

Interactive Object Definition in Medical Images Using Multiscale, Geometric Image Descriptions

Robin E. Fredericksen, James M. Coggins, Timothy J. Cullip, Stephen M. Pizer

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

ABSTRACT

A promising approach to medical image object definition involves automatic computation of a region-based image description along with a region containment directed acyclic graph (RCDAG) induced from the description via multiscale analysis of image structures [Pizer 1989]. The information resulting from this computation provides the basis for interactive object definition. During object definition the human user inserts semantics into the image description through additions to and alteration of the automatically computed RCDAG. This paper describes the object definition method and a tool for interactive object definition. Design criteria and resulting design decisions for this tool are presented, followed by a discussion of preliminary image segmentation and object definition results.

Introduction

Medical image objects consist of image regions characterized both by spatial and intensity structure and by semantic constraints understood by image analysts, including physicians and dosimetrists. We have developed a tool for defining image objects based on interactive user augmentation of an automatically-derived description of image structure. The image description consists of non-overlapping contiguous primitive image regions organized by region containment relationships into a directed acyclic graph (RCDAG). Each object in the original image is composed

of a set of primitive regions; object definition requires specification of the component primitive regions. Our object definition tool does this through manipulation and editing of the RCDAG, hence the name Image Hierarchy Editor (IHE). IHE operates on images organized by RCDAGs. Two approaches for creating such image structure representations are described in [Pizer 1990] and [Coggins 1990].

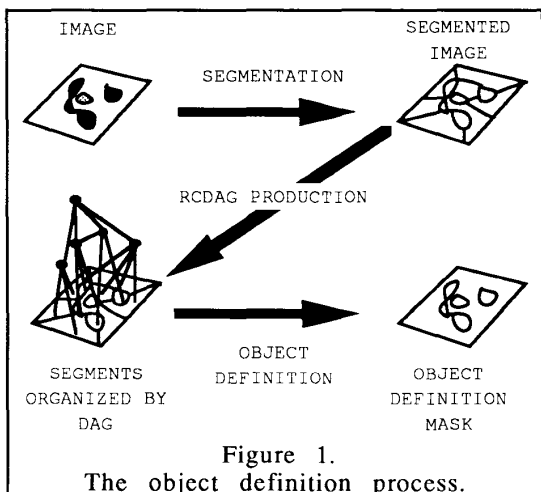
Motivation

Medical imaging seeks to measure quantitative or qualitative properties of objects. Such measurement requires that the image pixels portraying an object be distinguished from pixels not pertaining to the object. Known object definition methods use manual contouring (boundary specification), thresholding, or heuristics to label image regions according to the objects portrayed. Defining an object using manual contouring is tedious and time consuming. The requirement for accuracy demands pixel level operations such as painting and contouring; the requirement for speed demands higher level operations. Alternatively, automatic computation, for example thresholding or heuristic methods, can be fast but are much less accurate.

Our Approach to Object Definition

Our approach to object definition takes advantage of the strengths of both automatic bookkeeping of pixels, regions and labels, and human

perception and knowledge about the problems being solved. Figure 1 illustrates the main steps of our method.



We describe an image as a collection of primitive regions that can be organized by a region containment directed acyclic graph (RCDAG). These primitive regions may then be composed into coherent, visually sensible image objects. The primitive regions and hierarchies can be computed automatically [Pizer 1989, 1990] based on image intensity surface geometry. Some example methods for producing primitive regions and RCDAGs are based on maxima and minima [Lifshitz 1987], watershed (WS), watershed dual (WSD) and Intensity Axis of Symmetry (IAS) regions [Gauch 1988]. Another method produces regions based on spatial and scale coherence of scale and orientation features [Coggins, 1990]. IHE is designed to work independently from any particular algorithm.

The primitive regions and the RCDAG are computed without knowledge of expected object types in the image, so there are no objects explicitly defined in the RCDAG. Because people are much better than computer programs at recognizing objects in the varying context of medical images, IHE is

designed to allow the user to insert semantics into the RCDAG.

The IHE user inserts semantic information into an RCDAG by defining objects through aggregation of primitive regions. If the primitive regions and RCDAG are reasonable, then objects may be defined with a small number of actions. Region aggregates may be named, thereby adding new nodes representing these named objects to the RCDAG for later selection and reference. The object can be defined by selecting a pixel in each primitive region making up the object's image. IHE handles the bookkeeping of filling regions based on a user's pixel selection and maintains lists of selected regions. The definition of image objects can be accelerated further by exploiting the region relationships represented in the RCDAG. Each node in the RCDAG represents the region composed of all the primitive regions corresponding to its terminal node descendants. If the RCDAG connects primitive regions into visually coherent objects, then moving up the RCDAG from a terminal node (a primitive region) will aggregate primitive regions faster than selection of individual primitive regions making up an object.

To be effective, IHE must make the process of object definition simple, fast and accurate. Such performance requires 1) accurate primitive regions, 2) a good RCDAG, and 3) fast and easy use of IHE. The first two requirements are addressed by algorithms that compute the primitive regions and RCDAG [Cullip 1990]. The third requirement is addressed by careful design of the tool.

Designing IHE

Object definition operations implemented by IHE are: a) color labeling of primitive regions, b) primitive region aggregation, c) primitive region union and difference,

d) aggregate region union, difference and intersection, and e) RCDAG traversal. Each of these are discussed below.

Color Labeling of Regions

Color labeling is used to display the results of region selection and aggregation operations. Labeling is achieved by linearly mapping each pixel in the labeled region from its unlabeled intensity level to a corresponding intensity level in a color scale. This scheme allows the intensity structure of a selected region to remain visible during object definition, albeit at a lower contrast. Color labeling is supported by a color look-up table defined as shown in figure 3. The 256 colors supported by the workstation are divided into a 128-level grey scale and eight 16-level color scales. Seven of the colors are available to the user (see the IHE user interface in figure 2).

GREY SCALE	COLOR SCALES							
128	16	16	16	16	16	16	16	16

Allocation of color map entries.

Figure 3.
The IHE color map implementation.

Object Definition Operations

Users define objects by selecting and combining regions. Results of these operations are indicated by color labeling. Operations on regions labeled with different colors are also provided.

To label a primitive region, the user clicks a mouse button while pointing to a pixel in the image. The primitive region containing the selected pixel is painted with the currently selected color. Clicking on other pixels effects a set union of the primitive regions containing the selected pixels. The aggregate region thus composed need not be contiguous. Figure 4 shows a selected primitive region in the corpus callosum. A primitive region is removed from the labelling by clicking on a pixel in that primitive region.

Aggregated regions (labeled with distinct colors) can be manipulated to produce new regions using set union, difference, and intersection operations. The user selects two operand regions by color, specifies the color of the result region, and then specifies the operation to be performed. This suite of operations requires a maximum of four mouse clicks.

The potentially most powerful object composition operation in IHE allows the user to select regions based on the RCDAG. In a single step, the user can traverse the RCDAG to a parent node, thereby labeling all pixels in primitive regions that are descendants of that parent node. This traversal to the parent can be interpreted as selecting the next larger containing region in the RCDAG. Although a given node may have more than one parent, ambiguity is avoided by reducing the DAG to a logical tree (explained below).

Navigational Aid

IHE provides a simple but useful aid to navigation through the RCDAG (figure 5). While the right image window displays the currently selected object region (a portion of the brain stem in the figure), the left image shows in blue the primitive regions that will be added by performing a parent operation from the current node of the RCDAG. This allows the user to decide whether performing a parent operation is desirable. The navigational window is updated as the user traverses the RCDAG; if the RCDAG is reasonably good, regions of interest may be defined rapidly with a minimum of operations.

Learning and Performance Curves

IHE is designed for a user community of computer literate medical professionals. Minimization of learning time and of object definition time (as measured by the number of mouse operations) were primary design criteria.

In general, a point-and-click (event-driven) user interface minimizes memory load on the user by eliminating the need to memorize command line structures, and provides a simple interactive environment. The event-driven user interface developed for IHE allows use of the mouse both for selecting regions and for performing operations. IHE provides labeled buttons on the control panel, pop-up menus, and abbreviated mouse click semantics using all three mouse buttons and the shift and control keys.

Reducing Apparent Complexity

Our intended users can easily understand the concept of a binary tree. Visualizing and using a bushier tree is more difficult; visualizing a DAG is more difficult still because a node may have more than one parent. We therefore simplify the presentation of the RCDAG by traversing the RCDAG as a tree. The reduction of the RCDAG to a tree is accomplished by ignoring all but a node's first parent. This simplification could possibly remove important structure information; however the RCDAGs we have encountered are in fact trees, so we have yet to observe negative effects from this simplification.

The Need for Speed

Fast response is crucial for any interactive tool. Because IHE manipulates multiple images and abstract data structures simultaneously, the internal structure of IHE is carefully designed to achieve reasonable response rates.

The speed of object definition with IHE is determined largely by the quality of the regions and the RCDAG computed by the image description algorithms. Assuming that good primitive regions and a good hierarchy can be obtained, there remain several design issues facing IHE that affect its response time.

The first issue is to create a simple and fast command format. Commands to IHE

can be specified by clicking a mouse button (possibly shifted) while pointing to either a button on the main control panel or a portion of the right image window. Control panel buttons may bring up a menu or another control panel. Most button commands are duplicated with specific mouse actions. These abbreviated commands can help expert users to manipulate IHE rapidly, while the buttons and menus provide for the novice or less frequent user.

The speed of pixel operations is improved through use of two data structures. First, the RCDAG is implemented using an array-based node heap rather than explicit pointers. Array indexing is used to traverse the RCDAG rather than pointer evaluation. Second, bounding boxes on primitive regions are stored to limit pixel operations to image areas of interest. The use of bounding boxes increases the speed of pixel operations, such as color labeling of primitive regions, by an order of magnitude.

Definition Process ' Evaluation

Evaluation of our object definition process requires qualitative and quantitative evaluations of primitive regions and their organizing RCDAG. Some indication of the quality of the objects defined and the speed at which they may be defined is also necessary. The test image used in this evaluation is the MRI image of the head used in the figures in this paper.

Qualitative Evaluation

The range of primitive region sizes produced by the WSD and IAS methods differ widely. The WSD method produces primitive regions that are more uniform in size and shape. The IAS method results in regions that are large in wide, flat areas of the intensity surface, but small and more numerous in areas of the intensity surface that have finer structure. This is a good

indication of the sensitivity of the IAS to scale in the image.

The organizing RCDAGs are also qualitatively different. The IAS RCDAG is in general much flatter than the WS RCDAG; it moves to larger scales more quickly for a given next-parent action.

Quantitative Evaluation

The object definitions we speak of here generally include some portion of the object surround. Quantitative results of object definition in the test image using the WSD approach were reported in [Pizer, 1989] and are repeated here. Surprisingly, both the IAS and WSD methods yield an almost identical number of regions (1510 for the IAS and 1530 for WSD). Definition of the tongue region in the image required 5 atomic actions (mouse clicks) with the WSD method, but 13 atomic actions with the IAS method. Defining the cerebral cortex with the WSD method required 18 groups of 3-5 next-parent operations; using the IAS method required 10 groups of 2-3 next-parent operations and 10 individual primitive region union operations. This is a difference of about 30-40 atomic actions, or a 50% savings over the WSD method. Finally, definition of the brain stem using the WSD method required 11 atomic actions while requiring 17 atomic actions using the IAS method.

Volume Rendering

The general quality of the IAS segmentation method was evaluated by segmenting and volume rendering a cerebellum from a 3D MRI image of a human head. (The image used in the figures in this paper is one slice of that data set.) The object definition labels were treated as image masks, and were used to extract the selected volume elements from the original 3D data set. These volume elements were then used in a volume visualization process. The number of atomic actions required to define the cerebellum varied from slice to slice., ranging from 3 to 15, with an

average around 10. Each slice also required somewhere from 1-3 minutes, averaging slightly less than two minutes. The volume rendering was produced from 69 128x128 sized image slices. Some images from this process are displayed in figure 6. It is obvious that there are some regions produced by the segmentation process that are inappropriate (note the extra material on the bottom right of the cerebellum in figure 6b.). These deficiencies may be addressed with better primitive regions or with region editing capabilities (see Future Developments). Although the resulting volume visualization may not be pleasing to the eye, our radiation physicist evaluator believes that the resulting segmentation is good enough for most computational purposes, for example dose volume determination.

Conclusions

The IAS and WSD methods are computationally different, and therefore produce qualitatively different segmentations and hierarchies. These differences are reflected in the quantitative measurements of object definition discussed above. Segmentation using the IAS creates region sizes in closer correspondence with local surface scale, which results in good object boundary definition. We also note that in the cases where the WS method required fewer atomic actions for object definition than the IAS method, the IAS method produced better object boundaries due to its primitive region size variation corresponding to image structure scale. The shallowness of the IAS hierarchy is, we believe, a result of our implementation of the hierarchy induction method and is therefore correctable. Finally, the volume rendering produced using IHE is of a quality that appears to be satisfactory for computational purposes in radiation treatment planning.

Future Developments

Image roam-and-zoom is implemented in preparation for adding pixel level editing operations. These operations are required even when good RCDAGs are available because some semantically meaningful structures, e.g. radiation treatment volumes, are actually clinical or mathematical abstractions that are not present in the image data. Because IHE must support definition of such volumes, pixel level editing cannot be ignored. Implementation of pixel painting will require the ability to create new primitive regions because painting portions of a primitive region will force IHE to split that region.

The current version of IHE does not allow the user to select nodes directly from the RCDAG. A terminal node must be chosen first, then parent and child operations are used to move through the RCDAG. Object definition may be even faster if the RCDAG can be manipulated directly, but such a capability will require a 3-D visualization of the RCDAG connectivity.

IHE is implemented in AT&T C++ and uses the InterViews user interface library. Forthcoming enhancements of X11 (X11R4) such as a shared memory facility should make display updates faster, so we plan to continue use of X11. We are seeking possible alternatives to InterViews.

Acknowledgements

We thank Marc Levoy for use of his volume rendering software and Vicki Interrante for producing the volume renderings. We thank Bo Strain for photography and Graham Gash for

invaluable advice regarding X11 and Unix. We thank Drs. Julian Rosenman and Ed Chaney for their evaluation of the volume rendered images, and we thank Siemens, Inc. for providing the MRI image data. The research reported herein was carried out with the partial support of NIH grant number P01 CA47982.

REFERENCES

- [1] Coggins, J.M., "A Multiscale Description of Image Structure for Segmentation of Biomedical Images", in this volume, 1990.
- [2] Cullip, T.J., Fredericksen, R.E., Gauch, J.M., Pizer, S.M., "Algorithms for 2D and 3D Image Description Based on the IAS.", in this volume, 1990.
- [3] 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, 1988.
- [4] Lifshitz, L.M., *Image Segmentation Using Global Knowledge and A Priori Information*, Ph.D. Dissertation, UNC Chapel Hill, 1987.
- [5] Pizer, S.M., Gauch, J.M., Coggins, J.M., Cullip, T.J., Fredericksen, R.E., Interrante, V.L., "Multiscale, Geometric Image Descriptions for Interactive Object Definition", *Mustererkennung*, 1989 (Proc 11th Symposium of DAGM [The German Association for Pattern Recognition]), *Informatik-Fachberichte* 219: 229-239, Springer-Verlag, 1989
- [6] Pizer, S.M., Gauch, J.M., Cullip, T.J., Fredericksen, R.E., "Descriptions of Image Intensity Structure via Scale and Symmetry.", in this volume, 1990.

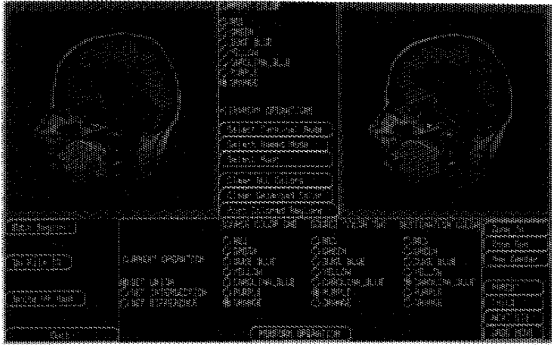


Figure 2.
IHE user interface with labeled objects.

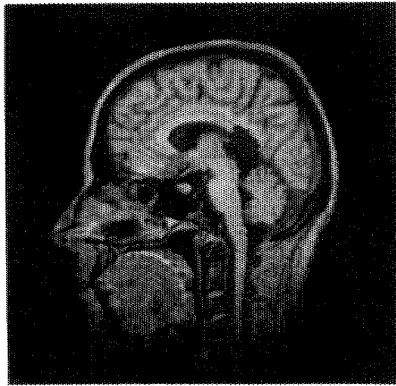


Figure 4.
A primitive region in the corpus callosum.

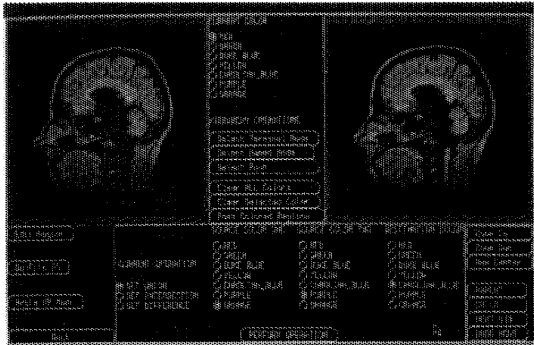


Figure 5.
The right image displays the labeled regions. The left image also labels the sibling regions of the current node in the RCDAG (blue).

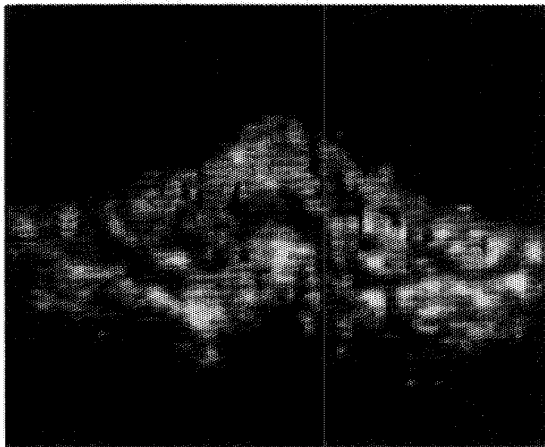


Figure 6
Volume visualization of the cerebellum. (a-c, top to bottom)