Semi-automatic medical image segmentation using knowledge of anatomic shape

James F. Brinkley

Department of Biological Structure SM-20
University of Washington, Seattle, Washington 98195

ABSTRACT

A program called SCANNER (version 0.6) is described for performing 2-D interactive medical image segmentation using knowledge of anatomic shape. The knowledge is implemented in a *radial contour model*, which is a flexible, generic model that can accurately deform to fit the data, but which also encodes the expected shape and range of variation for a 2-D contour shape class. The model, which can describe contours that are single-valued distortions of a circle, is learned from training sets of similarly-shaped contours. Variation in the learned model allows it to provide search regions for low level edge detectors, thereby reducing the incidence of false edges. Initial evaluation of this system was performed for structures seen in 111 2-D CT images from 12 patients undergoing radiation treatment planning for cancer. The results suggest that the model is able to capture the cross-sectional expected shape and range of variation for several clinically-important structures (the liver, kidney, eye, and some tumors), that the knowledge-based approach should reduce the segmentation time over current manual methods by a factor between two and ten, and that the usefulness of the model decreases as variability of the structure increases.

1. INTRODUCTION

One of the most ubiquitous problems in medical image analysis is segmentation of important biological structures from the background. This problem arises frequently because segmentation is a necessary step for any sort of processing other than enhancement of 2-D images for human visualization. With the increasing availability of 3-D and 4-D (time varying) image data, the segmentation problem becomes even more acute because of the added amount of data. In these cases manual segmentation by human experts is unacceptably time consuming and error-prone, and is a major reason why 3-D visualization and manipulation is not more widely used in clinical medicine.

One example of the need for improved segmentation is three-dimensional radiation treatment planning[1]. Patients with tumors are often treated by high dose radiation, and the goal of radiation treatment planning programs is to arrange the x-ray beams so that maximum dosage reaches the tumor while minimum dosage reaches the surrounding tissue. In particular, there are certain critical structures which must not receive high dosage. These structures include the liver, kidney, eye, brainstem, and spinal cord, among others. In order to be sure these structures receive minimal dosage, radiation dosimetrists manually segment them on a set of parallel CT images through the patient, after which the structures are reconstructed in 3-D and passed to the treatment planning programs. The soft tissue structures are segmented manually because commonly-available low level segmentation programs fail so often that they become useless. The result is that, in our institution, the segmentation process can take up to half the planning time. This situation is very common for other image segmentation problems, so that a large part of medical image segmentation is still done manually, particularly for soft tissue objects which are not well delineated from the background.

The approach to segmentation followed in this research is based on two major premises: 1)knowledge of anatomy is required in order to disambiguate structures in noisy images, and 2) the segmentation

problem is not likely to be completely solved in the near future, so any segmentation system must be interactive. Although several knowledge-based segmentation systems have been developed in the past several years[2,3,6,11,15,17,18,19,20], very few if any are in widespread clinical use. This lack of practical application is due in large part to the fact that segmentation remains a difficult problem, and that those knowledge-based systems that have been developed have not yet been integrated with interactive tools. Such tools, which are increasingly being developed for commercial image analysis products[8,9,14], allow the user to compensate for deficiencies in the computer methods while still gaining some improvement over purely manual methods.

This paper describes the initial implementation and evaluation of an interactive program called SCANNER (version 0.6), which attempts to combine a knowledge-based approach with good interactive tools. The current implementation has been tested primarily on images from patients undergoing radiation treatment planning for cancer, but the program should be applicable to many other segmentation problems as well.

2. REPRESENTATION OF ANATOMIC SHAPE

The particular form of knowledge employed in SCANNER is knowledge of anatomic shape. The knowledge is represented in a *generic*[5] model that is *flexible* enough to accurately fit the data, but which also captures the expected shape of an anatomic shape class as well as its range of variation. The ability to explicitly encode variation in a flexible, generic model allows a model-based system to predict search regions on the image, while at the same time providing an accurate fit of the model to the data. Although many of the currently-popular deformable models are also flexible and therefore accurate[7,10,12,15,17], they do not encode variation, so cannot be easily used to direct the search for edges. On the other hand, the generic models de-
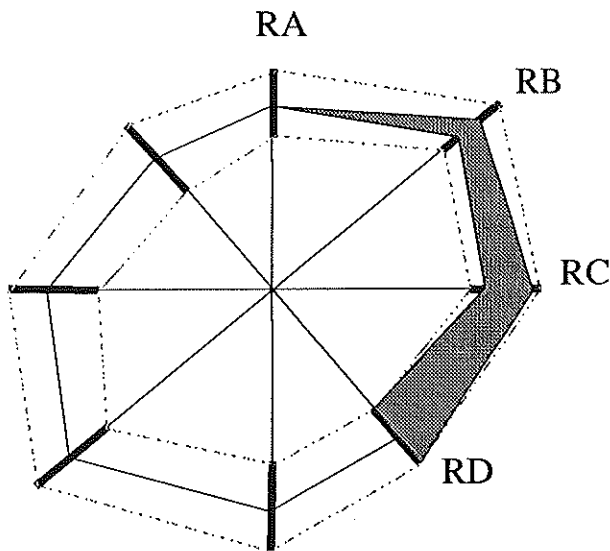

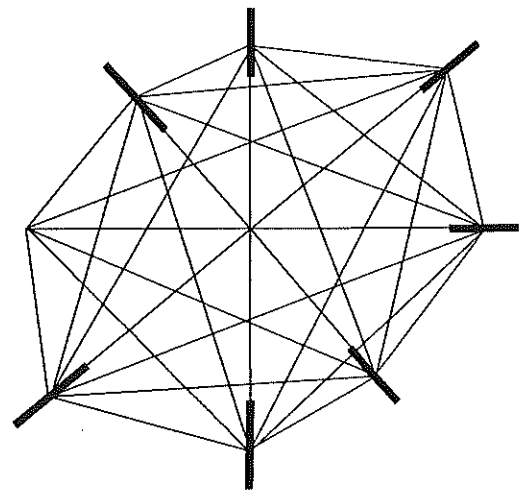Figure 1. Local radial contour model

Figure 2. Maximal radial contour model

veloped for modelling man-made objects are not flexible enough to accurately fit the data[5].

The model employed in SCANNER is called a radial contour model (RCM). This model is an example of a geometric constraint network, a type of constraint network[13] which has been proposed as a general representation for biologic objects[4]. The hypothesis behind this representation is that networks of local interacting

geometric constraints between structure subparts, when interacting together, are able to generate an overall representation of the essential shape and range of variation of the structure. The RCM is a 2-D implementation of a previously reported 3-D representation that was used in a knowledge-based 3-D ultrasonic organ modelling system[3]. However, the previous representation was not implemented in an interactive system, nor was it ever tested on real medical images.

Figures 1 and 2 show the radial contour model, which can be used to represent 2-D contours that are single-valued distortions of a circle. A general constraint network consists of a set of variables, a set of possible values for each variable, and a set of constraints that determine which of the possible values for the variables are compatible[13]. In the case of the radial contour model the variables are points on the contour boundary, each of which is restricted to lie along a set of fixed radials emanating from a local contour coordinate system. The value for a single variable is the distance from the origin to the contour boundary, and the possible values are given by one-dimensional *uncertainty intervals* along each radial (shown as darkened lines). Lines drawn between the inner endpoints of each interval, and between the outer endpoints, define a 2-D search region within which the computer always expects the contour on the image to lie (figure 1). Lines drawn between the midpoints of all the intervals defines a contour which represents the best guess at any one time as to the actual location of the contour in the image.

Constraints between radials are defined from a training set of similarly shaped contours. In a local radial contour model (figure 1), each radial is only constrained by its nearest physical neighbor, whereas in a maximal radial contour model (figure 2) every radial is constrained by every other radial. For each member of the training set, and for each pair of radials RA and RB connected by a constraint, the ratio is measured between the observed distance along RA to the contour and the observed distance along RB to the contour. The range of such observed ratios defines the constraint between RA and RB. That is, if the value of RA is given (say by an edge in the image) then the constraint states that RB must be within an interval given by the constraint.

These constraints interact with image data by a constraint propagation process, which is shown as the shaded area in figure 1. An edge or user input for radial RA causes the uncertainty interval for RA to be reduced to a single point (the possible values for RA in the constraint network now consist of only one point, the edge observed in the image). This edge information, when combined with the ratio constraint between RA and RB, propagates to RB, causing the uncertainty interval at RB to be narrowed from its original value. Although the interval is wider than at RA, the combination of knowledge of shape variation encoded by the constraint, together with measured data at RA, allows additional information to be inferred at RB. Since RB is now changed, its value can be propagated further to RC, and so on until an interval does not change, in which case the propagation wave stops (in this case, at radial RD). The result of this constraint propagation procedure is that edge information obtained in one part of the contour is able to reduce the search region for edges obtained in another part of the contour. This procedure is an application of relaxation labelling[21,13] to the problem of model based segmentation. Details of the procedure may be found in the previous report[3], along with a proof that, under a reasonable assumption, the procedure executes in $O(N)$ time for the local model. Under the same assumption, the maximal model should execute in $O(N^2)$ time, but the observed execution time is much less than that.

3. THE SCANNER PROGRAM

The radial contour model has been implemented in a program called SCANNER (version 0.6) in order to demonstrate and evaluate the utility of the model for interactive 2-D image segmentation. SCANNER is implemented in Objective-C on the NeXT computer, and makes extensive use of the interface development tools available on the NeXT. SCANNER consists of several interacting objects: an ImageManager for managing 2D images, an Edgefinder for finding edges along a 1-D line in the image, a ModelManager for creating and managing Radial Contour Models, a ContourManager for creating and managing Radial Contours (the output of the program), and a Segmenter for performing knowledge-based segmentation. The modules are controlled by the user with menus and inspector panels (figure 3).
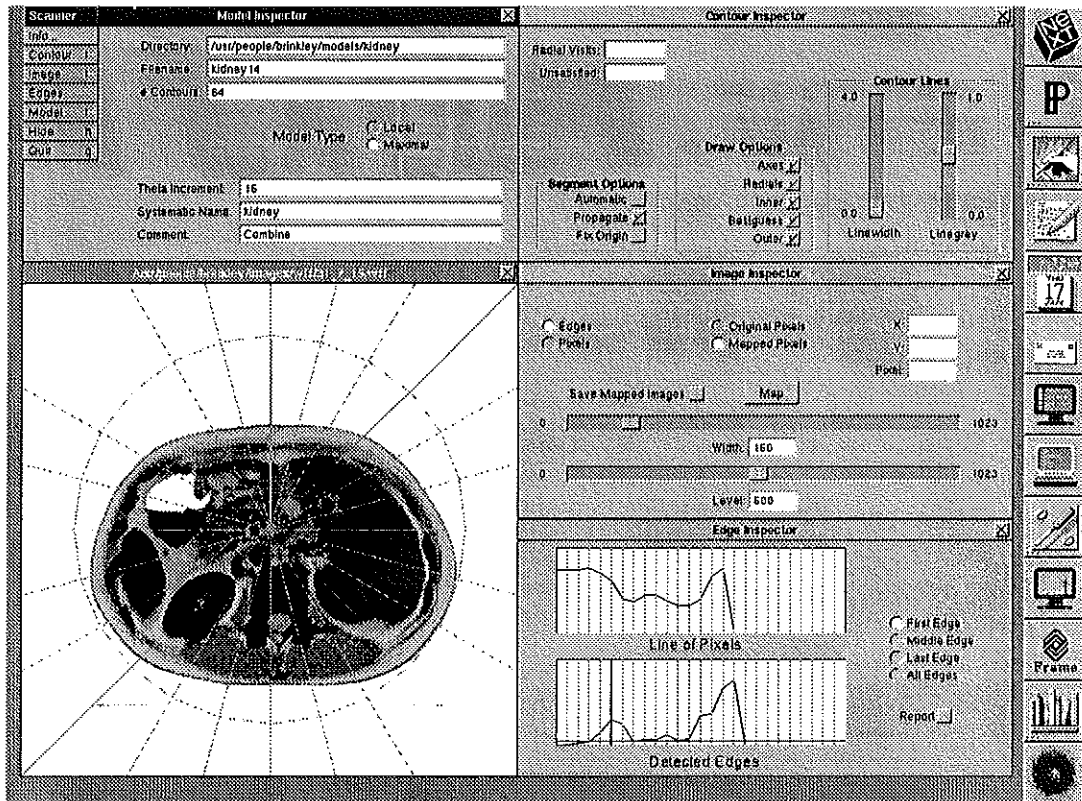

Figure 3. SCANNER user interface

Given a previously created radial contour model retrieved by the ModelManager, the program is able to perform interactive segmentation (figure 4 A,B,C). The figure shows model-based segmentation using the local radial contour model rather than the maximal model because the wider search regions make it easier to see how sequential interactions between the model and image data cause the search regions to become smaller. The New contour menu item causes an initial radial contour model to appear in the center of the image (figure 3), after which the user positions the mouse to move the center of the radial contour to the perceived center of the object on the image, in this case the kidney. This operation bypasses the problem of locating a frame of reference for the local coordinate system, which is an example of the matching problem that is not likely to be solved in the near future. Although this problem is difficult for the computer, it is not difficult or tedious for the user. The approach to this problem taken in SCANNER is typical of the approach to the segmentation problem in this research: let the user continue to perform the tasks that are difficult for
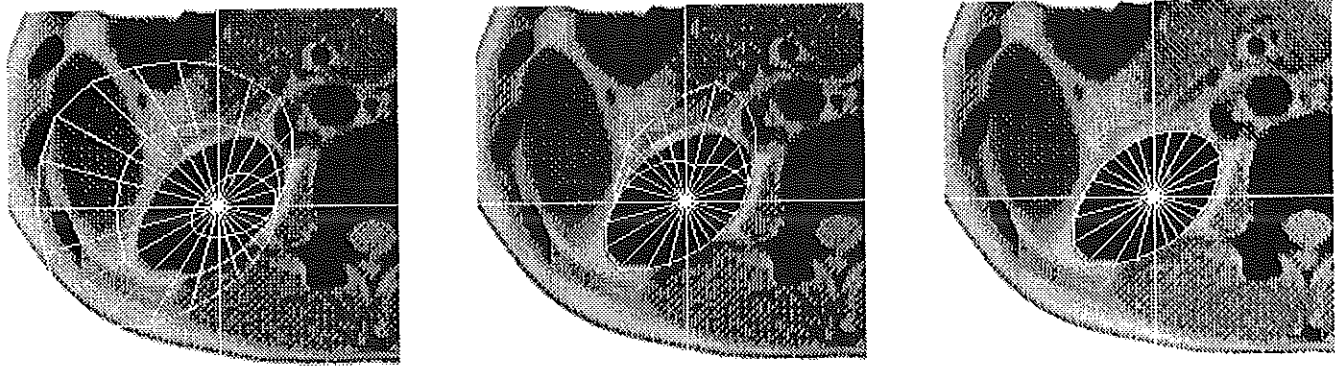
| A. Initial | B. Intermediate | C. Final |

Figure 4. Model-based segmentation of the kidney

the computer, but let the computer perform those tasks it can do, and which will result in an overall speedup in the segmentation time. As further research improves the model-based approach, more of the tasks can be transferred to the computer, thereby gradually iterating towards a completely automated system.

Once the user has moved the center of the contour to the center of the object on the image, he indicates the position of the contour along a single radial, and constraint propagation causes the search region to be narrowed from its initial value. In Figure 4A the user has indicated the position of the contour along the positive X-axis and constraint propagation has created the initial search region. The computer then sequentially chooses a radial (currently the radial with the smallest uncertainty), sends the line of pixels corresponding to the uncertainty interval along the chosen radial to the edge finder, and uses the returned edge to initiate the constraint propagation procedure. Figure 4B shows the model after several radials along the bottom part of the contour have been found. The search terminates when all the radials have been examined (figure 4C). If the computer incorrectly finds an edge the user can manually correct the offending radial.

4. INITIAL EVALUATION

An initial evaluation was performed to determine how useful the program might be to radiation oncologists. For this purpose a series of 111 2-D CT images were obtained from 12 patients currently undergoing routine radiation treatment planning in the Department of Radiation Oncology. SCANNER was used to build models for the eye, kidney, and liver from these patients, as well as a model for tumors seen on three patients, and the anterior horn of the lateral ventricle as seen on images of a cadaver brain. Since the maximal model produced smaller search regions than the local model, and since both models executed at about the same rate (contrary to expectations), results are presented for the maximal model only.

For each structure several models were built from training instances or from a combination of other models. Figure 5 shows best guess contour and confidence limits for several of these models. Note that the shapes appear recognizable as the corresponding structures, even though the only difference in the models is the content of the training sets. In fact, when anatomists were shown these shapes, together with a list of possible choices, they were able to match the object name with the shape.

Figure 5 shows that the different models form a hierarchy according to the expected amount of variability in the model. At the bottom of the tree ("Single Slice" level) are models formed from 6 training instances from the same image. Variation within this model arises from uncertainty as to the location of the origin of the contour coordinate system, as well as noise in the image. These were the only models created directly
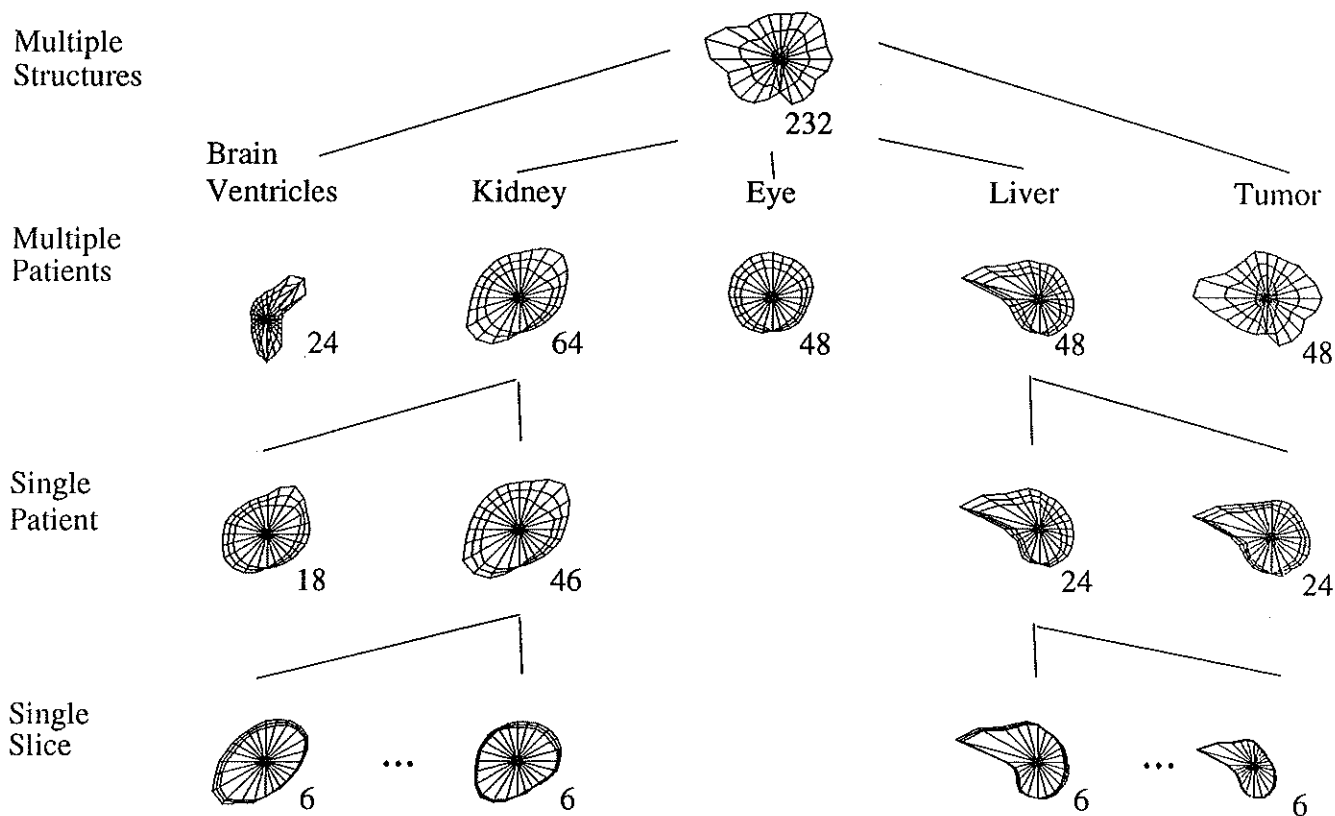
Figure 5. Essential shape and range of variation for various models

from training instances. All other models were combinations of these models. However, the combination method produces exactly the same result that would have occurred if each model were produced directly from training instances.

At the "Single Patient" level are models obtained from a single patient but from different slices at different levels through the organ in that patient. Variation in this case is due to two factors: variation at the single slice level, and variation in the shape of the contour at different levels in the organ.

At the "Multiple Patients" level are models obtained from the same structure but from more than one patient. These are the beginnings of generic models that capture the range of variation for a structure, in addition to variation at the lower levels.

Finally, at the top "Multiple Structures" level is a single model representing a combination of all models. This is essentially a control that can be used to assess variability and usefulness when no specific knowledge of organ shape is available.

Table 1 compares the different model types from figure 5 according to variability and expected usefulness. The first column shows the total number of constraints for all models at each level, where the number of constraints in a single model was 552 (24 radials x 23 constraints per radial in the maximal model). Thus, there were 2208/552 or 4 models at the Multiple Patients level. The second column shows pooled mean and standard deviation of variability for these different model types. *Variability* was measured for each model

by calculating the mean and standard deviation of the differences between minimum and maximum values for each constraint in the model. Thus, the wider the constraint limits the higher the measure of variability. All differences were significant at $p < .05$.

	#Constraints	Variability	#Trials	Usefulness
Multiple Structures	552	1.72(0.91)	20	13.5(8.0)
Multiple Patients	2208	0.57(0.32)	52	6.7(5.1)
Single Patient	6072	0.38(0.28)	78	4.6(4.6)
Single Slice	21528	0.15(0.10)	117	4.4(5.8)

Table 1. Comparison by source of variation

Table 1 also gives a measure of usefulness for each model. In this study *usefulness* was defined as the number of radials that needed to be corrected by manual radial updating after the computer had looked at all the radials. Since all models had 24 radials, and since one radial was always updated manually at the beginning to start the process, the least useful model would have a usefulness of 23 (all 23 radials had to be corrected), and the most useful model would have a usefulness of 0 (no radials had to be corrected). This is a measure of potential clinical usefulness that is independent of expected improvements in hardware. Actual execution times on the current 68030-based NeXT, using non-optimized code, were on the order of one to two minutes, which is too slow for clinical use. However, this time will dramatically improve with faster hardware.

Given this definition of usefulness, Table 1 shows that the most useful model type was from single slices or single patients (there was no significant difference between the single patient (4.6) and single slice (4.4) measure of usefulness. The table also shows that, as expected, usefulness decreases as variability increases. However, usefulness is still 13.5 for the control model, which shows that on the average the edge finder alone would still speed things up by a factor of two over a completely manual method. Usefulness increases on the average by another factor of two (13.5 versus 6.7) as specific knowledge is added.

	#Constraints	Variability	#Trials	Usefulness
Multiple Structures	552	1.72(0.91)	20	13.5(8.0)
Tumor	552	0.99(0.52)	12	11.1(6.3)
Liver	552	0.53(0.29)	16	8.3(5.7)
Kidney	552	0.44(0.17)	14	4.2(3.5)
Eye	552	0.33(0.10)	10	2.5(4.1)

Table 2. Comparison by structure

Models were also compared as to usefulness and variability when classified according to structure type, as shown in table 2. In this case only the multiple patient and multiple structures models were compared. Again, as expected, the most regular structures had the least variability and were the most useful. Thus, the eye, which is almost a circle in cross-section, had the least variability and proved most useful for model-based segmentation (expected speedup approximately 10 times over a completely manual method), whereas

the tumor and combined model had the most variability and were least useful (expected speedup approximately 2 times).

5. DISCUSSION

This paper has described a flexible, generic representation for 2-D contours which are distortions of a circle. The representation has been implemented in an interactive image segmentation program called SCANNER (version 0.6). Although the radial contour model certainly cannot represent all anatomic structures, the initial evaluation suggests that the model can represent several clinically useful shape classes, including many of the critical structures needed for radiation treatment planning. For these structures, the initial evaluation results suggest the following 1) the model captures the essential shape as well as the range of variation purely from a network of interacting constraints learned from training examples; 2) the model can represent abnormal objects such as tumors, but if the variability in the object is large (as it will be when more tumor types are included) the corresponding usefulness for model-based segmentation is reduced; 3) noise in the images is simply reflected in greater variability in the model. Thus, models built from ultrasound training sets are expected to be more variable and less useful than models built from CT images, but both models should be more useful than manual methods; 4) by employing a graphical user interface and good interactive tools it is possible to build useful semi-automatic knowledge-based contouring systems.

Additional evaluation with larger numbers of patients are necessary in order to show that variability does not become so large that the models provide no useful guidance. However, if additional patients show that the models remain useful, then this approach should greatly speed up the segmentation process for those types of structures that can be represented. A clinically-useful system would incorporate the knowledge-based approach, but would also incorporate other interactive methods for segmenting objects that could not be represented by the radial contour model. As the radial contour model is generalized to handle additional structures, these new model types can be added to the clinical system. In this way a practical, knowledge-based system can be developed that gradually becomes more automated as the difficult representation problems are solved.

6. ACKNOWLEDGEMENTS

7. REFERENCES

1. N.M. Bleehan, E.Gladstein, and J.L. Haybittle, editors, *Radiation Therapy Planning*, Marcel Dekker, New York, 1983.

2. A. Botto, A. Casali, S. Dellepiane, G. Ghilino, and G. Vernazza, "Expert segmentation for the extraction of biomedical structures in microscopical images", In SPIE Vol. 1092 *Medical Imaging III: Image Processing*, pages 194-198. SPIE, 1989.

3. J.F. Brinkley," Knowledge-driven ultrasonic three-dimensional organ modelling. *IEEE Trans. Pattern*

Anal. Mach. Intell., PAMI-7(4):431- 441, 1985. Also available as Stanford Technical Report KSL-85-33.

4. J.F. Brinkley, "Representing biologic objects as geometric constraint networks", In *Proceedings, AAAI Spring Symposium Series: Artificial Intelligence in Medicine*, pages 7-8, Stanford University, March 22-24, 1988.

5. R.A. Brooks," Model-based three-dimensional interpretation of two- dimensional images", *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-5(2):140-150, 1983.

6. S.-Y. Chen and W.-C. Lin, "An expert vision system for medical image segmentation", In SPIE Vol. 1092 *Medical Imaging III: Image Processing*, pages 162-172. SPIE, 1989.

7. R. Dann, J. Hoford, S. Kovacic, M. Reivich, and R. Bajcsy, "Evaluation of elastic matching system for anatomic (ct, mr) and function (pet) cerebral images", Journal of Computer Assisted Tomography,131(4):603- 611, July/August, 1989.

8. J. Ezrielev and S.C. Dzik, "Image editor for a 3-d ct reconstruction system", In SPIE *Medical Imaging IV: Image Processing*, pages 67-76, Society of Photo-Optical Engineers, 1990.

9. R.T. Fan, S.J. Zinreich, C.B. Quinn, and L.L. Fellingham, "Minimally interactive soft tissue segmentation for 3-d display from tomographic images", In SPIE Vol. 1092 *Medical Imaging III: Image Capture and Display*, pages 279-286. SPIE, 1989.

10. I. Kapouleas and C.A.Kulikowski, "A model-based system for the interpretation of mr human brain scans", In R.H. Schneider and S.J. Dwyer, editors, SPIE Vol. 914, *Medical Imaging II*, pages 429-437, Newport Beach, Jan 31-Feb 5 1988.

11. N. Karssemeijer,L.J.Th.O. van Erning, and Eg.G.J. Eijkman," Recognition of organs in ct image sequences: a model guided approach", *Computers and Biomedical Research*, 21:434-448, 1988.

12. M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models", In SPIE 1987, pages 259-268. Schlumberger Palo Alto Research, 1987.

13. A.K. Mackworth, "Consistency in networks of relations", *Artificial Intelligence*,8:99-118, 1977.

14. P.H. Mills, H. Fuchs, S.M. Pizer, and J.G. Rosenman, "Imex: A tool for image display and contour management in a windowing environment", In Proc. SPIE, Vol 1091, *Medical Imaging III: Image Capture and Display*, pages 132-142, 1989.

15. R.B. Schudy and D.H. Ballard, "A computer model for extracting heart surfaces from four-dimensional cardiac ultrasound data", In Proc. 6th Conference on Computer Applications in Radiology and Computer-aided Analysis of Radiologic Images, pages 366-376, Newport Beach, Calif., June 1979.

16. R.J. Seitz, C. Bohm, T. Greitz, P.E. Roland, L. Eriksson, G. Blomqvist, G. Rosenqvist, and B. Nordell. Accuracy and precision of the computerized brain atlas programme for localization and quantification in

positron emission tomography. J. Cerebral Blood Flow and Metabolism, 10:443-457,1990.

17. U. Shani, "Understanding three-dimensional images: the recognition of abdominal anatomy from computer axial tomograms (cat)", Technical Report 82, University of Rochester, Dept. Computer Science, August 1981.

18. S. Shemlon and S.M. Dunn," Rule-based image interpretation with models of expected structure", In SPIE Medical Imaging IV: Image Processing, pages 33-44, Society of Photo-Optical Engineers, 1990.

19. C. Smets, P. Suetens, A. Oosterlinck, and F. Van de Werf, "A knowledge-based system for the delineation of the coronary arteries. In SPIE Vol.1092 *Medical Imaging III: Image Processing*, pages214-219, 1989.

20. B.I. Soroka, "Generalized cones from serial sections", *Computer Graphics and Image Processing*, 15:154-166, 1981.

21. D. Waltz, "Understanding line drawings of scenes with shadows", In P.H. Winston, editor, *The Psychology of Computer Vision*, chapter 2, pages 19-91. McGraw-Hill, New York, 1975.

22. P.H. Winston, "Artificial Intelligence", Addison-Wesley, Reading, Massachusetts, 2nd edition, 1984.