R(x,y)$ = the maximum intensity level for which the radius function of a sheet point at that level includes x,y (see Figure 5).

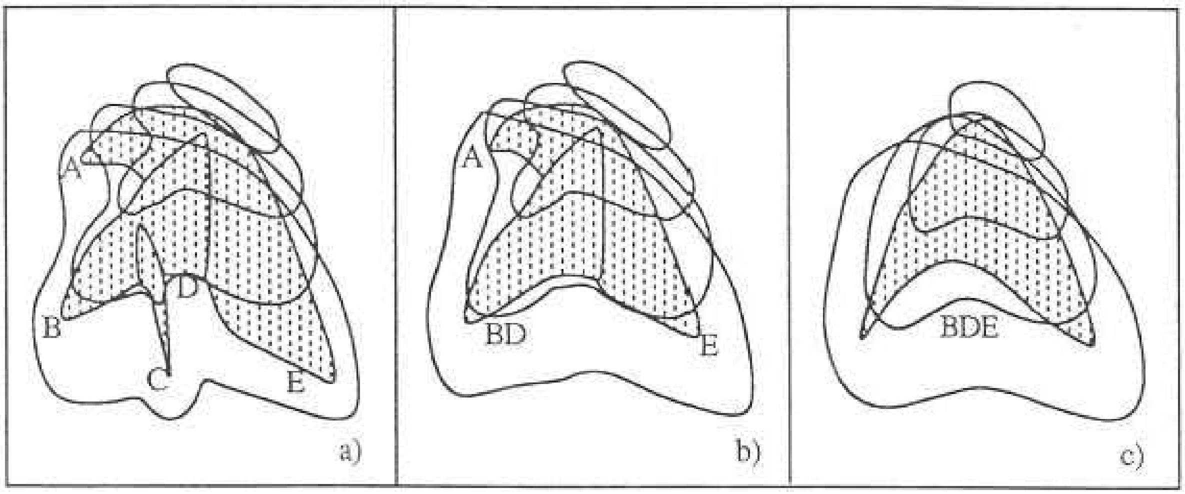

Figure 6: The effects of resolution reduction on the SAP of figure 6a are shown in 6b and 6c. When branch 'C' annihilates, we identify it as a subobject of the combined branch 'BD'. Similarly, branch 'A' is determined to be a sub-branch of 'BDE'.

The hierarchy induced by Gaussian blurring prevents detail, or noise, from destroying the naturalness of the decomposition that the SAP induces [Pizer, 1987a]. SAP branches corresponding to detail shrink and annihilate early into the limbs off which they branch, so the regions corresponding to these branches are defined as subregions and, more importantly, the limbs are restored to the natural correspondence with a single region rather than two regions interrupted by a detail. Such a hierarchical description applies not only to images but also to simple figures, when they are represented by characteristic functions.

The SAP for an n-dimensional image is an n+1-dimensional tree of sheets, a prodigious object to follow through image blurring. However, Gauch has pointed out that since each SA is terminated by a boundary point of maximum curvature magnitude (a *vertex*), each SAP sheet is terminated by a curve of level curve vertices (a *vertex curve* – see Figure 7). These vertex curves are simply tracks in the original image, corresponding to ridges or courses in the terrain map. They can be followed through image blurring, and when a vertex curve annihilates, the SAP sheet that it terminates must also annihilate. Therefore, it is possible to compute the SAP only for the original image, and for each vertex curve annihilation to follow the corresponding SAP sheet to its branch curve. The sheet defines a region image R and specifies R as a subregion of the region image corresponding to the limb sheet into which it connects.

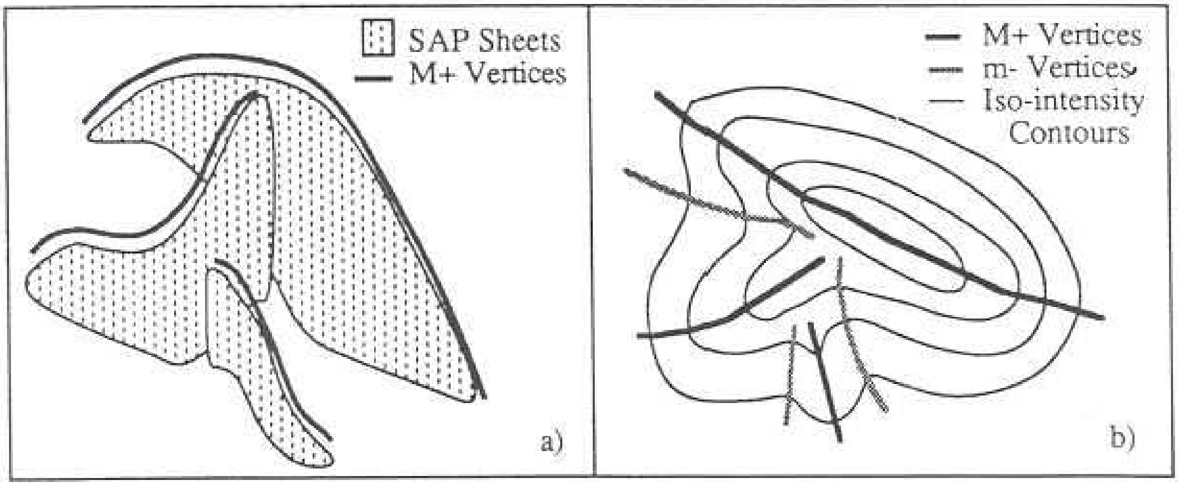

Figure 7: a) The relationship between the symmetric axis pile for an image and the vertex curves corresponding to the end curves of the 'earthen' SAP. For clarity, only the vertex curves corresponding to positive curvature maxima (M+), are shown. b) Iso-intensity contours and vertex curves corresponding to positive curvature maxima (M+) and negative curvature minima (m-).

Figure 8 shows level curve curvature and vertex curves for various degrees of Gaussian blurring of a digital subtraction angiogram, an image of blood vessels and also for an abdominal CT image. In the curvature images high positive curvature (M+) is shown as white, and low negative curvature (m-) is shown as black. The grey points on white and black curves correspond to saddle points in the image. Only M+ and m- vertex curves can be the tops of SAP sheets, and these SAP sheets can help in following the vertex curves across saddle points. The vertex curves move continuously to annihilation and induce the hierarchical description described in the previous paragraph.

The level curve curvature K at each image point is computed as $K = v^t\ hessian(I)\ v$, where I is the image and v is the unit vector in the direction of the level curve tangent, $(-\partial I/\partial y,\ \partial I/\partial x)$. The K values were computed by the multiresolution n-jet approach of Koenderink [1986]. This approach involves computing $\partial^n I/\partial x^m \partial y^{n-m}$ for all n less than some limit, all $m \leq n$, and all degrees of blurring. From these, many feature values, including level curve curvature, Laplacian value, etc. can be easily computed. Listing useful essential structures that can be computed in this way ought to be the subject of active research.

Vertex curves togther with the SAP of the original image seem to satisfy all of the criteria specified for an essential structure.

Region definition. It induces a subdivision into regions that also carry information on intensity variation.

Region sensibleness. The fact that it is based on ridges and courses seems to allow object curving to be followed and objects not to break into unrelated pieces. Like the essential structure of intensity extrema augmented with iso-intensity contour segments, the vertex curve / SAP structure captures the behavior of critical points, but it is more oriented to a whole object rather than one point.

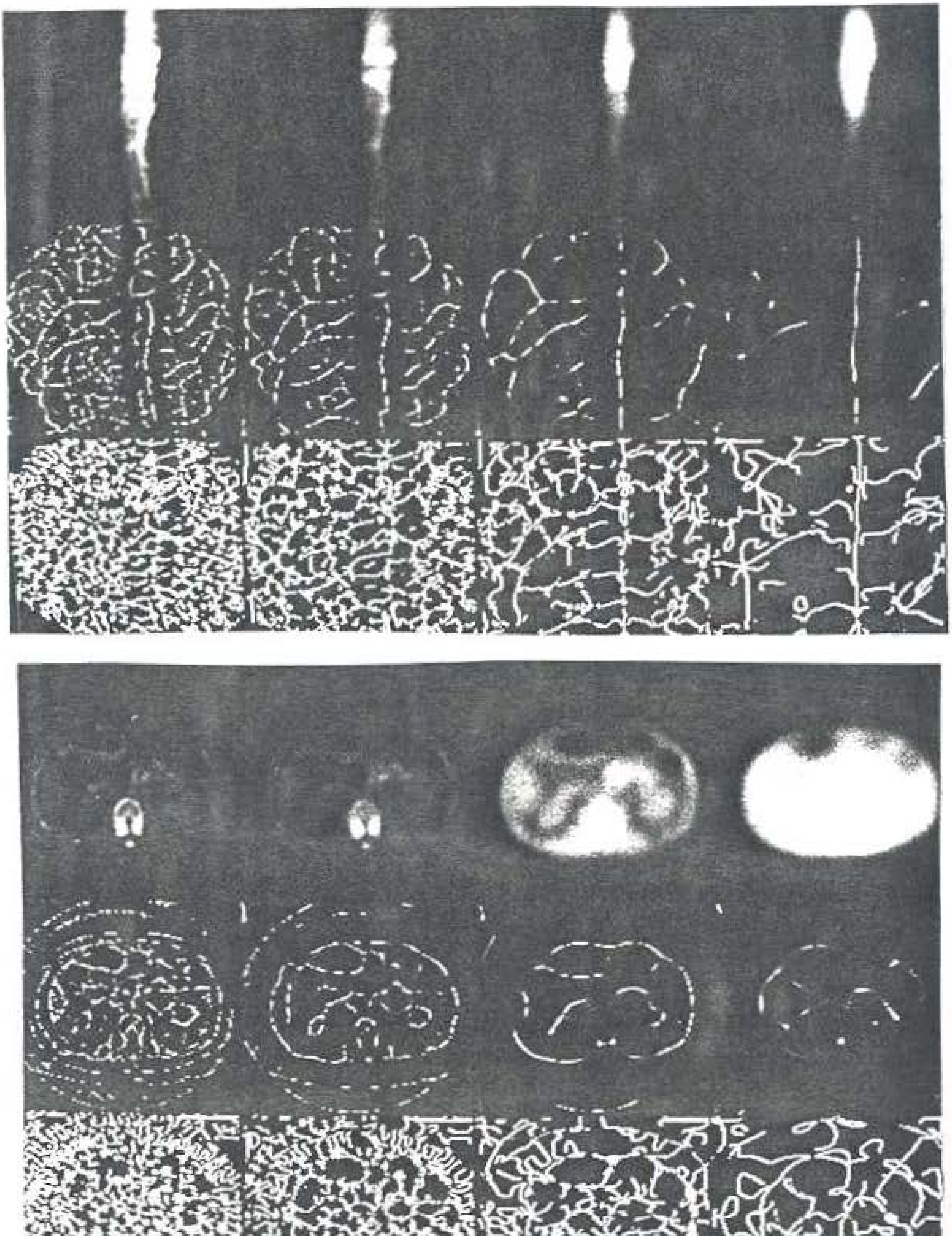

Figure 8: A sequence of blurred digital subtraction angiogram images (top row) with their corresponding level curve curvature (second row) and vertex curve images (third row). Rows four, five and six show analagous sequences for an abdomen CT image.

Consistent simplification. Under image blurring no new values of level curve curvature are created, but the topology of the associated vertex curves can change. These changes occur when saddle-extremum pairs annihilate (or form) and also when locally concave or convex regions on the side of hills and valleys are destroyed. By following the smooth evolution of vertex curves, the simplification of SAP structure can be deduced.

Region hierarchy. The hierarchy induced by image simplification involves only a selection among branch sheets of the SAP which are already in the form of a tree (or a forest of trees). Furthermore, the regions they induce are directly described in terms of intensity and spatial shape by the properties of the symmetric axis transform.

Generalization to all dimensions. The method seems extendable to higher dimensions, though details need to be investigated.

The vertex curves / SAP essential structure thus seems quite promising. However, the usefulness of this description and some of its mathematical properties are still under investigation. Furthermore, the dependence on intensity level curves seems unfortunate, and improved means of slicing the image surface need to be developed.

5. INTERACTIVE OBJECT DEFINITION

The production of an image description by the above-mentioned approaches is a completely automatic process. The result is a set of image regions and a hierarchical relation among them. Even for the simple extremum-based essential structure the regions in the image description frequently correspond to semantically meaningful image objects. For the vertex-curve-based structure, we expect this to occur even more frequently. For these regions object definition can be accomplished by a user simply by pointing to a pixel (voxel) in the region and displaying the resulting region next to the original image, for user verification (see Figure 2). We have shown this operation to be computable at interactive speeds. If the region displayed is a subregion of the desired object, a simple button push can cause the next larger containing region to be selected and displayed, also at interactive speeds. During image description each region can be labeled by its scale (degree of blurring to achieve annihilation) or intensity (of a critical point at annihilation). This allows the selection of a region by a scale window or an intensity window, an operation that has appeared useful. Selection could also be based on other region properties, such as area, mean intensity, and intensity variance.

Since these regions in the image description are defined only by image intensities, they cannot be expected always to reflect semantic information. Some means will be necessary either to edit the resulting descriptions, especially by moving regions within the hierarchy, or to edit objects defined from these descriptions, as follows. For the extremum-based descriptions we have found that operations of region union and difference, combined with the ability to divide a region into two by painting out a few connecting pixels, were adequate to define most objects of anatomic interest. Another operation that may sometimes be helpful is to change which of the branches intersecting at a specified branch point form the limb and which forms the branch. We must develop a method to allow such editing without requiring the user to comprehend the image description hierarchical structure for his particular image.

The definition of late annihilating (important) regions sometimes suffers from the fact that they include their subregions by definition, despite the fact that these subregions are not part of the corresponding semantically meaningful object. Means for subtracting subregions en masse, e.g., by the blurring level at which they annihilated, have proved necessary to obtain adequately fast operation.

A final tool necessary for the application of this interactive object definition approach to 3D images is a means of displaying a selected 3D region and the original 3D image data, for user verification of the region. The volumetric rendering work by Levoy [see this volume] and others seems to hold promise in this regard.

6. SUMMARY

We have shown that describing images hierarchically by following essential structures to annihilation is attractive if the essential structures satisfy a number of criteria. The idea can be applied to a wide range of essential structures. However, the vertex curve / SAP essential structure seems particularly attractive in meeting all of the criteria. Other structures based on geometrical features of the intensity surface might also have these strengths. Interactive object definition methods that operate in a few seconds on a 1 MIPS machine have been demonstrated.

This paper has left many open directions for exploration, including how cuts through terrain images should be made, how useful the vertex curve-based descriptions will be, and what other essential structures ought to be investigated. Means for 3D display to allow a user to compare 3D regions selected from an image description to the original image data need to be developed. We are confident that such research will lead to the production of object definitions in user times orders of magnitude faster than is now possible. Work on extending the ideas of this paper to vector-valued images or time series of images would also seem beneficial.

7. ACKNOWLEDGEMENTS

We are grateful to Jan Koenderink and William Oliver for useful discussions, to Bart ter Haar Romeny for collaboration on the implementation of n-jet-based methods for computing essential structures, and to Sandra Bloomberg for programs for computing symmetric axes. We thank Bo Strain and Karen Curran for photography and Sharon Laney for help with manuscript preparation. The research reported here was carried out under the partial support of NIH grant # R01 CA39060.

8. REFERENCES

1. Blum, H., and Nagel, R.N., "Shape Description using Weighted Symmetric Axis Features", *Pattern Recognition*, Vol. 10, pp. 167-180, 1978.

2. Burt, P.J., and Adelson, E.H., "The Laplacian pyramid as a compact image code", *IEEE Trans. on Communications*, Vol. 31, No. 4, pp. 532-540, 1983.

3. Crowley, J., and Parker, A., "A Representation for Shape Based on Peaks and Ridges in the Difference of Low-Pass Transform", *IEEE Trans. PAMI*, Vol. 6, No. 2, pp. 156-169, March 1984.

4. Gauch, J.M., Oliver, W.R., Pizer, S.M, "Multiresolution Shape Descriptions And Their Applications In Medical Imaging", Tech. Report 87-018, University of North Carolina - Chapel Hill, to appear in *Information Processing in Medical Imaging*, 10th Conf. Utrecht, Plenum, 1987.

5. Koenderink, J.J., "The Structure of Images", *Biological Cybernetics*, Vol. 50, pp. 363-370, 1984.

6. Koenderink, J.J, and van Doorn, A.J., "Representation of Local Geometry in the Visual System", *Biological Cybernetics*, Vol. 53, November, 1986.

7. Koenderink, J.J, "Operational Significance of Receptive Field Assemblies", *Biological Cybernetics*, Vol. 54, January, 1988.

8. Levoy, M.S., "Alias Free Rendering of Surfaces from Computed Tomography Data", See this volume.

9. Lifshitz, L.M., *Image Segmentation using Global Knowledge and A Priori Information*, Ph.D. Dissertation, UNC Chapel Hill Technical Report 87-012, 1987a.

10. Lifshitz, L.M., "A Multiresolution Hierarchical Approach to Image Segmentation based on Intensity Extrema", Tech. Report 87-019, University of North Carolina - Chapel Hill, to appear in *Information Processing in Medical Imaging*, 10th Conf. Utrecht, Plenum, 1987.

11. Pizer, S.M., Oliver, W.R., Bloomberg, S.H., "Hierarchical Shape Description Via The Multiresolution Symmetric Axis Transform", *IEEE Trans. PAMI*, Vol. 9, No. 4, pp. 505-511, 1987a.

12. Pizer, S.M., Amburn, E.P., Austin, J.D., Cromarti, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J.B., Zuiderveld, K., "Adaptive Histogram Equalization and Its Variations", *Computer Vision, Graphics and Image Processing*, Vol. 39, No. 3, pp. 355-368, 1987b.

13. Robson, J., "Frequency Domain Visual Processing", *Physical and Biological Processing of Images*, OJ Braddick and AC Sleigh, ed., Springer-Verlag, 1983.

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

15. Witkin, A., "Scale-Space Filtering", *Proceedings of International Joint Conference on Artificial Intelligence*, 1983.

16. Yuille, A.L., and Poggio, T., *Scaling Theorems for Zero-Crossings*, MIT A.I. Memo 722, 1983.