Interactive image segmentation for radiation treatment planning

by P. J. Elliott
J. M. Knapman
W. Schlegel

COVIRA (COmputer VIsion in RAdiology) is a project in the European Community's Advanced Informatics in Medicine program. The goal is to improve the diagnosis and planning of treatment (radiotherapy) for patients with brain tumors and other diseases. The aim of radiotherapy is to provide a high dose of radiation to a tumor while sparing as much as possible of the surrounding healthy tissue. A necessary first step is defining the target volume and organs at risk by manually outlining the required contours on magnetic resonance or computed tomography scans. For a full three-dimensional plan this is time-consuming, as 40 or more scans are used. Computer image segmentation speeds up the process, and a method that combines information from edge and region detectors is described. Since this method is not able to completely meet the clinical requirements, an interactive image segmentation algorithm has been developed that enables the operator to employ clinical judgment. Probabilities are assigned to edges and regions and presented to the user as a hierarchy of segmentations. The approach is being subjected to extensive clinical evaluation, using pilot applications running on IBM RISC System/6000™ workstations.

COVIRA (COmputer VIsion in RAdiology) is a project sponsored by the Commission of the European Communities under the Advanced Informatics in Medicine (AIM) research program. COVIRA uses computers to improve the diagnosis and planning of treatment for cancer patients, especially those with brain tumors.

The first phase of the AIM program ran from June 1989 to December 1990. The IBM United Kingdom Scientific Centre took part in this program, in a consortium with Philips Medical Systems, a clinical partner, and two academic partners. IBM contributed in the areas of image segmentation using edges and regions, texture analysis of magnetic resonance (MR) images, and the provision of a statistical knowledge source related to the MR imaging process. We developed working algorithms running on IBM RISC technology (RT*) 6150 workstations. In an evaluation of the clinical usefulness of the algorithms produced by the various partners in the project, ours were preferred by the clinicians.

For the main phase of AIM (January 1992 to December 1994), the consortium has been expanded to include Siemens, the other major European medical equipment manufacturer, in addition to other leading academic partners and a total of six clinics. This reflects the European Commission's desire for AIM to produce pilot systems that are subject to extensive clinical evaluation. Our pro-

posal was one of about 200 submitted in September 1991. Only 36 were accepted.

In the AIM main phase, the IBM UK Scientific Centre is concentrating on applications in radiation therapy planning, in conjunction with the Royal Marsden Hospital and Institute of Cancer Research (Sutton, UK) and the German Cancer Research Center (Heidelberg, Germany). These applications will run on IBM RISC System/6000* workstations. Other partners in the project will work on neuroradiological diagnosis and neurosurgery planning. IBM's specific contributions include interactive three-dimensional image segmentation, three-dimensional visualization, and development of the clinical user interface. We are also playing a significant role in the management of the COVIRA project.

The aim of radiation therapy is to provide a high dose of radiation to a tumor in order to eradicate it, while sparing as much of the surrounding healthy tissue as possible. Using modern imaging techniques such as CT (X-ray computed tomography) and MRI (magnetic resonance imaging), new three-dimensional methods of planning and applying the treatment have been developed in recent years. However, at present these methods are restricted to specialist centers because the planning stage is very time-consuming. Within COVIRA, a system will be developed that significantly speeds up the planning process through the use of computer visualization techniques to identify relevant structures in the images, together with knowledge-based systems to assist the user in generating the best treatment plan. The reduction of planning time and cost will allow many more centers to use the new techniques.

It is estimated that two-thirds of all tumor patients are referred to radiation therapy. About 40 percent of these can be treated effectively with current methods. Another 40 percent are not suitable for treatment because the disease has spread too far. The remaining 20 percent could be treated if the new three-dimensional planning methods were generally available. Therefore, up to 250 000 patients a year in Europe and the United States could benefit from the work being done in COVIRA.

In this paper, a brief description of the radiation treatment planning process is given. One of the most time-consuming stages, defining the therapy-relevant volumes, is done by manually outlining contours in the images; the current state of the art is described. A method for automatic image segmentation, developed by IBM during the exploratory phase of COVIRA, is discussed, followed by a detailed description of an extension of this work that allows the clinical user to interactively control the segmentation result.

Radiation treatment planning

The goal of radiotherapy, as stated previously, is to eradicate a tumor without causing severe damage to healthy tissue. Computer-assisted radiation treatment planning (RTP), and especially the newly developed three-dimensional treatment planning techniques,[1] provide a tool to find the optimum configuration of radiation beams for the treatment of an individual patient. In clinical practice, RTP has traditionally been based on two-dimensional sectional representations of the treatment of the patient. However, the shape of the tumor and adjacent uninvolved anatomic structures may vary substantially from one section to another. The sectional two-dimensional planning approach may lead to uncertainties in localizing the tumor volume and target volume as well as internal organs, a problem that can still be considered as one of the largest sources of error in the whole radiation treatment procedure.

Today, the situation is changing due to several technical developments, among which the advent of whole-body X-ray computed tomography (CT) and of magnetic resonance imaging (MRI) are the most important. With the availability of CT and MR imaging devices, the ability to visualize anatomical structures in detail has been tremendously enhanced. In order to take full advantage of these new imaging developments within the field of radiation therapy, new tools have to be developed and evaluated that permit easy and precise extraction of the therapy-relevant information from three-dimensional images; three-dimensional display of anatomical structures; and computer simulation for accurate placement of radiation beams. This paper focuses on the application of computers in the step that is currently considered the bottleneck of three-dimensional treatment planning: segmentation of therapy-relevant anatomical structures.

The radiotherapy planning procedure. The radiotherapy process can be separated into a number of steps, shown in Figure 1.

As can be seen from Figure 1, the process starts with an imaging procedure (CT and/or MRI). The main steps in the radiotherapy process are treatment planning (including segmentation of the therapy-relevant volumes and defining an arrangement of radiation beams to achieve the required dose), patient positioning, the (repeated) irradiation of the patient, and verification of the dose given. Computer assistance is essential in most of these steps.

The main steps of radiotherapy are treatment planning, patient positioning, irradiation, and verification.

The tumor volume includes all the tissue that can be recognized and identified as tumor tissue within the tomographic images. The identification of tumor tissue may be based on different criteria: e.g., sometimes enhanced, sometimes decreased soft tissue contrast in the MR or CT images, with sometimes well-defined, sometimes blurred or even invisible margins (after subtotal or total tumor resection). Further criteria for the definition of the tumor volume are the displacement of neighboring anatomical structures or the occurrence of edemous tissue.

After the definition of the tumor volume, the target volume has to be defined. This is the volume that has to be completely covered by the therapeutic dose during the irradiation of the patient. The target volume not only includes the tumor volume, but also possible tumor extensions, possible sites of metastasation such as lymphatic regions in the vicinity of the tumor volume, preferred paths of further tumor growth, and a so-called "safety region" around the tumor. The latter takes into account the uncertainties during the irradiation of the patient caused by patient movement and inaccurate patient alignment under the linear accelerator. The problematic nature of tumor and target volume definition is due to having to depend on histology and surrounding anatomy to determine tumor margins and safety margins. The goal of radiotherapy is not only to apply a high radiation dose to the tumor tissue, but also to avoid the occurrence of side effects in the surrounding healthy tissue, especially in those anatomical structures that will not tolerate radiation burdens above so-called "critical doses."

The third important type of therapy-relevant volume that has to be defined during the treatment planning process is the nontarget tissue, the "organs at risk." Depending on the location of the target volume, organs at risk in the brain may be the brain stem, white matter, the optic nerve, the eyes, or the lenses. There are several good reasons for the complete and accurate segmentation of organs at risk:

1. The geometrical shape and the location of the three-dimensional surfaces of the organs at risk can subsequently be used to optimize the position, shapes, and incidence angles of the radiation fields.
2. It is possible to precalculate the dose distribution within the target volumes, and, by further changes of the irradiation technique, minimize the physical dose in those critical areas.
3. With the help of radio-biological models, it is possible to estimate the risk of radiation injury (the treatment-related morbidity probability, or TRMP) of the critical organ and the probability for tumor control (tumor control probability, or TCP). To base the optimization of radiation treatment plans on TCPs and TRMPs seems much more promising than a pure geometrical and physical consideration of the treatment planning problem. In addition to more elaborate and realistic radio-biological models, what will be needed in the future to make this biological treatment planning a routine tool are procedures for individual three-dimensional image segmentation for the definition of critical organs or even functional subunits of critical organs.

Tumor tissue that is not included in the target volume will be underdosed with high probability; organs at risk that are not included during this process will probably not be considered in the

evaluation of treatment plans and eventually be overexposed. Errors, inaccuracies, and oversights that are made during the first step of image segmentation dramatically influence the effectiveness of treatment planning and the outcome of the whole radiotherapeutic process. This shows the importance of optimal treatment planning in radiotherapy.[2]

State-of-the-art methods in clinical practice: Manual segmentation of target volumes and organs at risk

Currently, the method used in clinical practice for extracting three-dimensional information is manual segmentation of CT and MR data sets. The number of slices comprising a three-dimensional image may vary from 20 to 60. As many as ten or more structures may need to be outlined on each of these images. This is a major effort to be carried out for each patient. Until now, there were few automatic tools available in clinical practice to make this process easier and faster. However, there are some groups developing fast and user-friendly three-dimensional image segmentation and modeling tools.[3-5] To characterize the state of the art in manual segmentation, one of these systems is now described in some detail.

Tool for manual segmentation. TOMAS (TOol for MAnual Segmentation) is an interactive workstation program that has been developed at the German Cancer Research Center since 1987.[5] It is an advanced program for the three-dimensional segmentation of tumor volumes, target volumes, and organs at risk and is routinely applied in radiation treatment planning. Originally developed on VAX** workstations under the VMS (virtual memory system) operating system, it has recently been ported by the UK Scientific Centre to run on the IBM RISC System/6000 workstation under AIX*. The software is written in the C programming language, and the object-oriented user interface is implemented with X Windows** and OSF/Motif**.

The three-dimensional reconstruction and presentation of the contour lines is realized by means of PHIGS (Programmers Hierarchical Interactive Graphics System). The PHIGS we used conforms to the international PHIGS standard ISO 9592: 1988(E), which is device-independent and promises good portability.

CT and MR images are transferred directly to the workstation through an internal data network (Ethernet with Transmission Control Protocol/Internet Protocol, or TCP/IP).

Functionality of the manual segmentation tool. A program for manual segmentation of tomographic images should take the following items into consideration:

- Image presentation
- Efficient segmentation tools for drawing on tomographic images
- Layout of the user interface
- Visualization of the segmentation results

Image presentation. Figure 2 shows the different sections for tomographic imaging. Three small windows appear at the top of the display screen (see Figure 3, top). One of these small windows is used to present the image data in the original form (transversal in CT; transversal, sagittal, or frontal depending on the image acquisition mode in MRI, see Figure 2). The two other small windows are used for multiplanar reconstructed views of the image data. The user can select one of the three views for presentation in the larger working window (see Figure 3, bottom left).

The selection of the spatial position of an actual section is indicated by reference lines in the two other corresponding views. The reference lines can also be used to select an interesting slice simply by moving the line with the computer mouse to the requested position. This method allows the therapist to get an impression of the patient's anatomy in all three dimensions. The working window is used for outlining anatomical structures on the images. For the segmentation of small structures, the images can be enlarged up to eight times their original size. The area to zoom in upon is selected interactively by moving a rectangle on the image. A presentation with a high zoom factor can be smoothed by low-pass filtering.

Segmentation tools. Segmentation of three-dimensional images into tumor volumes, target volumes, and organs at risk is performed in the working window by contouring "regions of interest" in successive transversal, sagittal, or frontal sections with a tomographic image as the background. The drawing instrument is the computer mouse. Three different drawing functions are

available: free-hand, polygon drawing, and ellipse drawing.

When drawing free-hand, a line on the screen follows the movements of the mouse. Unwanted lines can be deleted at any time. It is often necessary to keep a well-defined security distance to an anatomical structure, e.g., for the definition of the target volumes. For this purpose the cursor can be changed to a circle with the center being the active point and the radius corresponding to the desired safety margin. For drawing polygons, only the edge points have to be marked. With the help of a spline interpolation, additional points can be computed. In this way it is quite easy to create naturally formed curves that approximate anatomical structures. To segment structures with nearly elliptical shape, e.g., the eyes or the lenses, it is convenient to use the circle and ellipse-drawing functions. This significantly reduces the time needed for outlining these objects.

All manually drawn objects are mouse-sensitive and can be selected simply by clicking with the left mouse button on an object's outline. Selected objects can be interactively moved, copied, or deleted. It is also possible to scale and edit objects by clicking the mouse button on special control points and moving them to the required position.

As contours of an anatomical structure are often similar from one slice to the next, the copy and edit functions are very convenient for three-dimensional segmentation of medical images. After a region has been drawn manually in the first slice that cuts the anatomical structure under consideration, it is copied to a neighboring slice. In many cases, only small parts of the contours then have to be edited or redrawn to match the structure of the object in the new slice. (Spacing between slices is usually between 2 and 8 mm.)

Interpolation of contours. Because the interactive delineation of contours is a time-consuming task, a contour interpolation feature is included that calculates the preliminary contour in a slice, if contours in adjacent slices have already been defined. The interpolation algorithm is based on triangulation of the surface defined by two contours.[6] Using this feature, the definition of volumes of interest is performed as follows:

- Outline the region of interest in the first, central, and last slices of the CT or MRI data set.
- The program interpolates preliminary contours in all intermediate slices.
- Check the contour in an intermediate slice. If the interpolation is unsatisfactory, the user can edit the contour with the described editing functions.
- The program interpolates again in all slices where no user-defined contours exist.

Contouring along grey-level thresholds. Besides manual segmentation, objects with sufficient grey-level contrast can be segmented automatically using a contouring algorithm that has been taken from the literature.[7] This algorithm has been very helpful for the definition of the outline contour of the patient and for segmentation of bony structures or lung tissue in CT images. However, this algorithm completely fails for more complex objects with only small grey-level difference to the surrounding tissue.

User interface. The interaction is done completely with the computer mouse. All user interface objects may be manipulated by the mouse. Functions of the program can be started simply by activating the mouse on icons in a control panel or by selecting entries of pull-down menus. Dialogs with the user are handled with dialog boxes. A dialog box consists of interactive elements such as text entry fields, radioboxes, listboxes, and scaleboxes. The user has a good overview of all program functions and can fully concentrate on the work on the screen. This kind of user interface makes it fast and easy to use the program, a consideration that is particularly important for inexperienced users.

Visualization. During and at the end of the segmentation procedure, it is important to have the possibility of checking the shape and the consistency of the volume under consideration, preferably in a three-dimensional view. Currently, there are three visualization aids implemented in the program:

- While drawing a contour in the working window, the contour is also shown in the corresponding small overview window. Furthermore, all points where the slices perpendicular to the slice in the working window cut the previous and current contours are shown in the other two overview windows (Figure 3, top).
- A fifth window may be opened that shows a selection of up to 16 CT or MRI images with the

Figure 1 The steps involved in three-dimensional (3D) radiotherapy

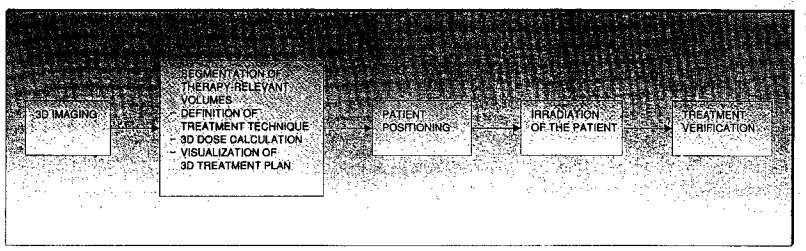


Figure 3 The user interface for manual segmentation (TOMAS program); overview windows with transversal, sagittal, and frontal views (top); working window and three-dimensional model window (bottom)

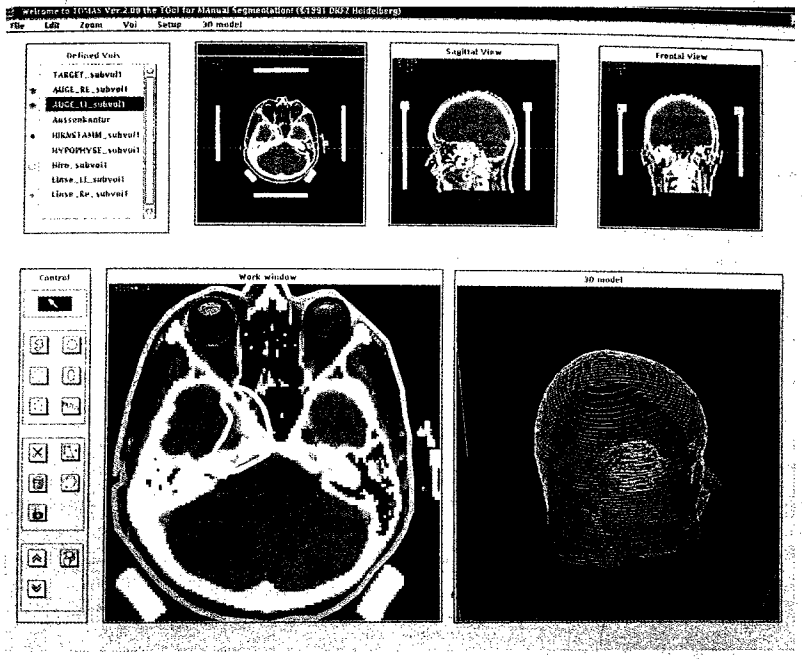


Figure 2 Transversal, sagittal, and frontal sections in tomographic imaging

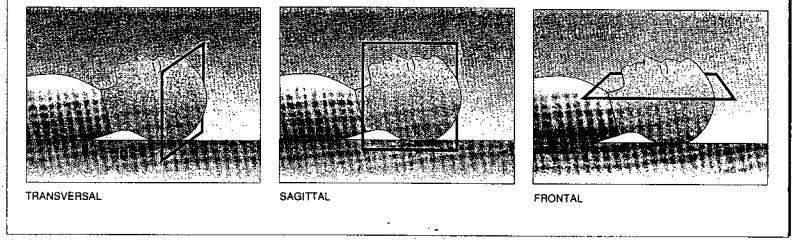
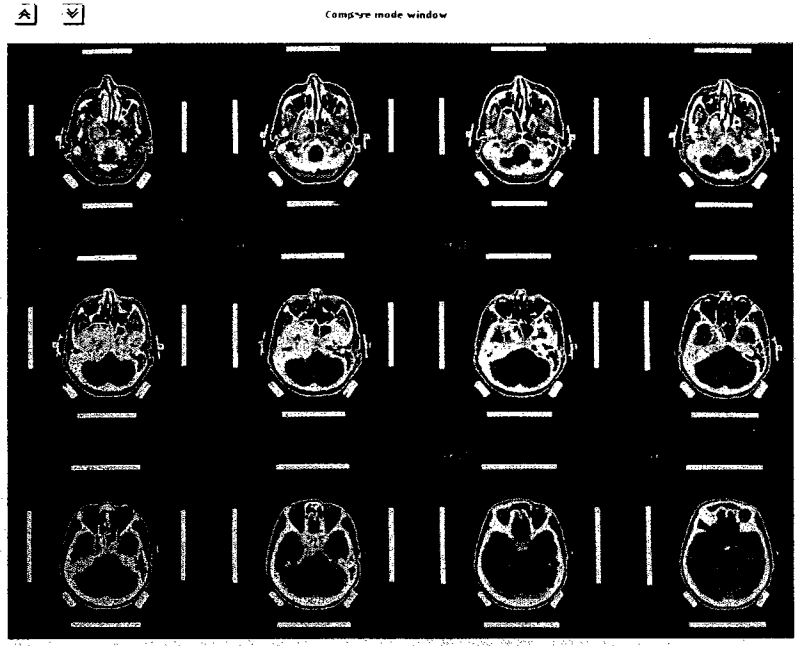


Figure 4 Compare mode window (TOMAS program) showing 12 CT slices



LIFT, PLEASE

Figure 5 Transversal cranial MR image (left); edge/region segmentation overlaid (right); Canny edges in green, region grower edges in yellow

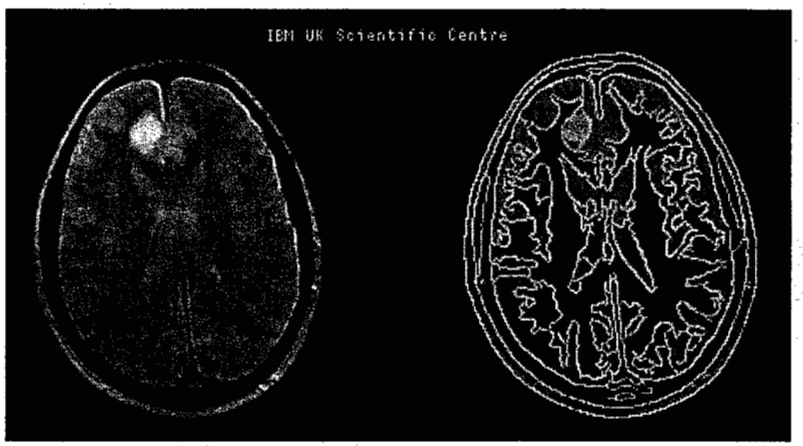


Figure 6 Comparison of manual segmentation with edge/region segmentation; red shows agreement, yellow is that part of the manual segmentation not found by the edge/region algorithm

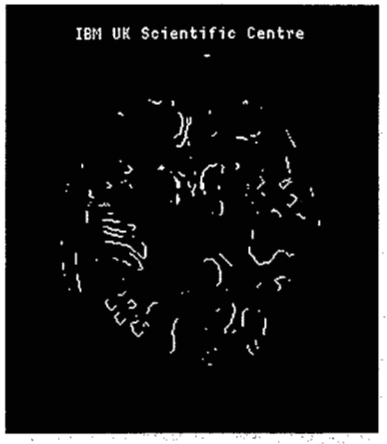


Figure 7 Differing segmentation at different slider positions

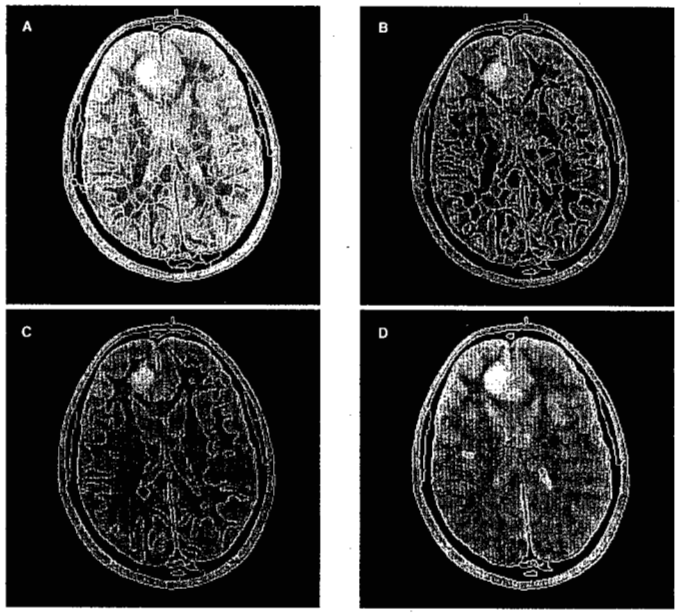


Figure 8 Defining a region

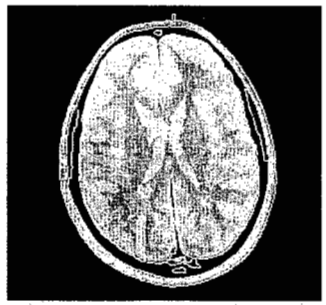
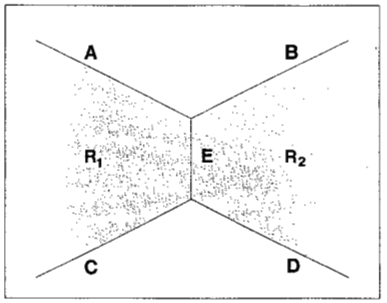


Figure 9 Merging adjacent regions and edges



contours of the different volumes that have been segmented previously (Figure 4).
- Another window may be opened that shows a perspective view of the different contour stacks, and thus provides a three-dimensional impression of the segmented volumes (Figure 3, bottom right).

Summary. Apart from the simple but limited contouring algorithm mentioned above, it is obvious this procedure is very time-consuming for the radiotherapist if the complete three-dimensional picture is to be built up. There is a clear requirement for more sophisticated computer algorithms to automate this process as far as possible. The saving of time that will result will enable more patients to benefit from the full three-dimensional treatment, with consequent improvement in tumor control and reduction in side effects.

The remainder of this paper reports on work that has been carried out toward this goal. A brief introduction to computer image segmentation is given first, followed by a description of a method for edge/region-based segmentation that was produced by the IBM UK Scientific Centre during the first phase of COVIRA. This has since been developed into an interactive system for hierarchical probabilistic segmentation, which is fully described in the following section.

Image segmentation

Segmentation is an important first step in analyzing an image. A segmented image represents the information in the image by delineating the boundaries between regions that correspond to different objects. This allows the amount of information to be reduced, and provides a description of the image contents better adapted to image analysis. The problem of segmenting an image into regions bounded by edges is classical in the fields of image interpretation and computer visualization, and no satisfactory, general solution has been found.

Two principal methods of image segmentation have been employed, using *edges* or *regions*. The first type corresponds to finding the local discontinuities in the grey-level intensity function of the image. These local discontinuities (or points of high gradient) characterize the abrupt intensity variations that often indicate the presence of an object in the image. Many techniques for edge detection have been developed, their purpose being to describe the image by a set of points that constitute the boundaries of different objects in the image.

The second type identifies homogeneous regions in the image. The idea underlying this approach is to describe the image by segmenting it into sets of points that possess certain properties of homogeneity. Different types of homogeneity have been used to characterize regions for the purpose of detection. For example, *smooth regions* can be detected on the basis of their roughly constant or perhaps smoothly varying intensity, while *textured regions* can be detected on the basis of the similarity of the texture across them. The use of texture detectors is not considered further here.

From a theoretical point of view, the two approaches (of edge and region detection) are complementary, in the sense that knowledge of the edges allows the regions to be deduced and vice versa. However, in practice the two methods give rise to different algorithms that produce different results, each having its own peculiarities. Some simple images can be satisfactorily segmented using edges alone, although gaps may be created when imaging noise disrupts the performance of the edge detector. The boundaries of smooth regions may identify edges not detectable by an edge detector; however, they may be false boundaries that have no physical significance.

It is better to combine information from the edge and region detectors, in order to produce a more reliable result. The edge detector uses information in a small locality to detect points at which the image intensity gradient is high, whereas the region detector exploits information gathered from many more points (and can, for example, fill gaps left by an edge detector). A system for edge/region-based segmentation was implemented by the IBM UK Scientific Centre for use in the first phase of the COVIRA project.

Edge/region-based segmentation

A survey of the literature has not revealed a definitive method for combining edge and region information to improve the segmentation. Basically, there are two possibilities to consider:

- To give priority to the edge information
- To give priority to the region information

It was decided to take the edge information first, and then use a region-growing algorithm to fill in the discontinuities in the edge map. The edge detector chosen is described in Reference 8 by Canny and is one of the most successful edge detectors today. One of the major advantages of using Canny's method in this application is that it gives clean edges with good localization. It has some deficiencies, notably in the area around junctions. However, this deficiency has been rectified, using the method discussed in Reference 9.

The Canny operator. The Canny operator finds edges by locating points of maximum gradient (edgels) in the grey-level intensity of the image. Thresholding is needed to separate the useful, stronger edges from weaker edges due to noise in the image. A hysteresis algorithm is used in which all edgels above an upper threshold are retained, as are weaker edgels that are connected to these strong edgels, provided that the weaker edgels exceed a lower threshold. The choice of suitable thresholds was originally made on an empirical basis, and later confirmed by a gradient descent method. These thresholds proved to give satisfactory results on more than 20 images from 6 different patients.

This process inevitably leads to gaps in the edges where there is insufficient contrast in the image. A region grower can be invoked in an attempt to repair these gaps by using information from a wider area, not just locally at the edge. A state-of-the-art region-growing algorithm was used, based on the work in Reference 10, in which smooth quadratic patches are fitted to the regions. (This is more general than simple thresholding, allowing second-order edges, missed by gradient-based edge detectors, to be recovered.) In practice we found that constant smooth patches gave better results on our MR images. The edge information is used to generate seed points for the region grower. It also limits the growth of the regions, thus intimately linking the two processes.

Region detection process. There are three distinct phases in the region detection process, as follow:

- A coarse resolution phase in which 2×2 parameterized patches (mean and variance) are merged on the basis of the likelihood ratio described in Reference 10. The Canny edges define the starting point for a new region to be as far as possible from the *end* of any edge, thus ensuring that the growing region is as large as possible before it encounters a broken edge. At each stage the most likely new patch is merged into the region. Regions are not allowed to cross Canny edges, and growth is stopped when there are no more patches that can be added without exceeding a threshold on the likelihood ratio.
- A second phase extends the above process to fine resolution (single pixel), continuing on from the first phase into unvisited areas of the image. (The unvisited areas are those 2×2 patches that contain Canny edges.)
- It is possible that parametrically homogeneous regions of the image are split into two or more regions, for example where the Canny edges form a bottleneck through which the region grower cannot pass. Therefore we have a third phase in which adjacent regions are merged, unless they have some small number of Canny edgels between them. (This reflects our strategy that the region grower is used to supplement the original edge information.)

A typical result of this process can be seen in Figure 5. On the left is a transversal MR image of the brain, in which a tumor is clearly visible. On the right is the computer segmentation, with the Canny edges shown in green. Gaps in these edges closed by the region grower are shown in yellow. For further details of this process, refer to Reference 11.

Clinical validation. In order to assess the clinical usefulness of the results, the results were compared to manual segmentations provided by radiologists from the Gregorio Marañon Hospital, Madrid. The radiologists were asked to mark boundaries in the images that they considered useful. Allowing an error of one pixel between the two segmentations, we typically observed that 85 percent of the edgels marked by the radiologists were found by the edge/region segmentation algorithm. This is shown in Figure 6, which shows the manual contours from the radiologists. Those parts found by the computer are marked in red, the remainder in yellow. Note that about 35 percent of the computer edgels did not correspond to any manual edge. It is clear that no amount of manipulation of the various thresholds in the algorithm would produce 100 percent correspondence. This is in part due to the fact that the radiologists used their clinical knowledge in deciding which boundaries to mark. This knowl-

edge is, of course, not available to our data-driven process.

Conclusion. We were led to the conclusion that an interactive process is required, in which the clinician is enabled to input clinical expertise in a simple, straightforward way. The interest then centers on how to exploit automatic methods for detecting edges and regions in order to minimize the manual interaction required. Ideally, one would like the human user to feel that the interactions are a way of exercising professional judgment—as, for example, adjusting a boundary to reduce radiation risk to the optic nerve—rather than a tedious chore that the machine ought to be able to carry out.

An interactive process is required to input clinical expertise.

A method developed by the UK Scientific Centre to do this is described in the next section.

Hierarchical probabilistic segmentation

The work described above, using a combination of edges and regions, gave good results and compared favorably with other methods, but it is not fast enough for interactive use. (On an IBM RT 6150 Model 135 workstation, the edge detector takes about one minute, and the region detection up to 60 minutes for a 256 × 256 image.) Also, no account was taken of the edge strength information. In the new method each edge is given a confidence in the form of a percentile (between 1 and 100 percent) based on its likelihood calculated from Bayesian belief updating, an application of methods described in Reference 12. Region variance information is incorporated into this likelihood in addition to edge strength. We have also experimented with including the results of symmetry matching and adding the additional information present in a second sensor channel in MR imaging.

The user interface exploits the confidence percentile by means of a slider. By adjusting it using the pointer, different segmentations can be obtained, as in Figure 7. These adjustments can be made in real time using a lookup table. In this way, a hierarchy of segmentations can be readily presented. Figure 7A shows the slider set to the lowest probability so that all regions are displayed. Moving the slider to increase the probability causes more and more regions to be combined (Figures 7B and 7C) until only the most likely are displayed (Figure 7D). The user can select a region at any particular setting of the slider, as in Figure 8. If desired, the region can be merged with others. Once a region is completed, any interior edges are removed and they no longer appear when the slider is readjusted.

Bayesian probabilities. A fundamental question is how information found from different sources using automatic methods can be combined in a manner that a human user can readily understand. In fact, integration of information is an important consideration whether human interaction is to take place or not. It is a topic of current research in machine visualization and in artificial intelligence work generally.

In principle, the method of Bayesian probabilistic belief updating[12] appears to provide a promising answer because it allows the weighing of different evidence strengths. For example, a weakly detected edge should be strengthened if there is supporting evidence from the surrounding regions. But it is not a trivial matter to apply the theory efficiently to two-dimensional data. One customary approach is to perform iterative convergence using the Markov random field model. Such iterations are notoriously slow to converge. The difficulty lies in the fact that belief propagation works well in (one-dimensional) chains but not in (two-dimensional) fields.

It is therefore more tractable to deal with the edges, since these are mostly arranged as one-dimensional chains. In Canny's method, a simplified form of belief updating takes place in the hysteresis algorithm, which propagates support from strong edge points to weaker neighbors. (We experimented with several more sophisticated forms of belief propagation but failed to better the results achieved with the standard hysteresis algorithm.)

Probabilities of edgels. The main result of the edge detector is to determine that certain points in the image are edge points, or *edgels*, and to give them a strength based on the local intensity gradient. The standard hysteresis algorithm then retains edgels that are stronger than an upper threshold and also retains weaker edgels that are connected to these strong edgels, provided that the weaker edgels are at least as strong as a lower threshold.

In our experiments, strengths ranged from 1 to over 150 (based on intensity values of 0 to 255) and the thresholds were typically 10 and 25. In the 256 by 256 images we often deal with, one may find perhaps 3000 edgels stronger than 25, another 2000 or so in the range 10 to 25, and maybe 10 000 below that strength. In the crucial middle range, the hysteresis algorithm will retain about two-thirds of the edgels, and we produced statistics to show a good correlation with results obtained manually. Below the threshold, we find a smaller proportion of useful edgels. (Some information about calibration against manual results appears in the section on calibration later in this paper.)

Rather than discard these useful edgels with the other weak ones as the standard algorithm does, we retain them with low probabilities, so that the information is not lost. Using the interactive slider technique, edges containing these unlikely edgels tend to appear at the bottom of the hierarchy (the lowest slider setting) and so may easily be discarded if the user desires. The edge strengths are converted to probabilities so that the information discovered by the hysteresis algorithm can be combined with the strength information in a consistent framework. Hence, edgels that satisfy the threshold criteria are given higher probabilities than the others. The probabilities are obtained from a table created by the calibration process (see the later section).

Meanings of the probabilities. The probability of an edgel needs to be defined. We adopt a task-oriented definition, based on the need to maximize the productivity of later processes and especially of the human interaction. If the probability of an edgel is 10 percent, this means that there is a 10 percent chance that it will be retained by later processes, including the human interaction.

Probabilities can be revised as new evidence is considered. This is detailed in the section on combining information.

Initial segmentation. We need to be able to combine evidence and weigh alternative hypotheses effectively, avoiding the classical AI bugbear of the combinatorial explosion. This needs to be done quickly, since a human user interaction is also required. The user interface also needs to be able to present choices without cluttering the screen with too much derived information.

In order for the Bayesian belief updating to work within these constraints, an initial segmentation is carried out automatically. This initial segmentation results in a set of regions, with each pair of regions separated by an edge (made up of edgels). The initial segmentation is at the bottom of the hierarchy of segmentations and is displayed when the user sets the slider to the lowest value. Although the eventual segmentation will be different from the one produced initially by this method, there is likely to be a useful number of reliable edges and regions detected that need no further change. For example, in the cranial MR image of Figure 7, the major boundaries of the skull and the tumor are reliably detected and the ventricular region is a combination of a small number of subregions.

It can be seen that, instead of reasoning at the level of pixels, edgels, or patches of an image, later processes—including human interaction—can work mainly at the level of regions and edges. This is likely to be much more convenient. Two regions can be merged, eliminating the edge that separates them. A region can be split by defining a new edge within it. These later processes avoid a complex interaction with the early (low-level) edge and region detectors by virtue of the summarized result provided in the confidence ascribed to each edge, which is calculated from the edgel and pixel data.

Closing gaps. A prerequisite for this approach is to be able to find closed boundaries of regions. Unfortunately the edge detection method used does not provide closed boundaries. The detector finds edgels that can then be joined into edge segments. Further work must be done to close the gaps between them in order to find the regions. There are other methods that do find closed regions, notably the one due to Marr and Hildreth,[13]

but the positional accuracy is poorer than is obtained with the method by Canny.

We have experimented with several ways of closing gaps between edges. So far, the most successful has been a straightforward search among neighboring feature points, followed by straight-line interpolation. A feature point may be an isolated edgel, a corner, or a terminator; a terminator is an end of an edge segment.

Once every terminator and isolated point has been joined to another edgel, the existing edge segments plus the interpolated edges form a network of closed boundaries. A flood algorithm can then be used to label every pixel (except the edgels) in the image as belonging to a region, each region being identified by a numeric label. This turns out to be a fast method of finding regions, taking approximately half a second (for a 256 by 256 image) on an IBM RISC System/6000 Model 530H. The complete process, starting with the image up to the point where the user can commence interactive work, takes about five seconds.

We can then label the edgels too, because we define every edge as the boundary between a pair of regions.

Combining information. Once a set of closed boundaries has been found, it is easy and efficient to gather statistics, such as mean and variance, from all the pixels in the regions. This information can be combined with the edge evidence in a Bayesian framework because a convenient basis now exists for hypothesizing and reasoning about new edges and regions.

Three factors are combined to determine the probability of an edge:

- The average probability of the edgels in the edge, a
- Whether the edge includes an interpolation, i
- A region variance statistic, s

The statistic s is found by considering the pair of regions R_1, R_2 which separate the edge E, together with the combined region $R = R_1 \cup R_2 \cup E$. If ν is the variance of R and m_1, m_2, m are the means of R_1, R_2, R, respectively, then

$$s = \frac{(m_1 - m)^2 + (m_2 - m)^2}{\nu}$$

Two tables are produced by the calibration program giving conditional probabilities $p(H|a,i)$ and $p(H|s)$ where H represents the hypothesis that the edge will be retained by later processes, including human interaction. Using these tables, we calculate the posterior odds on each edge given the statistics a, i, s, assuming that the edge and region evidence are conditionally independent, as

$$(H|a,i,s) = O(H)L(a,i|H)L(s|H)$$

Here

$$O(H) = \frac{p(H)}{1 - p(H)}$$

are the prior odds before looking at any evidence, and

$$L(s|H) = \frac{p(s|H)}{p(s|\neg H)}$$

is the likelihood of a particular value of the statistic s given a hypothesis H. To get the likelihood, we note that

$$L(s|H) = \frac{O(H|s)}{O(H)}$$

which gives an expression in terms of $p(H|s)$ and similarly for $p(H|a,i)$.

Hierarchy of segmentations. Once we have obtained a probability for each edge in the initial segmentation, we have defined a hierarchy of segmentations, because we can remove one edge at a time (the least likely), each time merging a pair of regions until the whole image becomes one region. At each stage we have a complete segmentation in which all the edges form closed boundaries to the regions.

Each time an edge is removed, two edges and one region will normally be merged. Thus if E separates regions R_1 and R_2 and A, B, C, D are other edges as illustrated in Figure 9, removing E

will cause two new edges $F = A \cup B$ and $G = C \cup D$ to be created as well as the combined region $R = R_1 \cup R_2 \cup E$. Statistics must be accumulated for these new entities and probabilities must be calculated for the combined edges F and G.

An individual edgel b in B could be part of several hierarchically superior edges above F. Eventually, though, as we progress up the hierarchy of segmentations removing edge after edge, the one containing b will disappear. Thus the edgel b belongs to several edges, one of which is the most superior in the hierarchy. Call it S.

When the interactive user moves the slider up, the less likely edges disappear, and, when it is moved the other way, they reappear. This must always be done in such a way as to preserve the integrity of the segmentation, i.e., so that regions are always enclosed and edges do not have dangling ends. In order to achieve this while avoiding ponderous data structures, individual edgels are given the probability of the most superior edge to which they belong. Hence edgel b is given the probability of edge S.

The edge probabilities are converted to percentiles in order to maximize the discrimination available by use of the slider. By assigning the values to individual edgels, a look-up table can be used to manipulate what is displayed on the screen during interactive sessions. This ensures much better performance than would be possible if the screen were redrawn every time the slider is moved.

Interaction. The interactive user may move the slider at any time to change the visible edges. In addition, the following actions can be selected. Each of these actions uses the edges visible at the time, according to the setting of the slider.

- *Define region.* After selecting this action, the user should click the mouse in one or more regions; if they adjoin, they will be merged.
- *Define edge.* After selecting this action, the user should click the mouse on two or more points. Straight lines will be drawn between them. The ends of the newly defined edge will be truncated or extrapolated to join with existing edges.
- *Save.* The edges and regions visible on the screen will be saved as the result of the interactive session. This applies both to the edges and region boundaries defined by user actions and to the edges visible as a result of the setting of the slider.
- *Quit.* No result is saved.

A newly defined region or edge can be undone by clicking the right mouse button before commitment takes place. An edge or region is committed when the user selects another action. After that, the edge or the boundary of the region appears in a different color and is unaffected by the slider. Edges inside the region no longer appear when the slider is moved. It is still possible for the user to make changes by defining new edges or merging regions.

The interactive user therefore has complete control while having access to the best segmentation results available from a combination of automatic algorithms.

Calibration. The result of the interaction can be used in our application for the preparation of a treatment plan. The result can also be used to calibrate the tables of probabilities for use on subsequent images. Initially the tables were prepared from the results of segmentations performed manually by collaborating radiologists in Madrid (Kuhn[14] et al., 1990) but now the probabilities can be refined from the more relevant statistics generated from using this tool.

Hence, the needs of particular specialities or the requirements of differing imaging modalities (MR, CT) can be accommodated adaptively.

The required conditional probabilities are as follows. The meaning is the probability that the edge or edgel will be retained.

- An edgel, given its strength and whether or not it is supported by connection to a stronger edge (used in Canny and hysteresis algorithms)
- An edge, given the average probability of its constituent edgels and whether or not it includes an interpolation
- An edge as the boundary between two regions, given the variance statistic

The calibration process matches edgels and edges in the automatic segmentation with those resulting from interactions. Counts of matched and unmatched edgels and edges are accumulated against the strengths and other statistics and the

probabilities are readily derived from these quantities. Counts can be accumulated from many images in order to obtain more reliable numbers. Where totals are too small, consolidation and interpolation are performed to ensure that the probability distribution functions are monotonic.

Matching. As a matter of terminology we say that matching is done from automatic edgels or edges to manual ones, the latter being produced either with this tool or otherwise.

Since the initial work was done on manual segmentations performed independently of this tool, there is some flexibility in allowing for matching edgels, provided that they are within a 1.5 pixel distance (the distance between two diagonally adjacent pixels).

When matching edges, we require that an automatic edge must overlap with one or more manual edges. Recall that an edge is the boundary between a pair of regions. If there is a manual region that has pixels on both sides of an automatic edge, then that edge does not match.

Conclusion

We have outlined one of the problems that exists in three-dimensional radiotherapy today, that of defining tumor volume, target volume, and organs at risk in three-dimensional CT and MR data sets (a necessary precursor to producing a complete three-dimensional radiation treatment plan). We have described work that has been carried out in an attempt to solve this problem by replacing time-consuming manual image segmentation by computer-assisted interactive image segmentation, in which the computer relieves the user of much of the tedious work, and enables the operator to use clinical judgment to achieve the desired result. Using a state-of-the-art workstation—the IBM RISC System/6000—we have shown that true interactive speeds are achievable in this context.

Other algorithms under investigation at the IBM UK Scientific Centre include another development of the original edge/region segmentation method. In this case the user interactively indicates seed points for a statistical region growing process, which can be inhibited by edges detected in the image. This method will be extended to the statistical growth of three-dimensional volumes on a slice-by-slice basis. Particular attention is being given to the development of direct three-dimensional segmentation algorithms, and we have produced a three-dimensional version of the Canny edge detector with hysteresis. The main research problem here is how to visualize and interact successfully with the three-dimensional data.

Although this work is primarily concerned with medical images, the same algorithms can of course be applied to applications in any field where it is necessary to extract information from complex data sets prior to further analysis.

In the main phase of COVIRA, our methods, in company with algorithms produced by other partners in the project, will be subjected to extensive clinical evaluation at the German Cancer Research Center (Heidelberg, Germany) and the Royal Marsden Hospital and Institute of Cancer Research (Sutton, UK). By this means we ensure that the work has a clinical focus, for the ultimate benefit of the patient.

Acknowledgments

We acknowledge the support given to this work by the Commission of the European Communities under the AIM program. Other partners in the AIM Exploratory Action Project A1011 COVIRA were Philips Medical Systems (Hamburg, Germany), the University of Hamburg (Germany), the University of Genoa (Italy), and the Gregorio Marañon General Hospital (Madrid, Spain) who supplied the MR images and manual segmentations.

We would also like to acknowledge the contributions of G. Brelstaff and M. Ibison to the edge/region-based segmentation work, and of W. Dickson to the hierarchical probabilistic segmentation work.

Cited references

1. W. Schlegel, "Computer Assisted Radiation Therapy Planning," in NATO Series **F 60**, *3D Imaging in Medicine*, K. H. Hohne, Editor, Springer, Berlin, Heidelberg (1990), pp. 399–410.

2. H. Suit and W. Du Bois, "The Importance of Optimal Treatment Planning in Radiation Therapy," *International Journal of Radiation Oncology and Biological Physics* **21**, 1471–1478 (1991).
3. M. Goitein and M. Abrams, "Multi Dimensional Treatment Planning: I. Delineation of Anatomy," *International Journal of Radiation Oncology and Biological Physics* **9**, 777–787 (1983).
4. B. Bauer-Kirpes, W. Schlegel, R. Boesecke, and W. J. Lorenz, "Display of Organs and Isodoses as Shaded 3D Objects for 3D Therapy Planning," *International Journal of Radiation Oncology and Biological Physics* **13**, 135–140 (1987).
5. R. Bendl, J. Pross, and W. Schlegel, "Application of Computer Graphics in 3D Radiotherapy Treatment Planning: Manual Segmentation and Virtual Therapy Simulation," in *Proceedings of the Third Joint Seminar on Medical Informatics and Bioengineering*, H. Dickhaus, T. Leo, and P. Russo, Editors, University of Ancona (1990), pp. 20–32.
6. H. N. Christiansen and T. W. Sederberg, "Conversion of Complex Contour Line Definitions into Polygonal Element Mosaics," *Computer Graphics* **12**, 187–192 (1981).
7. W. V. Snyder, "Contour Plotting," *ACM Transactions on Mathematical Software* **4**, No. 3, 290–294 (1978).
8. J. F. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-6**, No. 6, 679–698 (1986).
9. D. Li, G. D. Sullivan, and K. D. Baker, "Edge Detection at Junctions," *Proceedings of the 5th Alvey Vision Conference*, Reading, UK (1989), pp. 121–125.
10. J. F. Silverman and D. B. Cooper, "Bayesian Clustering for Unsupervised Estimation of Surface and Texture Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-10**, 482–495 (1988).
11. G. J. Brelstaff, M. C. Ibison, and P. J. Elliott, "Edge-Region Integration for Segmentation of MR Images," *Proceedings of the British Machine Vision Conference*, Oxford (1990), pp. 139–144.
12. J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufman Publishers, San Mateo, California (1988).
13. D. Marr and E. Hildreth, "Theory of Edge Detection," *Proceedings of the Royal Society of London B* **207**, 187–217 (1980).
14. M. H. Kuhn, I. C. Carlsen, S. Dellepiane, P. Elliott, F. Galvez-Galan, H. Neumann, and H. S. Stiehl, "Computer Vision in Radiology (COVIRA): Knowledge Based Segmentation and Interpretation of Cranial Magnetic Resonance Images," *AIM Euroforum*, Sevilla, Spain (December 1990), pp. 169–180.

Accepted for publication September 17, 1992.

Peter J. Elliott *IBM UK Scientific Centre, Athelstan House, St Clement Street, Winchester, Hampshire SO23 9DR, United Kingdom (electronic mail: pjelliot@winvmd.vnet.ibm.com).* Dr. Elliott graduated with first-class honors in physics at the University of Sussex, UK, in 1968. He joined IBM UK Laboratories Ltd., Hursley Park, to work on thin film magnetic recording heads for disk files. Sponsored by IBM, he carried out research on amorphous semiconductor thin films at the Cavendish Laboratory, Cambridge, UK, leading to his Ph.D. in 1975. He subsequently worked on the development of a number of disk storage devices, gaining experience in technical project management. In 1989 he joined the IBM UK Scientific Centre, working on medical image analysis. He was instrumental in setting up the European consortium for the main phase of COVIRA (COmputer VIsion in RAdiology), and is currently project manager of the consortium.

John M. Knapman *IBM UK Scientific Centre, Athelstan House, St Clement Street, Winchester, Hampshire SO23 9DR, United Kingdom (electronic mail: knapman@winvmd.vnet.ibm.com).* Dr. Knapman graduated in mathematics as a wrangler at the University of Cambridge, UK. He joined IBM United Kingdom as a systems engineer, eventually specializing in databases and fourth-generation languages, providing customer consultancy across Europe and in the United States. He earned his Ph.D. in artificial intelligence at Edinburgh University during this time. He joined the Scientific Centre in 1985 to work on image analysis and computer vision, applied to robot guidance, industrial inspection, and medical image understanding.

Wolfgang Schlegel *German Cancer Research Center (DKFZ), Im Neuenheirmer Feld 280, D-6900 Heidelberg, Germany.* Dr. Schlegel studied physics, mathematics, and chemistry at the Free University of Berlin from 1964–1968. He received his diploma in physics at the University of Heidelberg in 1970, and his Ph.D. in nuclear physics at the Max Planck Institute for Nuclear Physics in 1972. After a position as a research associate at the Max Planck Institute in Heidelberg he joined the German Cancer Research Center in 1973, where he initiated digital image processing of biomedical images and computer simulation of procedures in cancer therapy. Since 1976 he has held the position as head of the group for radio-oncological computing at the Institute of Radiology of the German Cancer Research Center. Since 1988, he has had the Venia Legendi for teaching medical physics at the University of Heidelberg. Dr. Schlegel is author of more than 100 scientific papers on image processing and computer-assisted treatment planning in radiation therapy and neurosurgery.

Reprint Order No. G321-5489.