



0031-3203(94)00124-3

# KNOWLEDGE-BASED ORGAN IDENTIFICATION FROM CT IMAGES

MASAHARU KOBASHI and LINDA G. SHAPIRO\*

Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195, U.S.A.

(Received 29 July 1993; in revised form 8 September 1994; received for publication 23 September 1994)

**Abstract**—This paper describes a new knowledge-based procedure for identifying and extracting organs from normal CT imagery. Our procedure differs from previous attempts in its use of a wide variety of knowledge about both the anatomy and the image processing operations. The system features the use of constraint-based dynamic thresholding, negative-shape constraints to rapidly rule out infeasible segmentations, and progressive landmarking that takes advantage of the different degrees of certainty of successful identification of each organ. The results of a series of tests on training data of 100 images from five patients plus additional test data of 75 images from three more patients indicate that the knowledge-based approach is promising.

Knowledge-based vision  
Dynamic thresholding

Medical imaging

CT images

Object recognition

## 1. INTRODUCTION

Patients who are scheduled to receive radiation treatment for cancer undergo CT scans which produce a sequence of images, each representing one slice through the three-dimensional (3D) organs and vessels being scanned. From this sequence of 2D images, it is possible to estimate the 3D structures through which the slices were taken and, if the estimate is good, to determine their approximate locations and volumes. This extraction of structures from parallel CT images of the patients is an important first step in the creation of patient-specific models that can be used by treatment planning programs to deliver maximum dosage to the tumor and minimum dosage to critical anatomical structures. Currently this step is performed manually by technicians called dosimetrists who use an interactive device, such as a mouse, to trace the contours of each organ of each image of a patient data set. Although many attempts have been made to automate the extraction of anatomy, no techniques have been successful enough to replace the current manual methods. Since these methods take up to half the planning time, there remains a great need to speed up, if not completely automate the process.

Automatic segmentation of CT images is a challenging problem in computer vision. While the positions of the organs within each slice can be predicted the organs are flexible and the 2D contours they exhibit can vary. Furthermore, the boundaries separating organs from their surroundings are not always clear; even humans have to guess where some portions of

boundaries are located in some images. Standard techniques such as absolute thresholds, edge-finding, or region growing acting blindly on the gray tones of an image are not powerful enough. Instead the system needs a knowledge-based control structure that can use standard techniques in a goal-directed and informed manner and can evaluate its own success or failure.

The goal of our work is to develop a knowledge-based recognition system that utilizes knowledge of anatomy, knowledge of the imaging process, and knowledge of the effects of various image processing operators to extract the organs from parallel CT images. To this effect we have developed an experimental system that locates the major organs in sets of images through the abdomen. The major features of our system include (1) dynamic thresholding controlled by feedback information on various properties of image regions, (2) the use of negative shape constraints (constraints that rule out certain impossible shapes), and (3) progressive landmarking that extracts organs in order of predicted success and uses already-found organs to help locate other organs. This paper describes the knowledge-based system we have implemented. Section 2 describes the characteristics of the problem. Section 3 summarizes the related literature on the knowledge-based approach. Section 4 indicates the knowledge used by our system. Section 5 discusses the algorithms we developed, and Section 6 describes the performance of the system.

## 2. CHARACTERISTICS OF THE PROBLEM

Our task of constructing an automated dosimetry system is equivalent to producing a definition in computable terms for each organ as a homogeneous region

\*All correspondence to: Prof. Linda G. Shapiro, Department of Computer Science and Engineering, FR-35, University of Washington, Seattle, Washington 98195, U.S.A.

distinguishable from other organs/tissues in CT images and implementing a segmentation according to that definition in a computerized image processing system. Although the history of CT imaging is decades old, consistent definitions for each organ in precise, computable terms do not exist. However, physicians are able to recognize real organs easily and fairly accurately. Since physicians' jargon is not necessarily convertible to precise and computable terms, our first step is to construct computable definitions of the organs.

Secondly, we must consider the mapping done by the CT imaging process whose domain is the real human body and whose range is the CT image. CT imaging transforms real 3D objects into 2D gray tone images of their slices. The transformation is not straightforward in that the anatomical classifications of organs/tissues do not precisely correspond to simple segmentations based on gray tone similarities in CT images. In reducing our task to the above two steps, we encounter the following difficulties:

(1) Standard gray-tone-based segmentation procedures do not necessarily produce regions corresponding to organs;

(2) There are very few shape invariants;

(3) The absolute gray tone levels observed in each organ vary widely with the instance of observation;

(4) There is no precise and objective ground rule for performance evaluation.

The first difficulty is due to the fact that two different organs can have the same or very close gray tones in CT images. Some organs have wide ranges of internal gray tone variations, while others have narrow ranges of variation. CT images can also have some aberration due to the limitations of the individual CT scanner. Our system cannot produce a segmentation when two different organs having the same gray tone are adjacent. This shortcoming is present even in human work, since humans do some intelligent guess work in such cases.

The second difficulty prohibits an approach that tries to find shape invariants for each organ. Unlike the rigid objects that are employed in many computer vision studies, most human organs have few *computable* and *stable* invariants. Although physicians may claim that they can identify organs by their shapes, what they may describe as invariants are lacking in *computational feasibility and stability*. (There are a few excep-



Fig. 1. Various shapes of kidney.



Fig. 2. Concept of negative shape constraints. Clearly these shapes are not of kidney's. Each of the above three shapes is formed by connecting a kidney with other organs/tissues. Such connections take place if two low a threshold is used in thresholding.

tions such as the eye sockets and the aorta, whose shapes are nearly perfect circles.) Figure 1 illustrates the variety of shapes that can exist for a single organ (the left kidney). In our system this problem is overcome by the principle of "avoiding positive shape commitment".

The third difficulty rules out the possibility of using a simple thresholding technique. The absolute gray tones of each organ depend on the individual instances. The same organ can display different gray tones if the setup of the CT scanner is different and/or a different scanner is used. This variation of gray tone is also caused by differences in the chemical contents of each organ, which depend on patients as well as on their physical conditions. Normalizing the gray tone level is one way to get around this problem. However, it does not always give satisfactory results, since for some organs even the normalized gray tone level has a wide enough range to confuse the identity of the organ. This problem is overcome in our system by taking advantage of a more stable property, the ordering of gray tone levels of organs. In the case where two organs have very close gray tones, other properties are used to distinguish one from the other.

Finally, just as the definition of each organ in the CT image is very difficult to formulate in computational terms, the performance of the system is difficult to measure objectively. One method of evaluation is to compare the results against the work of human dosimetrists. However this still has a problem in evaluating the degree of misidentification. Simply counting the number of mismatched pixels does not make a consistent and reasonable measurement, since qualitative differences in mismatch cannot be reflected this way. We used conservative subjective evaluation with three ranks for the results of our system as explained in Section 6.

To solve the above problems, several properties of organs and images are used. First, each organ has a fairly stable vertical and horizontal location and adjacency with other organs. Secondly, the ordering of organs by their gray tones is fairly stable, even though their absolute gray tones vary widely. Thirdly, each biological substance has a relatively narrow range of gray tones.<sup>(1)</sup> This property is used to determine a termination point in repeated feedback thresholding. Fourth, analysis of CT images is simpler than the usual 3D natural object recognition, because in the CT image analysis no complications are added by shading, range measurement, light sources, or reflection. The segmentation of each object in a CT image is similar to extracting a region of quasi-homogeneous particular gray tone range that reflects the electron density of the organ of interest.<sup>(2)</sup> This property justifies the basic approach of the proposed system, dynamic thresholding.

The fifth advantage is that there are relatively small numbers of objects to be recognized in a single CT image in our task. Therefore, the difficulty of the task can be substantially reduced by the use of progressive

landmarking, which reduces the search space efficiently. The sixth advantage is the existence of some stable landmarks, such as the spine and the aorta.

Finally, there are some useful shape constraints. There are few stable positive shape constraints: circularity for eye sockets and aorta, symmetry for the spine. On the other hand all organs have negative shape constraints (i.e., constraints on what shape an organ does *not* take). This is a constraint that defines only extremely abnormal shapes as impossible for each organ. Therefore, its reliability is high. Negative shape constraints are useful when used with repeated incremental thresholding, since when the threshold goes below an acceptable level, the shape of the region of the organ usually becomes extremely abnormal due to being connected to other organ/tissue. Figure 2 illustrates the concept of negative shape constraints.

### 3. RELATED RESEARCH

Medical image analysis is one of the areas of computer vision where domain knowledge plays a very important role, because localized pixel information obtained from CT images is often ambiguous and unreliable. Therefore, many systems, including the one described in this paper, use a knowledge-based approach.

The history of knowledge-based medical image analysis is older than the history of practical usage of CT imaging. One of the early studies in knowledge-based medical image analysis was done by Harlow and Eisenbeis<sup>(3)</sup> on radiographic image segmentation, when CT imaging was not yet available in hospitals. They proposed a top-down control system using a tree-structured model description containing knowledge about locations and spatial relations of parts/organs of the human body. In his thesis work, Selfridge<sup>(4)</sup> discussed image understanding systems in general and divided the causes of difficulties into problems of model selection, segmentation techniques, and parameter setting. In earlier work<sup>(5)</sup> he proposed an algorithm for detecting contours of the kidney and stressed the importance of domain knowledge for further enhancement of his system.

Shirai<sup>(6)</sup> developed a method for extraction of stomach regions from radiographs. He used a combined thresholding and edge-finding method to extract the stomach region with initial thresholds determined from analysis of the histogram of the intensity values over the image and knowledge of the usual size of the stomach. When a candidate region has been extracted, it is tested for size, position, and abnormal extension. If the region fails, the thresholds are changed in a feedback loop. This approach is very much in the spirit of our own, but is specific to a single organ and does not use as many different kinds of knowledge as we do.

More recently, the research conducted by Karssemeijer *et al.*<sup>(7)</sup> made effective use of domain knowledge incorporated in a semantic network of spatial relations among organs. This system has some

similarities to our system in the following respects: (1) it advocates the integration of the segmentation process with the recognition process, (2) it uses mathematical morphology operators effectively, and (3) the target slices are of the human abdomen, although their system is only for recognizing the spleen.

Bright proposed an interesting algorithm for segmenting objects in micrographs.<sup>(8)</sup> He used multiple thresholds to take advantage of the difference in gray tone between target objects and their background. Despite its computational expensiveness, this system was robust in applications concerning ambiguous and flexible objects such as biological cells and human organs. It is similar to the dynamic thresholding of our system with respect to its use of multiple thresholds to separate target objects from the other regions. A major difference is that our dynamic thresholding is controlled by feedback information on various properties of the focused regions, while Bright's method does not use a knowledge-based feedback loop.

Brinkley<sup>(9)</sup> uses a geometric constraint network approach to segmentation of CT images. He models the flexible two-dimensional shapes of organ contours by the allowable ranges of the lengths of radials from the centroid of the organ to points at evenly-spaced intervals around the boundary. He then uses a relaxation algorithm to compute the boundary points of the shape at each radial. His algorithm combines prior knowledge of organ shape with dynamic edge detection along the selected radials.

Finally, a promising approach to medical image analysis, although expensive in terms of data acquisition, was proposed by Chen *et al.*<sup>(10)</sup> whose system processes multi-modal imaging data. It employs images made by CT (computed tomography), MRI (magnetic resonance imaging), and PET (positron emission tomography) together with various domain knowledge and the knowledge of the correlation between biological substance and imaging results. A blackboard architecture is used to handle all the different types of knowledge.

The multi-modal approach has an advantage over the conventional single-mode methods, since different tissues having very close gray tones in a single mode cannot always be distinguished reliably with only one mode of data. Most of the current single-mode systems (including human dosimetrists) make an educated guess when the gray tones are ambiguous. Therefore, in order to improve the results of medical image analysis, additional reliable constraints are required.

#### 4. KNOWLEDGE FOR ORGAN RECOGNITION AND EXTRACTION

Our system was designed under the following two modeling principles. First, it employs only those properties that are meaningfully reliable and stable. This principle has significance particularly in the medical domain, since there are many quasi-stable properties that can be useful in some cases, while very misleading

in others. Most shape properties were discarded by this principle. Secondly, the system's capacity to discriminate one organ from others is enhanced by increasing the number of loosely constraining properties, rather than by using probabilistic refinement or setting precise constraints on a small number of properties. This principle is justified by the fact that no small number of properties can unmistakably distinguish an organ from others and that precise constraints on any property reduce the tolerance of the system to wide variations of instances. Thus, the modeling is done by describing each organ with a number of properties, each of which imposes very loose constraints. The specific properties used by our system to describe an organ include:

(1) Position in the ordering of gray tone levels among the organs. For example, bones have the highest gray tones.

(2) Relevant gray tone range. The relevant range is the range of gray tones containing the threshold that segments the organ correctly.

(3) Height of gray tone cliff. Suppose there is a gray tone that can be used as a threshold to make the whole contour of a target organ show up. (Every organ has such a gray tone, although the gray tone may not necessarily distinguish the organ from its surrounds.) Such gray tones usually fall in a certain range. This range is the height of the "cliff" for the target organ. While bones have very high cliffs, livers and spleens have relatively low cliffs.

(4) Location in terms of stable landmarks (the aorta and the spine). The coordinate system relative to these landmarks is used, since the absolute coordinate system does not work if the body is not positioned correctly in the center of the image. The location of an organ is represented by the center of gravity of the organ. The possible locations are restricted to those points that are within an acceptable range from the mean gravity center of the landmark.

(5) Adjacency with other organs. The relative position is specified by the angle of direction relative to the body-based coordinate system and the distance between the two organs in question.

(6) Size in terms of the area of the slice for each approximate level of slice. The size is given by the number of pixels in the area of the organ.

(7) Relationships with slices at other levels. The constraint used for this relation is the overlap ratio of the regions of interest between adjacent slices. This ratio must fall within a certain range depending on the width between adjacent slices. If two adjacent slices having an outrageous inclusion ratio are encountered, it would be a sign of an incorrect segmentation.

(8) Positive and negative shape constraints. Positive shape constraints describe the shape of acceptable objects, while negative shape constraints describe what is unacceptable. The only positive shape constraints we have found stable enough to use are the circularity of the eye sockets and the aorta and the symmetry of

the spine. The negative shape constraints used in this system are abnormal extension and abnormal compactness. These are useful constraints for employing with dynamic thresholding, since lowering the threshold a single step can cause the candidate region to connect with other tissues and thus to appear extremely abnormal.

The location, size, adjacency, overlap ratio and shape constraints are formulated from the anatomical definitions so that they are computationally executable. The properties are defined for each level of slice, and information concerning appropriate levels of slice for each organ is an input to the system. Each of the above constraints, except the first and the last, is given as a range that spans from 3–6 S.D.'s from the mean. Due to the very generous tolerance range, each constraint alone may not be able to distinguish a target organ. However, each target organ can usually be unambiguously detected by the conjunction of all the constraints.

### 5. ALGORITHMS

The system is composed of two major categories of module: the *brain* modules for control and the *engine* modules for processing. The engine modules implement the image processing and features extraction operators. The operators currently supported by the system are (1) thresholding, (2) morphological opening and closing, (3) the set operations of union, intersection, and difference, (4) the connected components operation, (5) property computation operations for location, size, positive shape constraints (such as circularity), and negative shape constraints (such as compactness) and (6) region identification/elimination based on these properties. The brain is made up of the domain knowledge base and the performance control module. The domain knowledge base contains statistical data for all the properties used in organ recognition. The domain knowledge base includes properties of the spine, the aorta, the right kidney, the left kidney, the spleen and the liver. The spine and the aorta are used as landmarks for locating the other organs, and the ordering is determined by the principle of progressive landmarking, locating the most reliable organs first. The current ordering is (1) spine, (2) aorta, (3) kidneys, (4) spleen and (5) liver. The knowledge base was compiled from a training data set of 113 slices from five patients.

The system was designed with the assumption that the original CT images are assigned consecutive slice numbers from the top slice to the bottom slice. The expected range of slice numbers for each organ is an input to the system. The extraction process proceeds vertically; that is, each organ is extracted from all slices before the extraction of the next organ begins. For each organ, except the spine and aorta, the initial slice to be processed is the slice that is closest to 2/5 of the organ's expected height from its bottom-most expected slice. The "2/5 from the bottom" heuristic for finding a

reliable starting position is based on our observation of the training set. The process goes through two iterations from that initial level, one upward to the topmost expected slice of the organ, the other downward to the bottom-most expected slice. For the spine and aorta, the extraction process goes from the lowest to the highest slice. At each slice the system goes through a series of dynamic thresholding operations that is controlled by feedback information on various properties of the candidate regions produced by the thresholding. The essence of the dynamic thresholding is as follows.

The system first sets the initial threshold to the high end of the relevant gray tone range predefined for the organ of interest and performs a thresholding operation with the initial threshold. It then performs a connected components operation on the binary image produced by the thresholding and checks to see if there is a region that satisfied all of the constraints. If there is, it reduces the threshold by a single step and repeats the same procedure, executing the connected components operation and checking all the properties.

The system repeats this process until the low end of the relevant gray tone range is reached or candidate regions disappear (due to becoming too large or connected to other organs/tissues). After exiting this loop of repeated thresholding, if there is an acceptable range of thresholds that provide candidates, the system selects a threshold from them that is most likely to produce the best-fit region to the organ of interest.

The selection of such a threshold is done by checking the rate of increase of the area of the candidate region with respect to the reduction of threshold. The rate of increase is computed for the gray tone interval that reflects the height of the gray tone cliff of each organ. The best threshold is selected out of the interval that has the least rate increase. This process is referred to as the "slope check" in the following descriptions. The final output is the candidate region created by that selected threshold.

The morphological routine was designed to be a dot clustering operator that can compensate for the weakness of the conventional connected components operator. In the morphological routine, the connectivity among pixels does not depend on a direct adjacency criterion. Instead, it is able to group together some separate dots whose connection closes up a hole in an organ, while cutting apart some narrow connections that would be acceptable to the connected components operator. In our system, if there is too narrow a range of thresholds that generate candidate regions, the system invokes the morphological routine with the binary image obtained from the lowest threshold that did not produce too large a region. A range of such thresholds is considered too narrow if it is less than the preset width for each organ. The preset widths reflects the heights of the gray tone cliffs of the organs.

The morphological routine first connects tiny regions that are separate, but close enough to each other, by a closing operation, since closely-situated disconnected

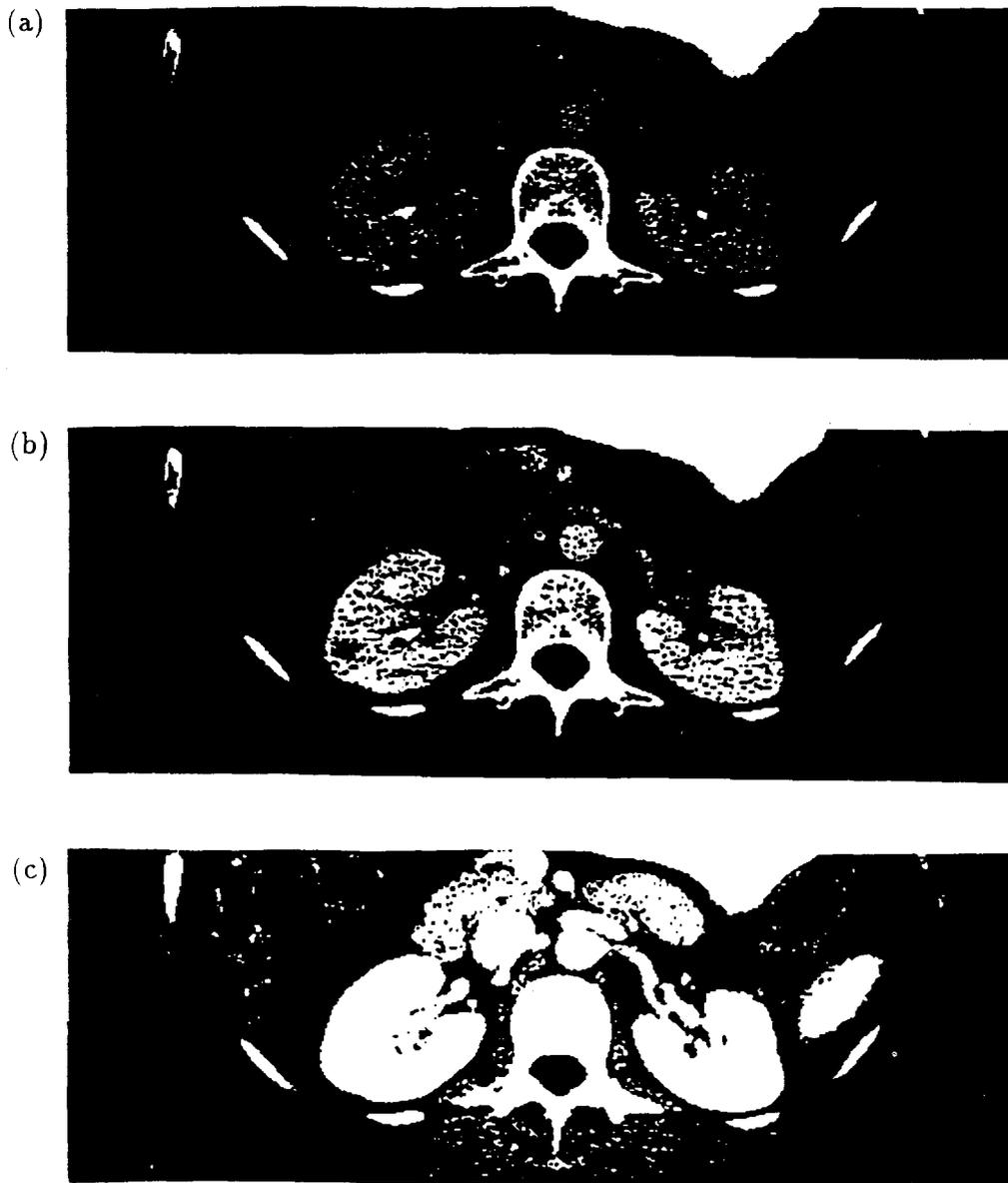


Fig. 3. Dynamic thresholding process. (a) At the initial (highest) threshold for kidneys, the kidneys show up only as a set of small, disconnected regions. (b) At 3 steps (gray tone width 30) lower than the initial threshold, the kidneys become large enough to be in the range of acceptable size. (c) At 11 steps (gray tone width 110) lower than the initial threshold, both kidneys connect with other organs/tissues and violate the negative constraints.

regions that belong to the same organ tend to come about when the threshold range is too narrow. Then it separates regions that are connected by narrow unwanted bridges using an opening operation. More stress is placed on cutting narrow connections than on connecting close separate regions, based on the heuristic, which seems effective in processing CT images, that the size of the structuring element used for the opening is greater than that used for the closing. It performs the connected components operation on the resulting image and checks all the properties of each region. The output for most organs is the largest region

that satisfies all the constraints. For some organs adjacency to another organ overrules the area size in the final decision. The following algorithm describes the basic procedure followed by the system.

*Step 1.* Set the initial threshold to the high end of the relevant gray tone range predefined for the organ of interest. From the second iteration on reduce the threshold by a constant value (10 was used in our experiments). If the threshold reaches the low end of the relevant gray tone range and if there is no candidate, then invoke the morphological routine (Step 11) for

the binary image computed by thresholding with the lowest threshold that keeps all the regions within the target area of the image smaller than the maximum acceptable size.

*Step 2.* Execute the thresholding operation on the given image with the current threshold.

*Step 3.* Execute the connected components operation on the binary image produced by Step 2 to produce a set of regions. Figure 3a shows (in white) the connected components found at the initial threshold for kidney extraction.

*Step 4.* (Area check) Check if there is a region of acceptable size within the search area that was pre-defined for the organ of interest. If there are some, record them as candidates. If there is none, go back to Step 1. In Fig. 3a there are no regions of acceptable size in the kidney search area. Three iterations later, as shown in Fig. 3b the kidney regions are large enough to become candidates.

*Step 5.* (Location check) Check if there is among the candidates a region that satisfies the location condition defined by the position of the gravity center. If there are some, keep them in the record of candidates. If there is none, go to Step 1. If the candidate object is the left kidney, then the left kidney candidate in the image should satisfy the location check, while the right kidney and spine should not.

*Step 6.* (Positive/Negative shape check) Check if there is among the candidates a region that satisfies the positive/negative shape conditions. These conditions are expressed by the properties such as spatial extension (an extension to an extremely abnormal region of the body-based coordinate system, violating the negative shape constraints), vertical and horizontal lengths, and their ratio. If there are some, keep them in the record of the candidates. If there is none, go to Step 1. In Fig. 3b, the kidney candidates satisfy the shape check, but in Fig. 3c, which illustrates the eleventh iteration, the kidney candidates have merged in with other organs and violate the abnormal extension constraint.

*Step 7.* (Overlap check) Check if there is among the candidates a region that satisfies the overlap condition with an adjacent slice that is already correctly segmented. The overlap condition is expressed as the minimum required overlap ratio which is 50% of the smaller region of the overlapping pair. If there are some, keep them in the record of the candidates. If there is none, go to Step 1.

*Step 8.* (Collision check) Check if there is among the candidates a region that does not collide with the other already recognized organs in the same slice. If there are some, keep them in the record of the candidates. If there is none, go to Step 1. Neither of the two kidney candidates in Fig. 3b collide with anything else.

*Step 9.* Choose the best candidate from the candidates at the current threshold, record its area, and go

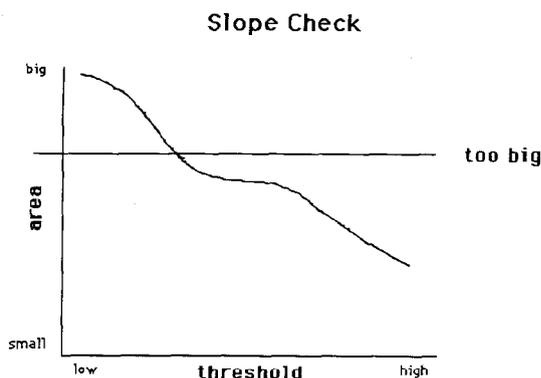


Fig. 4. The slope check test.

to Step 1. For most organs, best means most suitable in area. For some organs, however, adjacency to landmarks is the criterion used. Depending on whether the left or right kidney was sought, one of the two kidney candidates in Fig. 3b will be selected as the best candidate for the third interaction.

*Step 10.* (Slope check) This step checks the record of the change of the area size of the candidates with respect to the threshold change. Suppose the area of the candidate is plotted along the vertical axis and the corresponding threshold along the horizontal axis, as shown in Fig. 4. This step looks for the flattest part of the curve that gives an acceptable area for the organ in question and chooses the midpoint of the flat part as the correct threshold. The candidate at that threshold is output as the region of the organ of interest. However, if the slope of the flattest range is not within the preset acceptable range for the organ of interest, the system sends the binary image made by that middle-level threshold to the morphological routine (Step 11).

*Step 11.* Execute the following morphological operations.

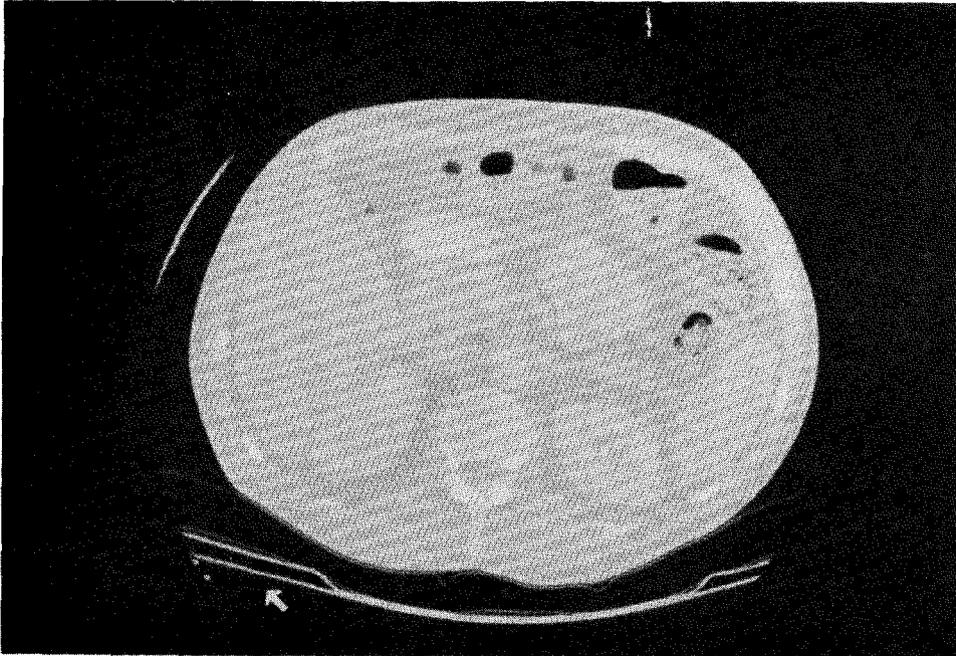
- (1) Close with a disk of diameter 3 pixels.
- (2) Open with a disk of diameter 5 pixels on the above result.
- (3) Execute connected component operation on the above result.
- (4) Extract the region that satisfies the conditions in Steps 4, 5, 6 and 7.
- (5) Close the extracted region alone with a disk of diameter 3. The result is output as the region of the organ of interest.

The extraction of the liver and the spleen is more difficult than that of the other organs, since their gray tone cliffs are low and their gray tones are very similar to adjacent tissue. For this reason, we have added some special additional processing for these two organs that takes advantage of the merits of the progressive landmarking and the stable ordering of the gray tones of some organs/tissues. If the regions of other organs/tissues having higher gray tones than that of the target organ are eliminated from the image, the segmentation of the target organ becomes easier by the thresholding

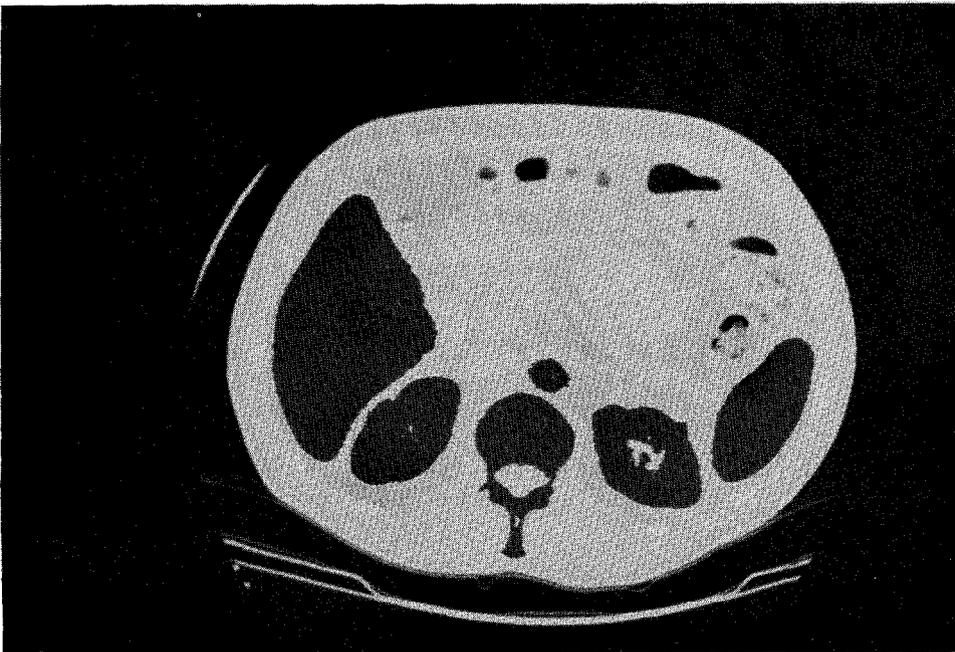
process described above, since there is less danger of connecting a candidate region with other organs/tissues. Since the liver and the spleen have lower gray tones than bones and many other organs/tissues, the elimination of higher gray tone regions makes the isolation of these organs from other organs/tissues easier.

#### 6. PERFORMANCE OF THE SYSTEM

One peculiarity of this task domain is that there is no precise ground rule of truth. Although most professional dosimetrists produce very similar results, there can be considerable differences among them in drawing very ambiguous contours of complex regions. It is



(a)

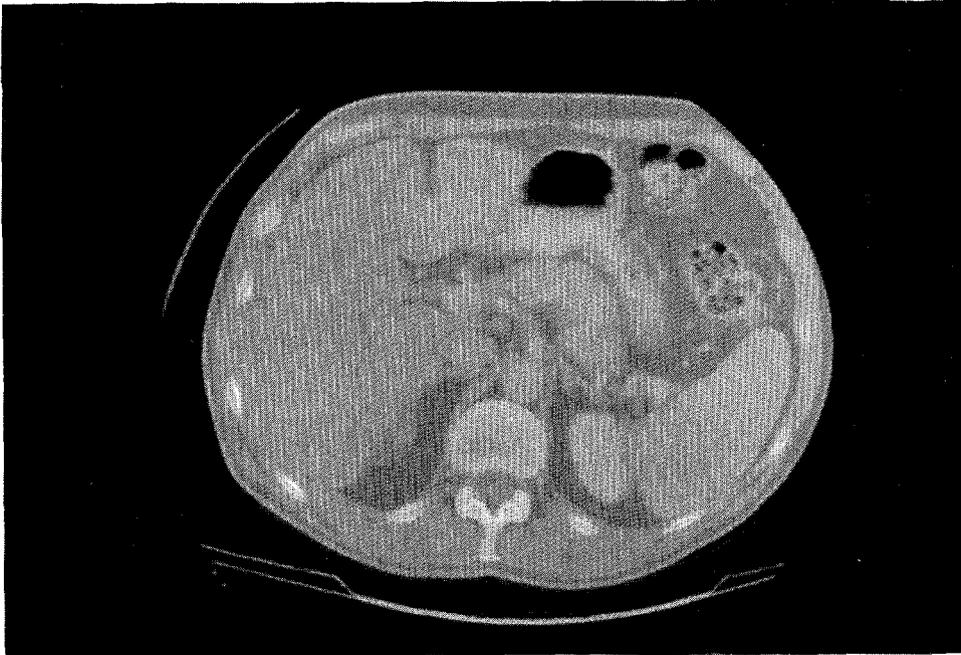


(b)

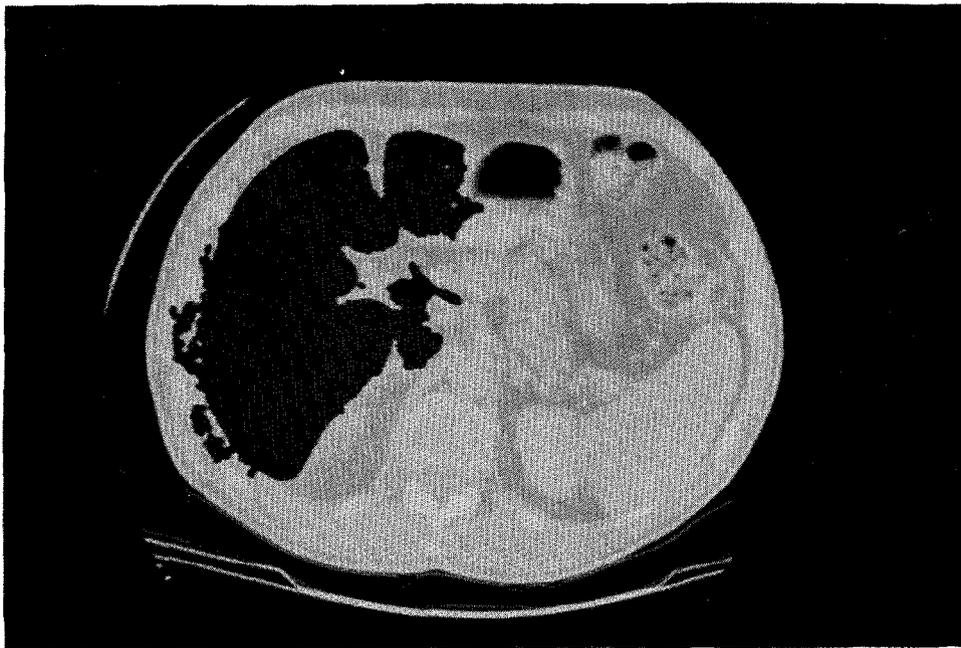
Fig. 5. Grade A extraction of all organs. (a) Original image. (b) Homogeneous dark regions are kidneys, liver (leftmost), spleen (rightmost), aorta (round one in the center) and a part of spine.

not unusual that differences between the results of dosimetrists can extend to around a 5 pixel width in an image of size  $512 \times 512$ . In difficult cases the mismatch can be more than 10 pixels wide, and a whole region of substantial area may be included or discarded depending on the individual.

In this work, the evaluation was done by the author, but with an attempt to be fairly conservative. Although the decision on some parts of the liver can be controversial in some instances, the contours of the kidney are in all instances very clear and unambiguous. The outputs of the system are evaluated against the correct



(a)

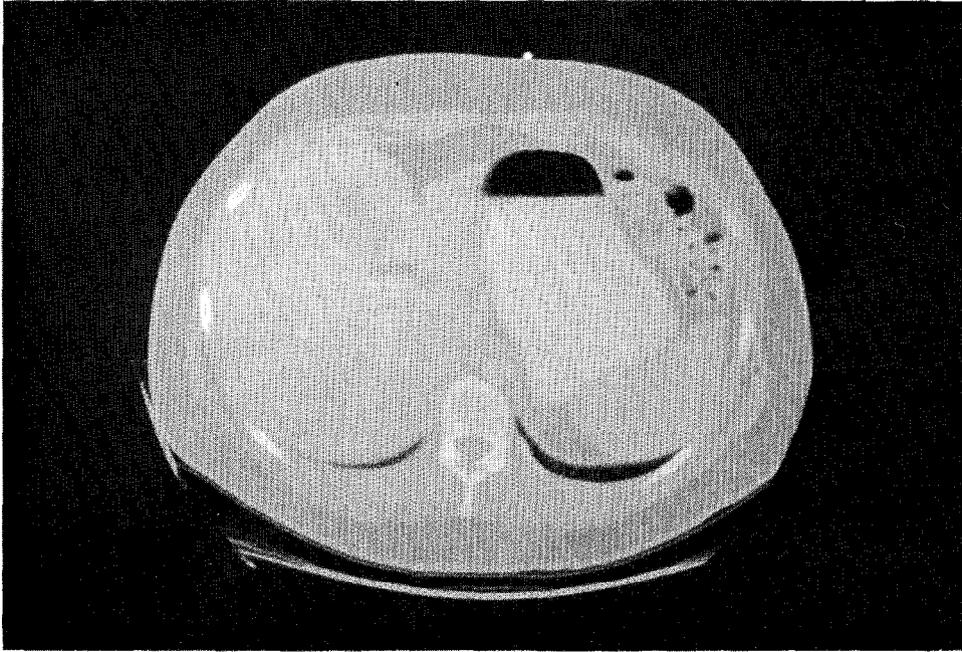


(b)

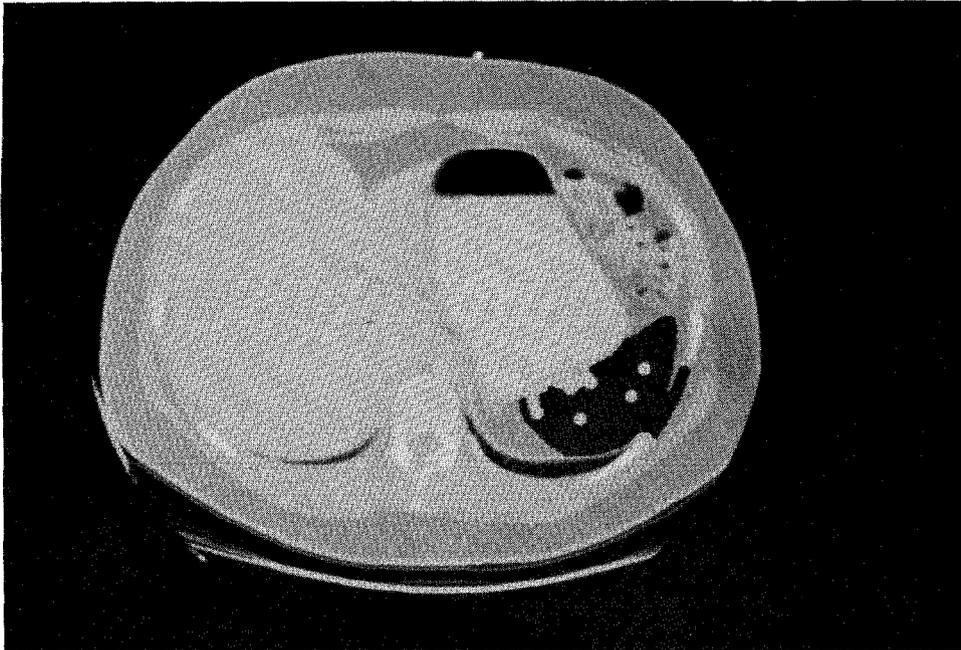
Fig. 6. Grade B extraction of liver. (a) Original image. (b) Homogeneous dark region is the extracted liver.

result indicated by the author. Evaluation was done on the accuracy of the outermost contours only. Some holes inside organs are ignored, since the objective of this system is to provide correct outermost contours of the organs. Inner holes can be, if necessary, filled by a hole-filling algorithm.

After checking for mismatch, each extraction result is graded as A, B or C. The grading is not done by objective measurements but by incorporating qualitative assessments. Qualitative considerations are necessary since, for example, misidentifying the aorta as a part of the liver, which makes the mismatch area 1%



(a)



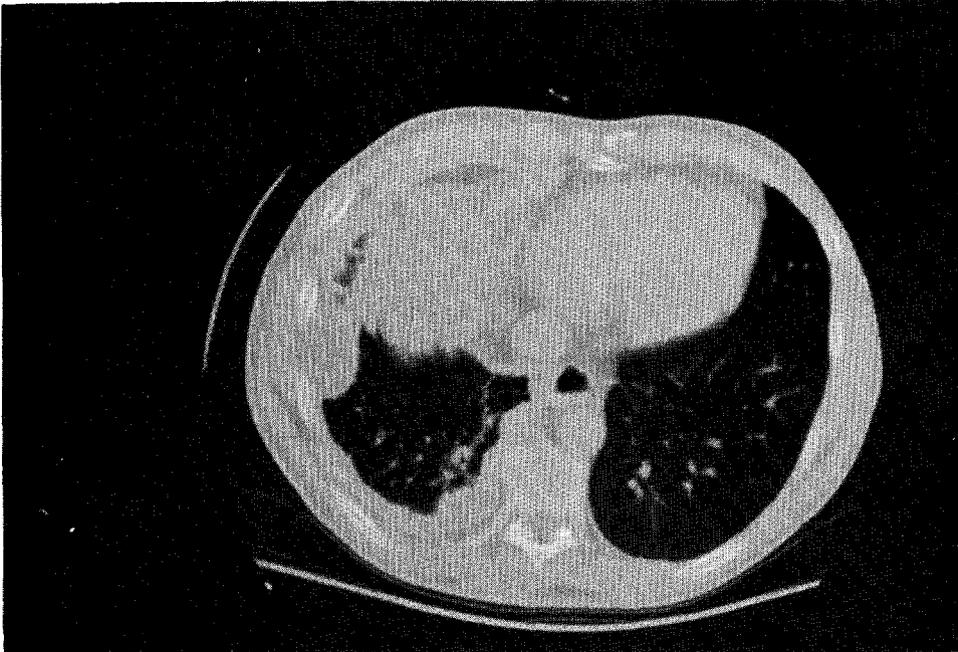
(b)

Fig. 7. Grade B extraction of spleen. (a) Original image. (b) Homogeneous dark region is the extracted spleen.

of the correctly identified region is worse than having thinly distributed contour mismatch amounting to 1.5% of the correctly identified region. Therefore, the criteria for grade A include the judgement of whether or not the identified region includes a part of other organs (excluding non-organ tissues like surrounding fat). The rules for grading a result are as follows.

(A) Comparable to human dosimetry work within 5 pixel wide mismatch regions in  $512 \times 512$  images. If there are more than 5 pixel wide mismatches, the area of such mismatch is less than 1% of the total area and the mismatch is of benign quality (i.e., not an inclusion of a part of a wrong organ).

(B) Worse than A, but at least 70% of the correct



(a)



(b)

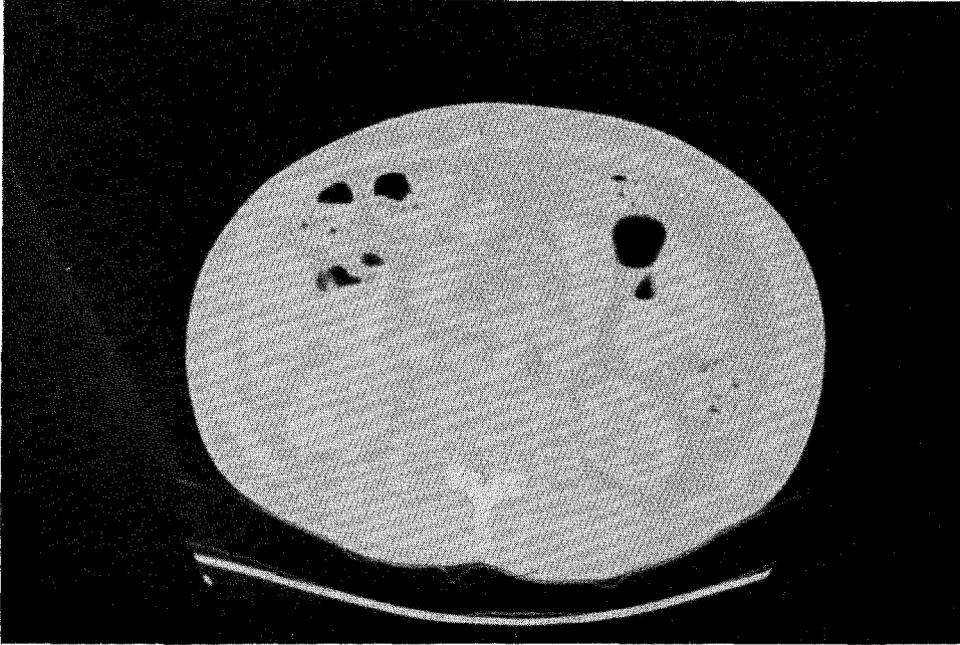
Fig. 8. Grade C extraction of liver. (a) Original image. (b) Homogeneous dark region is the extracted liver.

region of the organ is detected and the area of other tissues recognized as parts of the target organ is less than 30% of the correct area of the target organ.

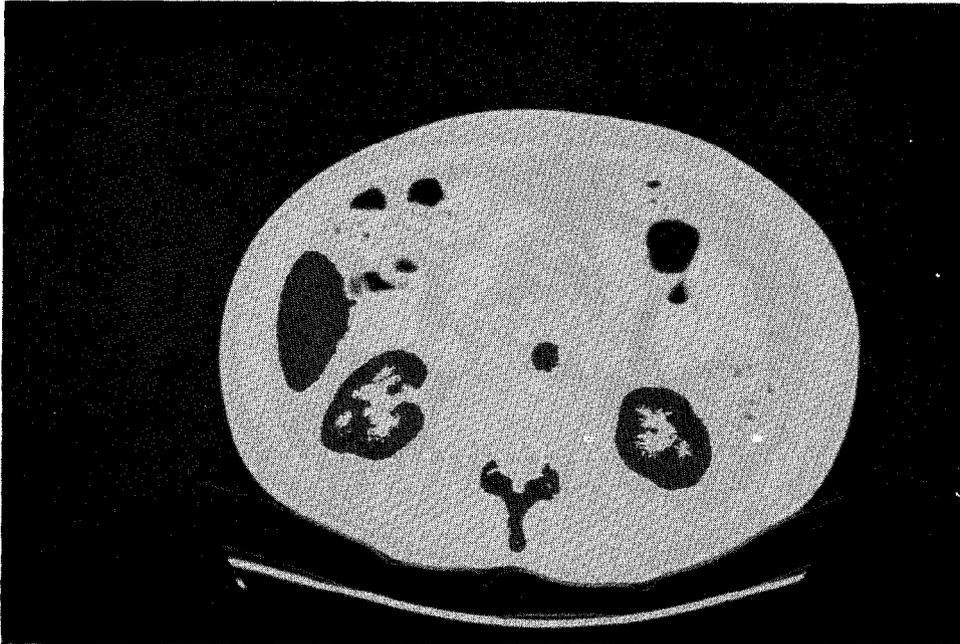
(C) Worse than B, less than 70% of the correct region of the organ is detected or the area of other regions included as parts of the target organ is more than 30% of the correct area of the target organ.

The system was first tested on the 113 training images of the abdomen from five patients. Our results were:

Kidneys: grade A—94%, grade B—3%, grade C—3%,  
 Spleen: grade A—62%, grade B—30%, grade C—8%,  
 Liver: grade A—60%, grade B—20%, grade C—20%.



(a)



(b)

Fig. 9. Extraction result of a low level slice. (a) Original image. (b) Homogeneous dark regions are the extracted organs.

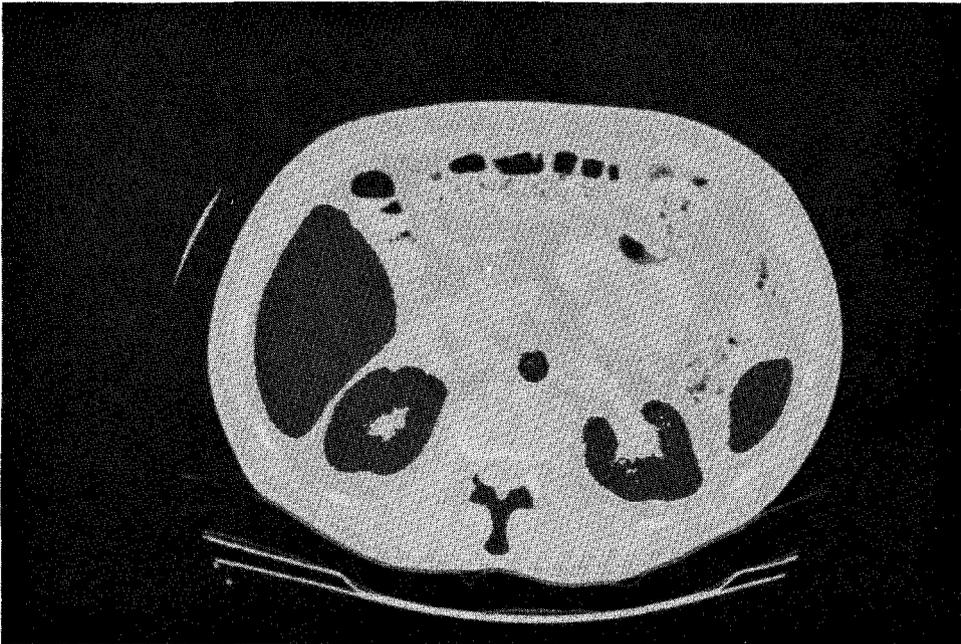
The system shows very good performance in extracting kidneys while its success rate is substantially lower with liver and spleen segmentations, especially in the high-level slices of the liver. Examples of the graded results are shown in Figs 5–8 along with the original images.

The kidneys can be correctly segmented in most cases by the dynamic thresholding technique alone, as

they tend to have high contrast of gray tone against their adjacent region. This high contrast reflects the tendency that the kidneys are covered with fat, which has very low gray tones. On the other hand the upper part of the liver tightly contacts tissues of gray tones very close to itself. Therefore, many of those images with which the system failed are difficult to analyse even for humans.



(a)



(b)

Fig. 10. Extraction result of a mid level slice. (a) Original image. This slice is five slices (about 5 cm) higher than the slice of Fig. 9. (b) Homogeneous dark regions are the extracted organs.

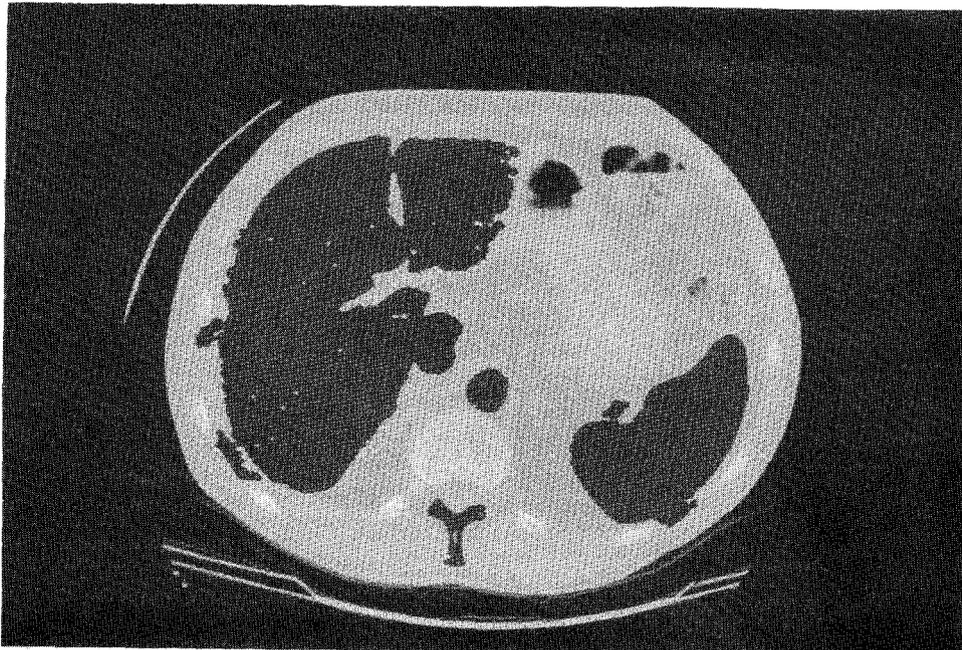
The system failed in only one slice out of 113 to locate the spine correctly (success rate 90%). For the aorta, it failed in 17 slices out of 113 (success rate 85%). However, for the spine and the aorta, a complete failure can be compensated for by using data from the nearest slice with successful extraction. This is because the column index of the spine's gravity center and the

row index of the aorta's gravity center do not vary much between slices, and providing these figures is their major role in this system. Figs 9-12 show the extraction results of four slices of a patient.

The system was next tested on 75 more images of the abdomen from three patients, that were not a part of the original training set. This was a particularly



(a)



(b)

Fig. 11. Extraction result of a mid-high level slice. (a) Original image. This slice is five slices (about 5 cm) higher than the slice of Fig. 10. (b) Homogeneous dark regions are the extracted organs.

stringent test, because many of these images were of rather poor quality. Our results on these images were:

Kidneys: grade A—85%, grade B—0%, grade C—15%,

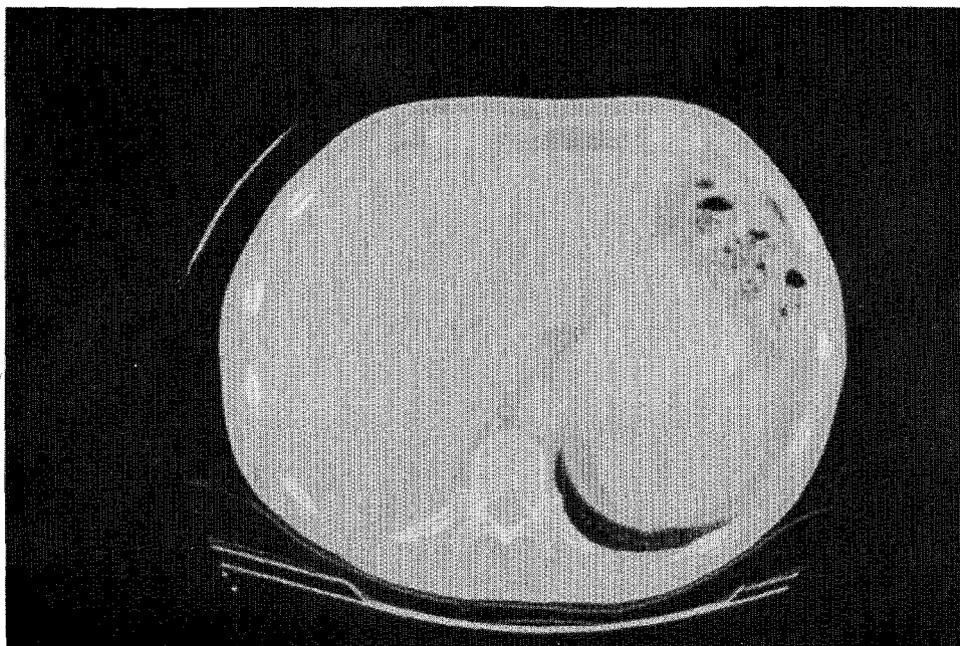
Spleen: grade A—70%, grade B—6%, grade C—23%,

Liver: grade A—52%, grade B—31%, grade C—17%.

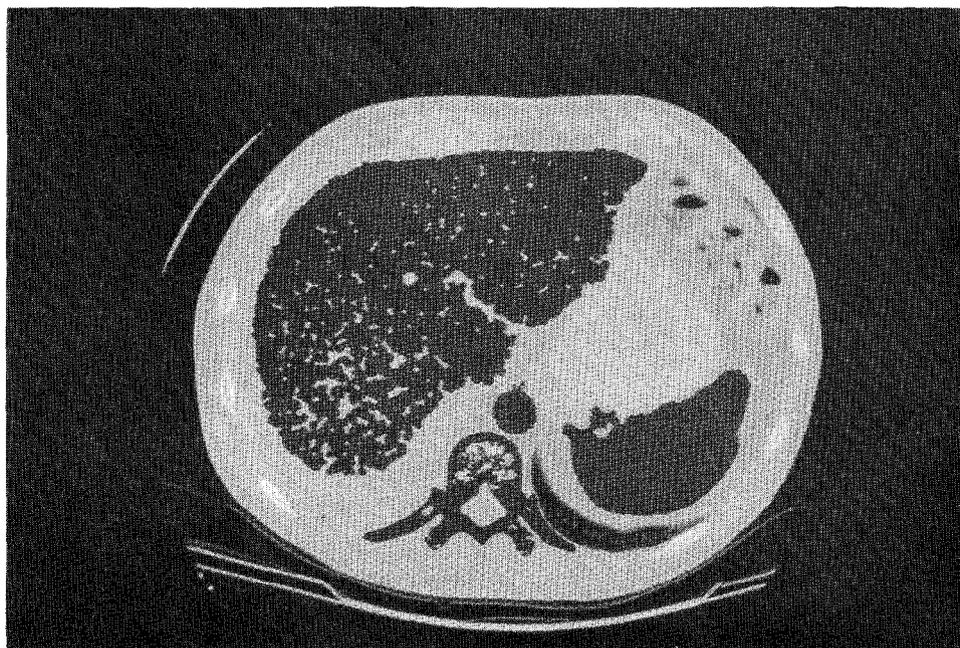
Most of the detections that failed were due to the poor quality of the images. The gray levels of the non-

extractable organs were very close to the gray levels of the surrounding tissues. Figure 13 shows examples of some of the poor quality images that failed in the test set.

We believe that the overall performance of the system is very good. The two principles of modeling and flexible control make it possible to achieve a difficult combination of powerful discriminatory capability



(a)

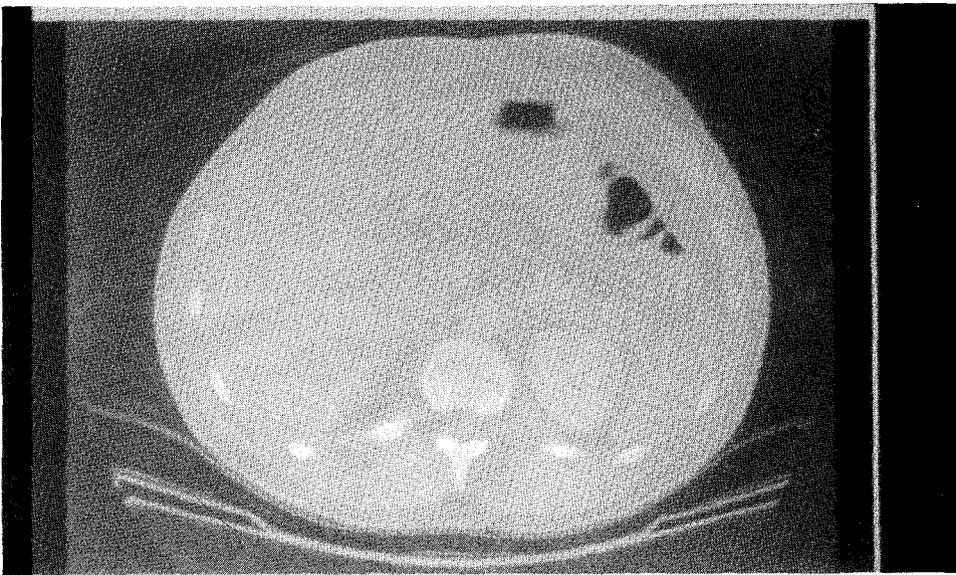


(b)

Fig. 12. Extraction result of a high level slice. (a) Original image. This slice is five slices (about 5 cm) higher than the slice of Fig. 11. (b) Homogeneous dark regions are the extracted organs.



(a)



(b)

Fig. 13. Several poor quality images from the test set.

and tolerance to wide variation of instances, both of which are required in this task domain. However, the system exhibits some weakness, especially in making distinctions between different adjacent tissues of very close gray tones. This weakness is caused by lack of shape knowledge with which to make the educated guesses that human dosimetrists use. The use of several different modes of data might be useful here. Despite its weaknesses, the basic approach of the current system is very effective as a first approximation to successfully locating the core region of an organ of interest.

#### REFERENCES

1. C. L. Morgan, *Basic Principles of Computed Tomography*, University Park Press, Baltimore (1983).
2. J. K. T. Lee *et al.*, *Computed Body Tomography with MRI Correlation*, Raven Press, New York (1989).
3. C. Harlow and S. A. Eisenbeis, The analysis of radiographic images, *IEEE Trans. Computers* **C-22**, 678–688 (1973).
4. P. G. Selfridge, *Reasoning about success and failure in aerial image understanding*, Ph.D. thesis, University of Rochester (1982).
5. P. G. Selfridge and J. M. S. Prewitt, Organ Detection in

- Abdominal Computerized Tomography Scans: Application to the Kidney, *Comput. Graphics Image Process.* **15**, 265–278, (1981).
6. Y. Shirai, A flexible method of extraction of approximate stomach region from radiograms, *Proc. Sixth Int. Conf. Pattern Recognition* pp. 906–908 (1982).
  7. N. Karssemeijer, L. J. von Esking and E. G. Eijkman, Recognition of organs in CT-image sequences: a model guided approach, *Comput. Biomed. Res.* **21**, 434–448 (1988).
  8. D. S. Bright, An object-finder for digital images based on multiple thresholds, connectivity and internal structure, *J. Comput.-Assisted Microsc.* **1**, 307–329 (1989).
  9. J. F. Brinkley, Representing biological objects as geometric constraint networks, *Proc. AAAI Spring Symp. Series: Artifi. Intell. Med.* Stanford University, pp. 7–8 (1988).
  10. S. Y. Chen, W. C. Liu and C. T. Chen, An expert vision system for medical image segmentation, *SPIE, Medical Imaging III: Image Process.* **1092**, 162–172 (1989).
  11. S. Shemlon and S. Dunn, Rule-based image interpretation with models of expected structure, *SPIE Medical Imaging IV: Image Process.* **1233**, 33–44 (1990).
  12. H.-H. Ehrlicke, Problems and approaches for tissue segmentation in 3D-MR imaging, *SPIE Medical Imaging IV: Image Process.* **1233**, 128–137 (1990).

**About the Author**—MASAHARU KOBASHI was born in Tsuyama, Japan, in 1952. He earned a B.A. in economics in 1975 and a second B.A. in business in 1976, both from The University of Tokyo. In 1981 he earned an MBA from Stanford University. In 1990, he entered the graduate program in computer science at the University of Washington, where he earned the M.S. degree in 1992. His research interests are: artificial intelligence, computer vision and knowledge representation in general.

**About the Author**—LINDA G. SHAPIRO was born in Chicago, Illinois, 1949. She received the B.S. degree in mathematics from the University of Illinois, Urbana, in 1970, and the M.S. and Ph.D. degrees in computer science from the University of Iowa, Iowa City, in 1972 and 1974, respectively.

She was an Assistant Professor of Computer Science at Kansas State University, Manhattan, from 1974 to 1978 and was an Assistant Professor of Computer Science from 1979 to 1981 and Associate Professor of Computer Science from 1981 to 1984 at Virginia Polytechnic Institute and State University, Blacksburg. She was Director of Intelligent Systems at Machine Vision International in Ann Arbor from 1984 to 1986. She is currently Professor of Computer Science and Engineering and of Electrical Engineering at the University of Washington. Her research interests include computer vision, artificial intelligence, pattern recognition, robotics and spatial database systems. She has co-authored two textbooks, one on data structures and one on computer and robot vision.

Dr Shapiro is a senior member of the IEEE Computer Society and a member of the Association for Computing Machinery, the Pattern Recognition Society and the American Association for Artificial Intelligence. She is past Editor of *CVGIP: Image Understanding* and is currently an editorial board member of *IEEE Transactions on Pattern Analysis and Machine Intelligence* and of *Pattern Recognition*. She was co-Program Chairman of the IEEE Conference on Computer Vision and Pattern Recognition in 1994, General Chairman of the IEEE Workshop on Directions in Automated CAD-Based Vision in 1991, General Chairman of the IEEE Conference on Computer Vision and Pattern Recognition in 1986, General Chairman of the IEEE Computer Vision Workshop in 1985, and Co-Program Chairman of the IEEE Computer Vision Workshop in 1982; and she has served on the program committees of a number of vision and AI workshops and conferences.